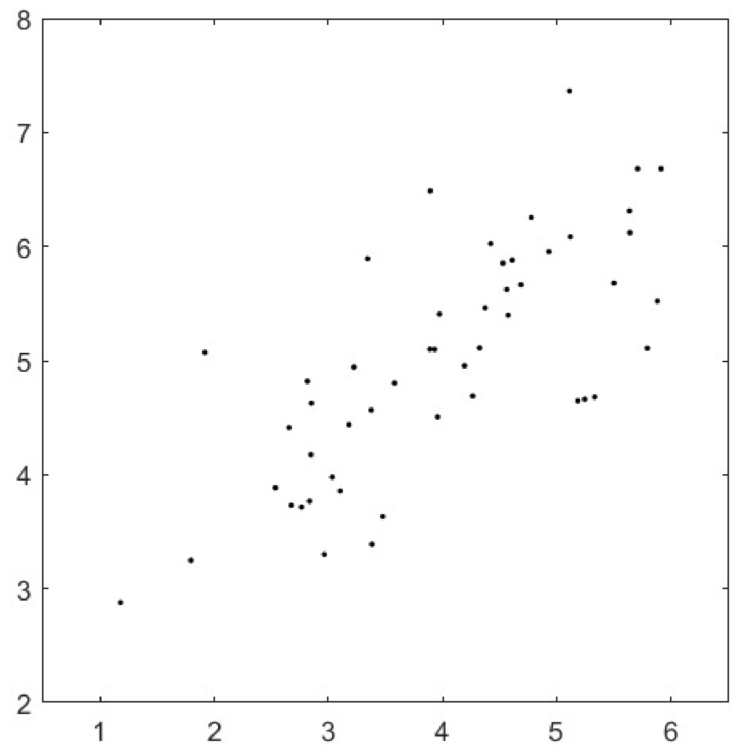
ΟΝΟΜΑΤΕΠΩΝΥΜΟ : Ιωάννης-Ιάσων Γεωργακάς
ΑΜ: 2017030021

Στατιστική Μοντελοποιήση και Αναγνώριση προτύπων (ΤΗΛ 311)
Αναφορά 1ου Σέτ Ασκήσεων

Θέμα 1ο : Principal Component Analysis (PCA)

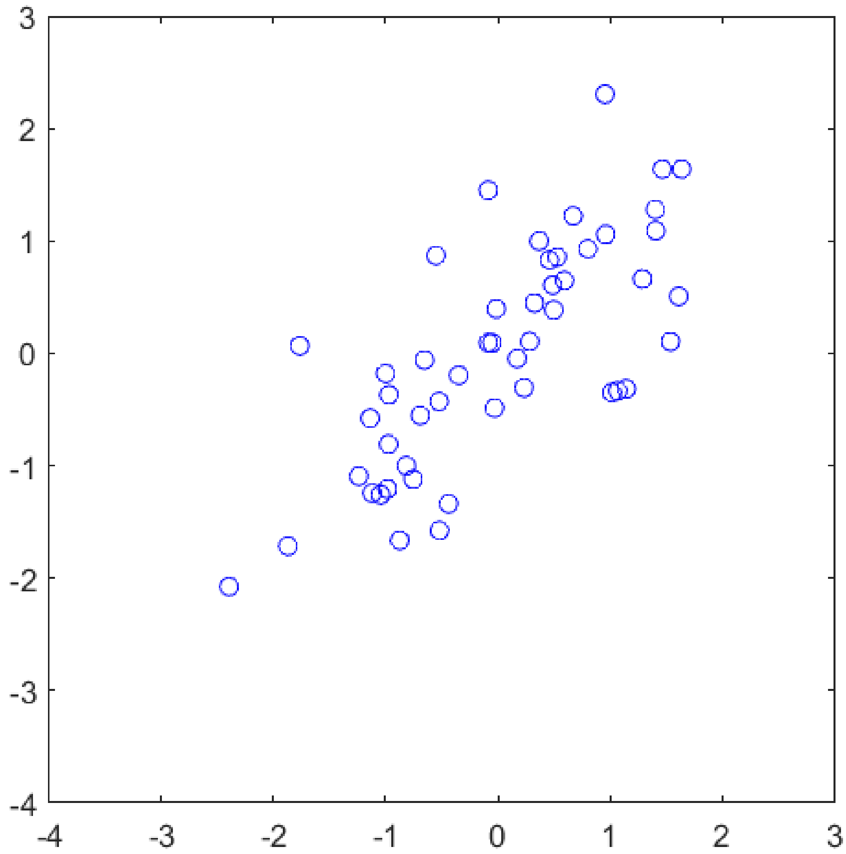
Στην 1η άσκηση εφαρμόστηκε η unsupervised μέθοδος PCA η οποία αναπαριστά τα N-διαστατά δεδομένα σε ένα χώρο λιγότερων διαστάσεων. Ο στόχος της χρήσης της συγκεκριμένης μεθόδου είναι να γίνει κατανοητή η βασική αρχή λειτουργίας της η οποία είναι να μειωθούν οι βαθμοί ελευθερίας μειώνοντας έτσι τις χωρικές και χρονικές πολυπλοκότητες.

<u>Μέρος 1ο:</u>
Άρχικά γίνεται φόρτωση των δεδομένων μέσω της εντολης load και η απεικόνιση τους στον δισδιάστατο χώρο (2D).



Σχήμα 1α : Τα αρχικά δείγματα

Στη συνέχεια υλοποιώντας την συνάρτηση featureNormalize πραγματοποιείται η προεπεξεργασία των δεδομένων κανονικοποιώντας τα δείγματα με μέση τιμή μηδέν και διασπορά ίση με 1 ώστε δεδομένα διαφορετικών κλιμάκων να μην οδηγήσουν σε παραπλανητικά στοιχεία και να συνεισφέρουν ισότιμα στην ανάλυση . Στο σχήμα 1β φαίνονται τα κανονικοποιημένα δεδομένα, συγκρίνοντας τα με τα αρχικά δείγματα, παρατηρείται ότι τα δεδομένα απεικονίζονται σε μικρότερη κλίμακα από ότι στην αρχή.

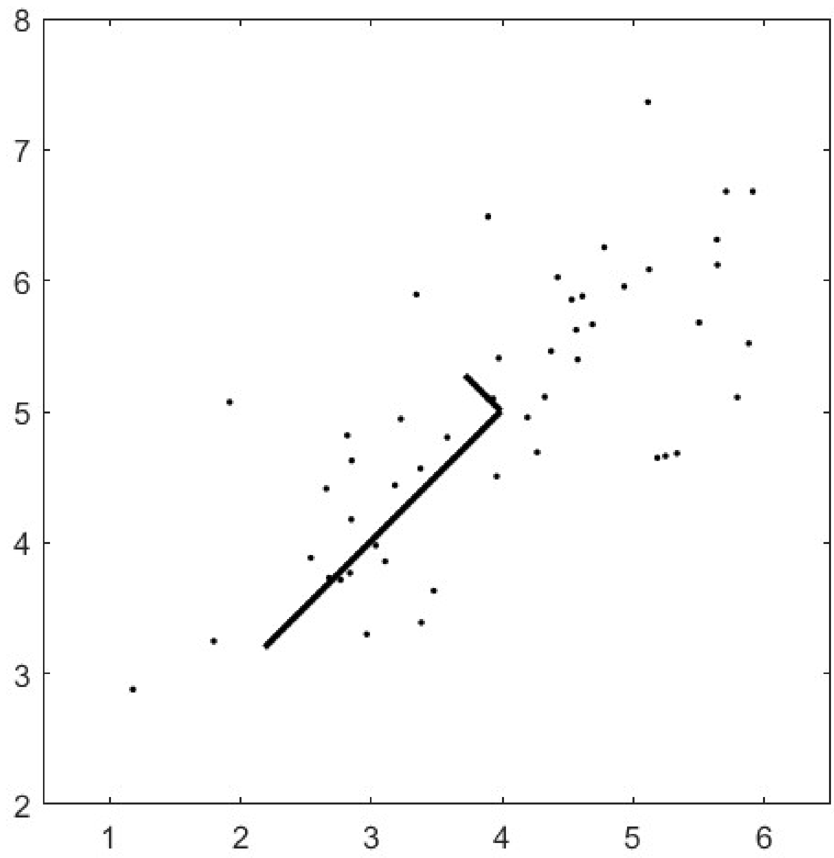


Σχήμα 1β : Κανονικοποιήση των δεδομένων

Κατόπιν υλοποιώντας τον αλγόριθμο PCA υπολογίστικαν οι κύριες συνιστώσες του αλγορίθμου ώς τα ιδιοδιανύσματα με χρήση της μεθόδου singular value decomposition. Κατά την υλοποιήση του αλγοριθμού PCA χρησιμοποιήθηκε ο παρακάτω τύπος για τον υπολογισμό του πίνακα συνδιασποράς Σ :

$\Sigma = \frac{1}{m} X^T X$, όπου $X \in \mathbb{R}^{mxn}$ είναι ένας πίνακας όπου κάθε γραμμή αντιστοιχεί σε ένα παράδειγμα με n χαρακτηριστικά.

Στο σχήμα 2 απεικονίζονται οι κύριες συνιστώσες του αλγορίθμου ως οι κατευθύνσεις των δεδομένων οι οποίες αντιπροσωπεύουν το μέγιστο ποσό διακύμανσης δηλαδή οι γραμμές οι οποίες συλλέγουν τις περισσότερες πληροφορίες των δεδομένων.



Σχήμα 2: Οι κύριες συνιστώσες των δεδομένων

Ακολούθως υπολογίζεται η συνεισφορά που έχει κάθε κύρια συνιστώσα στην συνολική διακύμανση μέσω του τύπου :

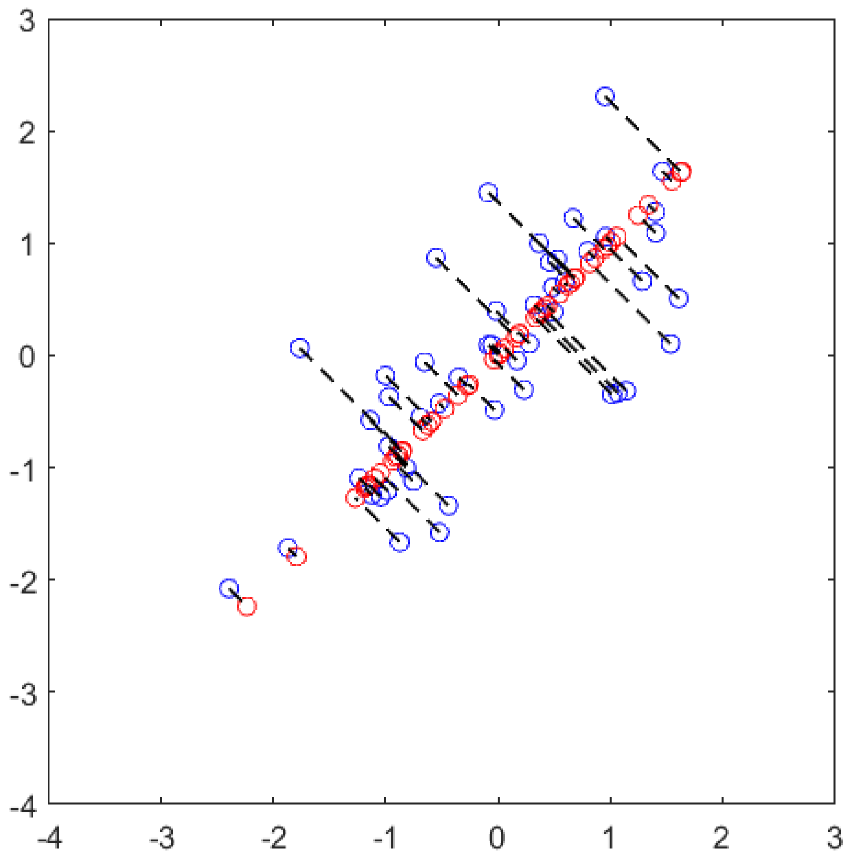
$PC_{variance} = \frac{eigenVal}{\mathrm{sum}(eigenVal)}$, δηλαδή η συνεισφόρα κάθε συνιστώσας είναι ίση με το πηλικό της ιδιοτιμής κάθε συνιστώσας προς το συνολικό άθροισμα των ιδιοτιμών όλων των συνιστωσών.

Επιπλέον μειώνεται η διάσταση των δεδομένων από 2D σε 1D στα αρχικά δείγματα υλοποιώντας τη συνάρτηση projectData η οποία χρησιμοποιώντας έναν γραμμικό μετασχηματισμό προβάλλει τα δεδομένα σε λιγότερες διαστάσεις. Ο τύπος του γραμμικού μετασχηματισμού που χρησιμοποιήθηκε είναι :

$Z_i = U^T X_i$ όπου X_i είναι κάθε δείγμα.

Στη συνέχεια υλοποιώντας τη συνάρτησης recoverData πραγματοποιείται η ανάκτηση των αρχικών δειγμάτων προσεγγιστικά αναπαράγοντάς τα ξανά στον αρχικό χώρο υψηλών διαστάσεων(2D). Αυτό γίνεται με την προβολή τους πάνω σε όλες τις κύριες συνιστώσες (principal components) με τον ακόλουθο μετασχηματισμό και φαίνεται στο σχήμα 3:

$X_{rec} = U \cdot z$



Σχήμα 3 Τα αρχικά δείγματα προεπεξεργασμένα (μπλέ κύκλοι) και τα δείγματα ανακτόμενα (κόκκινοι κύκλοι)

<u>Μέρος 2ο:</u>

Γενικεύωντας την μέθοδο PCA ,χρησιμοποιήθηκαν 5000 εικόνες προσώπων από το αρχείο 'ex1faces.mat' με σκοπό να εφαρμοστούν οι προαναφερθείσες συναρτήσεις σε ένα γενικευμένο dataset.

Αρχικά σχεδιάζονται οι 100 πρώτες εικόνες από το dataset με τη χρήση της συνάρτησης displayData() και πραγματοποιείται προεπεξεργασία των δεδομένων εφαρμόζοντας κανονικοποιήση με μέση τιμή μηδέν και διασπορά ίση με 1 χρησιμοποιώντας τη συνάρτηση featureNormalize().

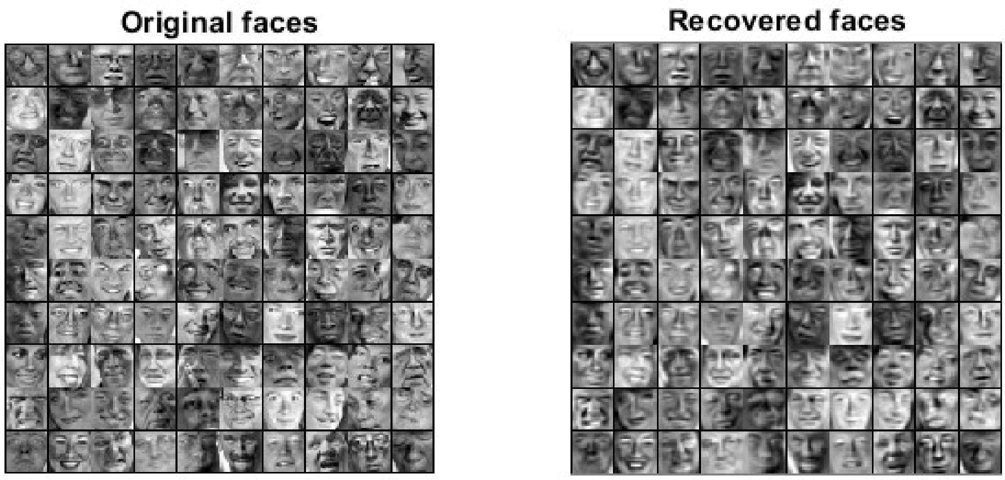
Σχήμα 4: Οι 100 πρώτες εικόνες

Στη συνεχεία με χρήση της συνάρτησης mypca() υπολογίζονται οι κύριες συνιστώσες των δεδομένων και απεικονίζονται στο σχήμα 5 οι 36 κύριες συνιστώσες χρησιμοποιώντας της συνάρτησης displayData().

Σχήμα 5: Οι 36 κύριες συνιστώσες

Κατόπιν εφαρμόζοντας την συνάρτηση projectData() μειώνονται η διάσταση των δειγμάτων χρησιμοποιώντας τις 100 πρώτες κύριες συνιστώσες.

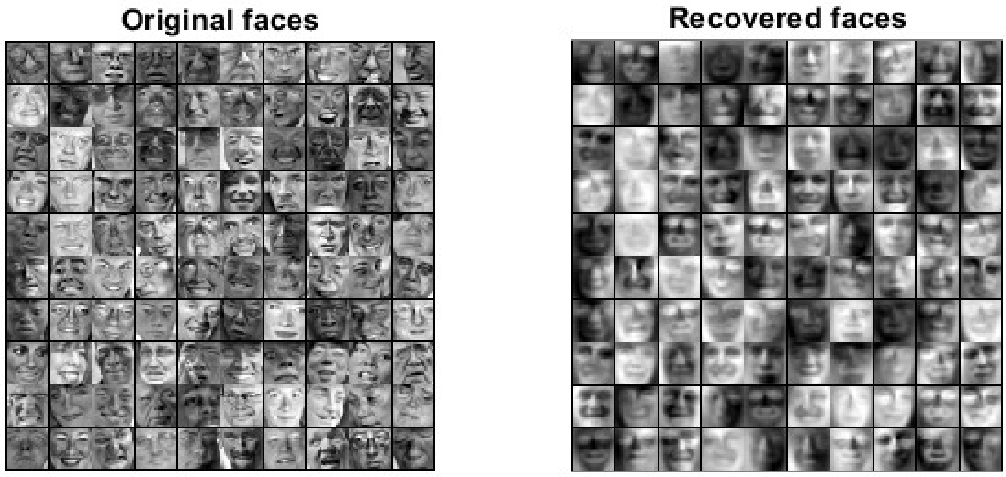
Τέλος αναπαριστόνται στο σχήμα 6 τα δείγματα μειωμένης διάστασης αφού προηγουμένως έχουν προβληθεί στον αρχικό χώρο χρησιμοποιώντας τη συνάρτηση recoverData().



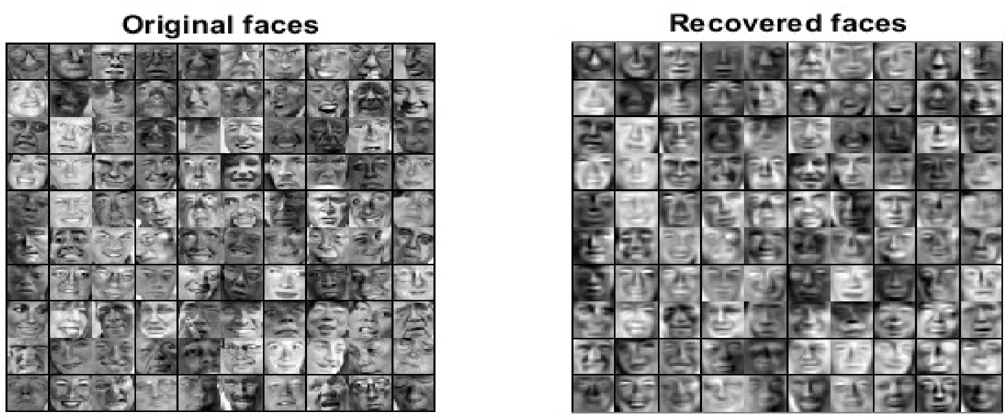
Σχήμα 6: Οι αρχικές εικόνες σε αντιδιαστολή με τις ανακτημένες εικόνες

Στο σχήμα 6 παρατηρείται ότι η ανάκτηση των δεδομένων έχει επιτευχθεί καθώς η μέθοδος PCA χρησιμοποιεί για την προβολή και ανάκτηση των δεδομένων τις κύριες συνιστώσες οι οποίες περιέχουν τη περισσότερη πληροφορία σε όσον το δυνατόν λιγότερα ιδιοδιανύσματα.

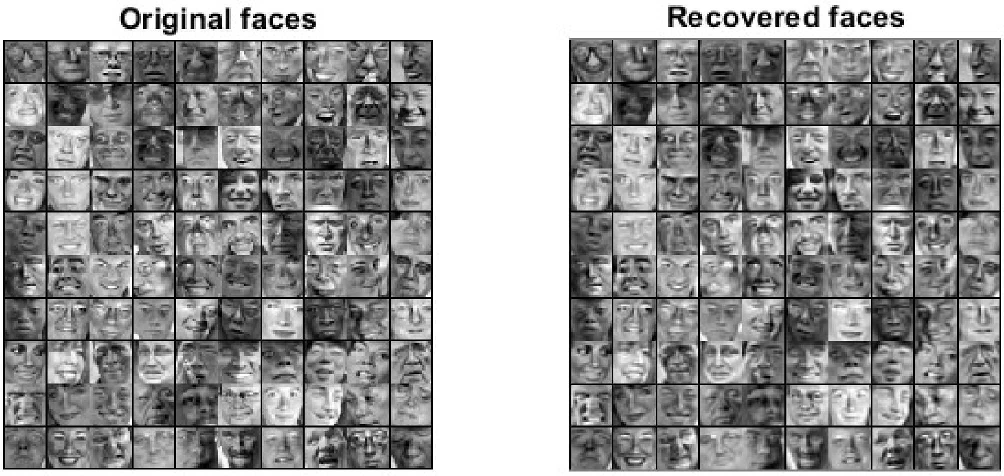
Επαναλαμβάνοντας την παραπάνω διαδικασία για διαφορετικό αριθμό από κύριες συνιστώσες , συγκεκριμένα για K = 10, 50 ,200 σχεδιάζονται τα ανακτημένα πρόσωπα σε αντιπαραβολή με τα αρχικά. Παρατηρείται ότι χρησιμοποιώντας μεγαλύτερο αριθμό κύριων συνιστωσών η λεπτομέρεια των προσώπων αυξάνεται αφού παρέχεται περισσότερη πληροφορία μέσα απο τα ιδιοδιανύσματα.



Σχήμα 7 : Κύριες συνιστώσες K = 10



Σχήμα 8: Κύριες συνιστώσες K = 50



Σχήμα 9 : Κύριες συνιστώσες K = 200

Θέμα 2ο : Ταξινομητής LDA (Linear Discriminant Analysis)

Είναι γνωστό ότι οι εκ των προτέρων(a priori) πιθανότητες των δύο κλάσεων είνα ίσες δηλαδή ισχύει ότι $P(\omega_1)=P(\omega_2)=\frac{1}{2}$.

$$\mu_1=\begin{bmatrix}-5\\5\end{bmatrix}\ ,\ \mu_2=\begin{bmatrix}10\\15\end{bmatrix}$$

$$\Sigma_1=\begin{bmatrix}11 & 9\\9 & 11\end{bmatrix}\ ,\ \Sigma_2=\begin{bmatrix}2 & 0\\0 & 2\end{bmatrix}$$

Για τον υπολογισμό του S_w θα χρησιμοποιήσω τον ακόλουθο τύπο :

$S_w=\sum_{i=1}^{M} P_i\cdot\Sigma_i$, όπου M είναι ο αριθμός των κλάσεων.

Άρα

$$S_w=\sum_{i=1}^{2} P_i\cdot\Sigma_i=\frac{1}{2}(\Sigma_1+\Sigma_2)=\begin{bmatrix}\frac{13}{2} & \frac{9}{2}\\ \frac{9}{2} & \frac{13}{2}\end{bmatrix}$$

Για την εύρεση του αντίστροφου πίνακα S_w^{-1} χρησιμοποιείται ο παρακάτω τύπος :

$$\begin{pmatrix}a & b\\ c & d\end{pmatrix}=\frac{1}{a\cdot d-b\cdot c}\begin{pmatrix}d & -b\\ -c & a\end{pmatrix}$$

Επομένως ο πίνακας S_w^{-1} είναι ίσο με :

$$S_w^{-1}=\begin{bmatrix}\frac{13}{2} & \frac{9}{2}\\ \frac{9}{2} & \frac{13}{2}\end{bmatrix}^{-1}=\frac{1}{\frac{169}{4}-\frac{81}{4}}\begin{bmatrix}\frac{13}{2} & -\frac{9}{2}\\ -\frac{9}{2} & \frac{13}{2}\end{bmatrix}=\frac{1}{22}\begin{bmatrix}\frac{13}{2} & -\frac{9}{2}\\ -\frac{9}{2} & \frac{13}{2}\end{bmatrix}\ (1)$$

Το διάνυσμα προβολής για M = 2 υπολογίζεται με βάση το τύπο :

$$w=S_w^{-1}(\mu_1-\mu_2)\ \ (2)$$

Άρα το διάνυσμα προβολής είναι,

$$(1),(2)\Rightarrow w=\frac{1}{22}\begin{bmatrix}\frac{13}{2} & -\frac{9}{2}\\ -\frac{9}{2} & \frac{13}{2}\end{bmatrix}\begin{bmatrix}-15 & -10\end{bmatrix}=\begin{bmatrix}\frac{13}{2}(-15)+(-\frac{9}{2})(-10)\\ (-\frac{9}{2})(-15)+\frac{13}{2}(-10)\end{bmatrix}=\frac{1}{22}\begin{bmatrix}-\frac{105}{2}\\ \frac{5}{2}\end{bmatrix}=\begin{bmatrix}-\frac{105}{44}\\ \frac{5}{44}\end{bmatrix}\Rightarrow$$

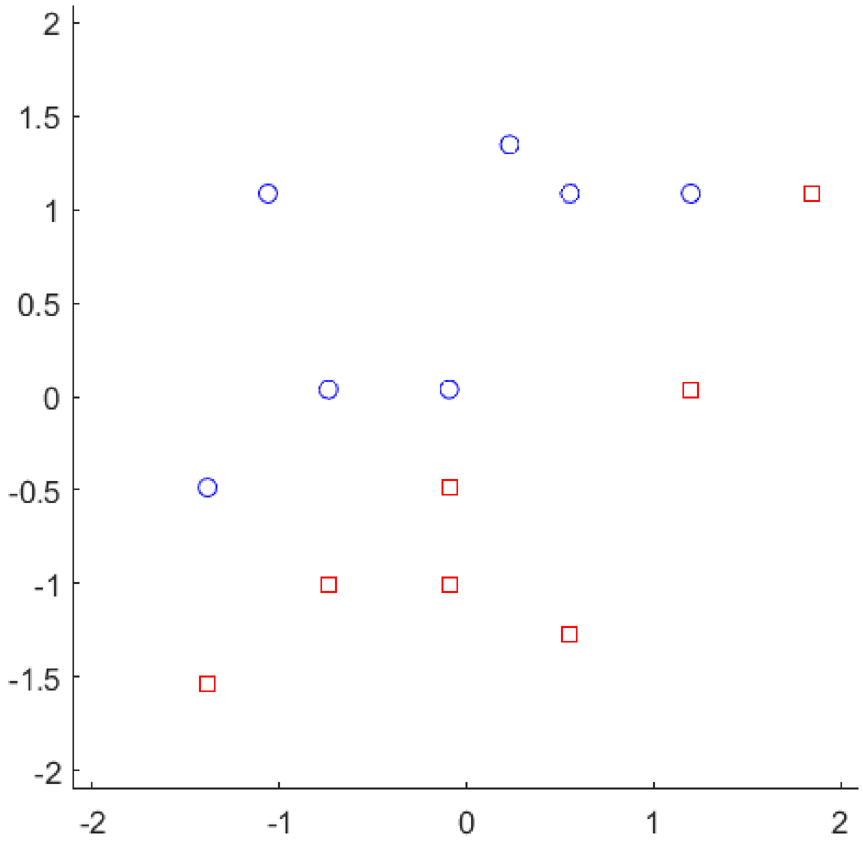
$$w=\begin{bmatrix}-\frac{105}{44}\\ \frac{5}{44}\end{bmatrix}$$

Θέμα 3: Linear Discriminant Analysis (LDA) vs PCA

Στην 3η άσκηση εφαρμόστηκε η supervised μέθοδος LDA η οποία αναπαριστά τα N-διαστατά δεδομένα σε ένα χώρο λιγότερων διαστάσεων. Όμως σε αντίθεση με τη μέθοδο PCA υπολογιζει τις κατευθύνσεις οι οποίες αντιπροσωπεύουν τους άξονες που μεγιστοποιούν τον διαχωρισμό μεταξύ των πολλαπλών κλάσεων.

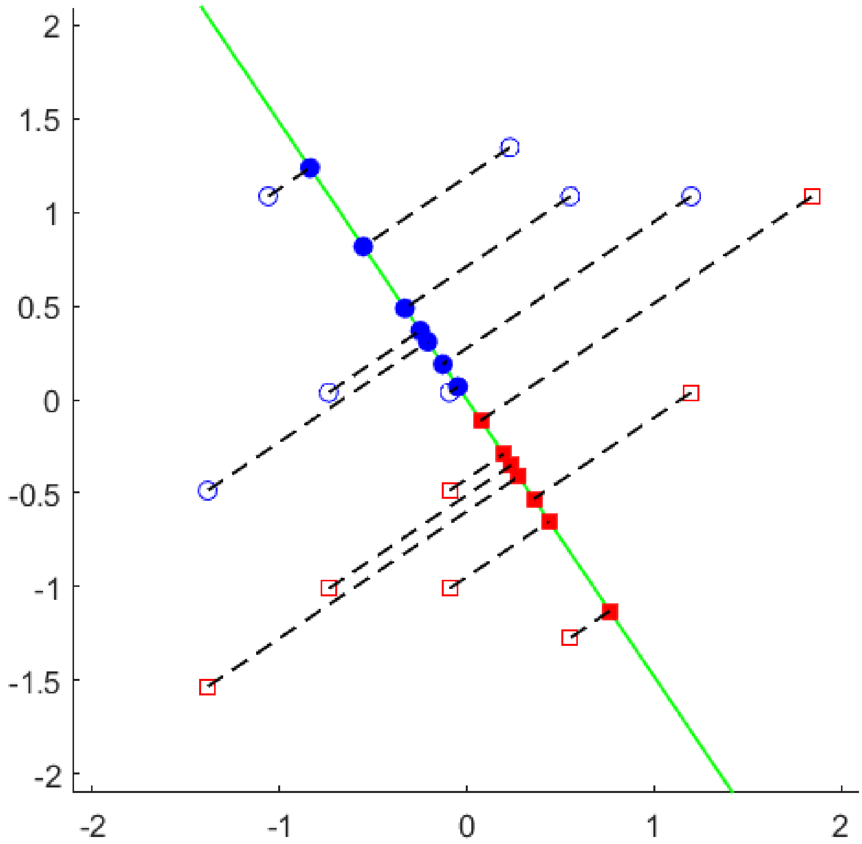
<u>Μέρος 1ο:</u>

Αρχικά φορτώνονται τα δεδομένα μέσω της εντολής load και δημιουργούνται οι μεταβλητές c,X οι οποίες περιέχουν την κλάση και τα χαρακτηριστικά διανύσματα αντίστοιχα. Στη συνέχεια πραγματοποιείται κανονικοποιήση των δειγμάτων με μέση τιμή μηδέν και διασπορά ίση με ένα χρησιμοποιώντας τη συνάρτηση featureNormalize που είχε υλοποιηθει στο θέμα 1, παράλληλα σχεδιάζονται οι κλάσεις στο σχήμα 10 με διαφορετικό χρώμα η κάθεμια έτσι ώστε να γίνονται εύκολα διαχωρίσιμες.



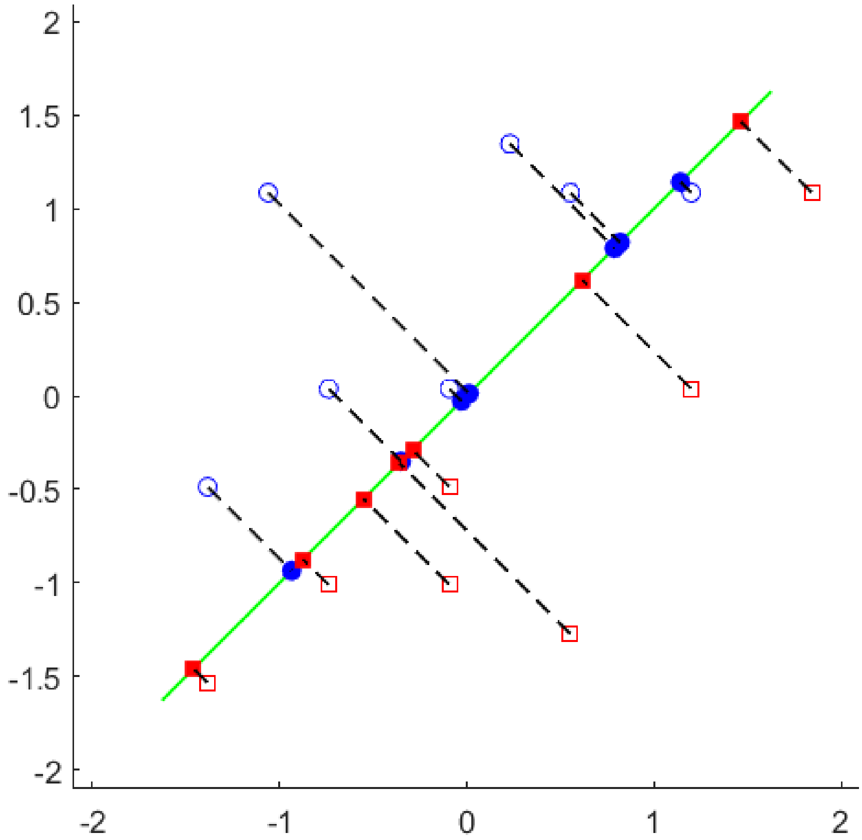
Σχήμα 10 : Αρχικά δείγματα κανονικοποιημένα

Κατόπιν υλοποιείται η συνάρτηση fisherLinearDiscriminant έτσι ώστε να εφαρμοστεί ο αλγόριθμος LDA για προβλήματα δύο κλάσεων. Δημιουργούνται δεδομένα μειωμένης διάστασης(1D) με την υλοποιήση της συνάρτησης projectDataLDA και ύστερα ανακατασκευάζονται τα δειγμάτα μειωμένης διάστασης στον 2D χώρο προβάλλοντας τα πάνω στην κατεύθυνση του διανύσματος προβολής LDA . Στο σχήμα 11 φαίνονται τα δείγματα μειωμένης διάστασης όπως προβάλλονται στον 2D χώρο μαζί με τα αρχικά δείγματα και παρατηρείται ότι έχει επιτευχθεί ο πλήρης διαχωρισμός των κλάσεων.



Σχήμα 11 : Αρχκά δείγματα κανονικοποιημένα και τα ανακτημένα δείγματα με τη χρήση της μεθόδου LDA

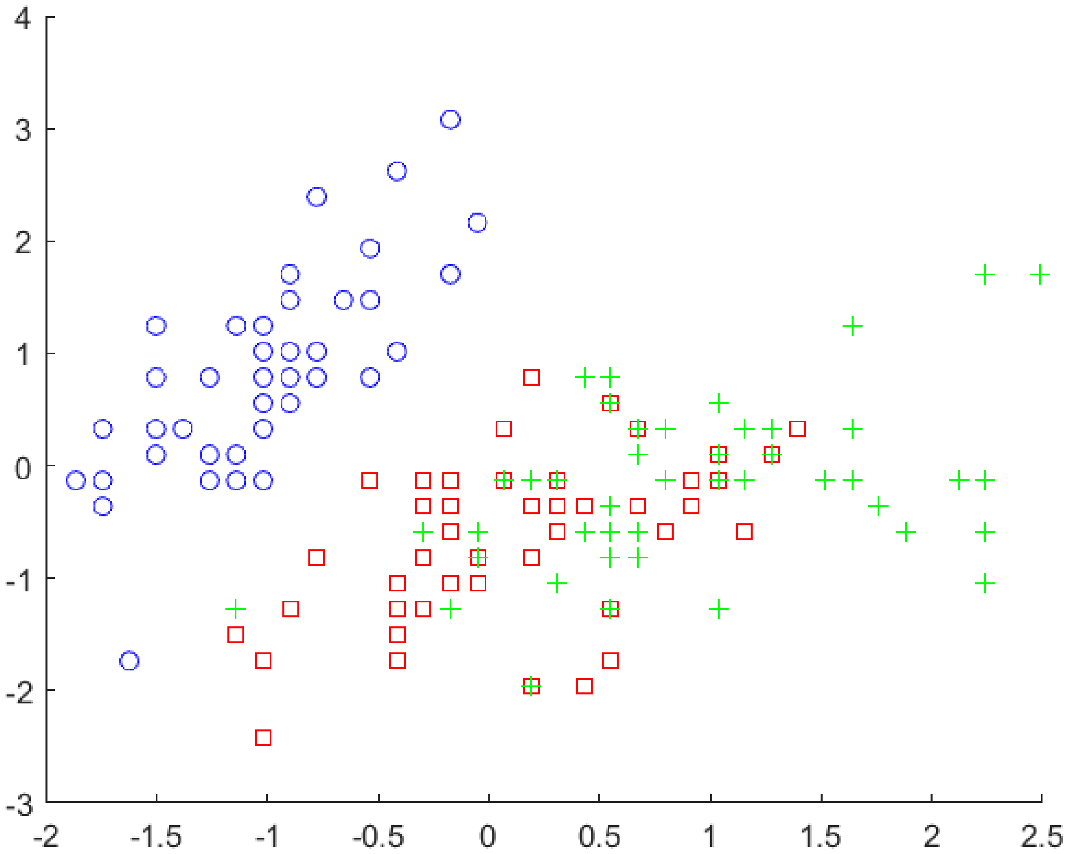
Εφαρμόζοντας την αντίστοιχη διαδικασία μείωσης διάστασης με τη μέθοδο PCA φαίνονται στο σχήμα 12 ότι δεν εχει καταστεί ικανός ο διαχωρισμός των κλάσων κάτι το ποίο είναι λογικό αφού η συγκεκριμένη μέθοδος δεν διαχωρίζει τα δεδομένα με βάση τις ετικέτες τους αλλα με βαση τη διασπορα τους . Συγκρίνοντας την τελευταία μέθοδο με τη μέθοδου διαχωρισμου LDA συμπεραίνεται ότι η μέθοδος linear discriminant analysis προτιμάται για data set τα οποία είναι επιθυμητό να διαχωριστούν με βάση κάποια labels έναντι της μεθόδου principal components analysis κατά την οποία δίνεται έμφαση στην έυρεση των κύριων συνιστωσών των δεδομένων δηλαδή τα δείγματα με τη μεγαλύτερη διασπορά.



Σχήμα 12 : Αρχικά δείγματα και ανακτόμενα δείγματα με τη χρήση της μεθόδου PCA

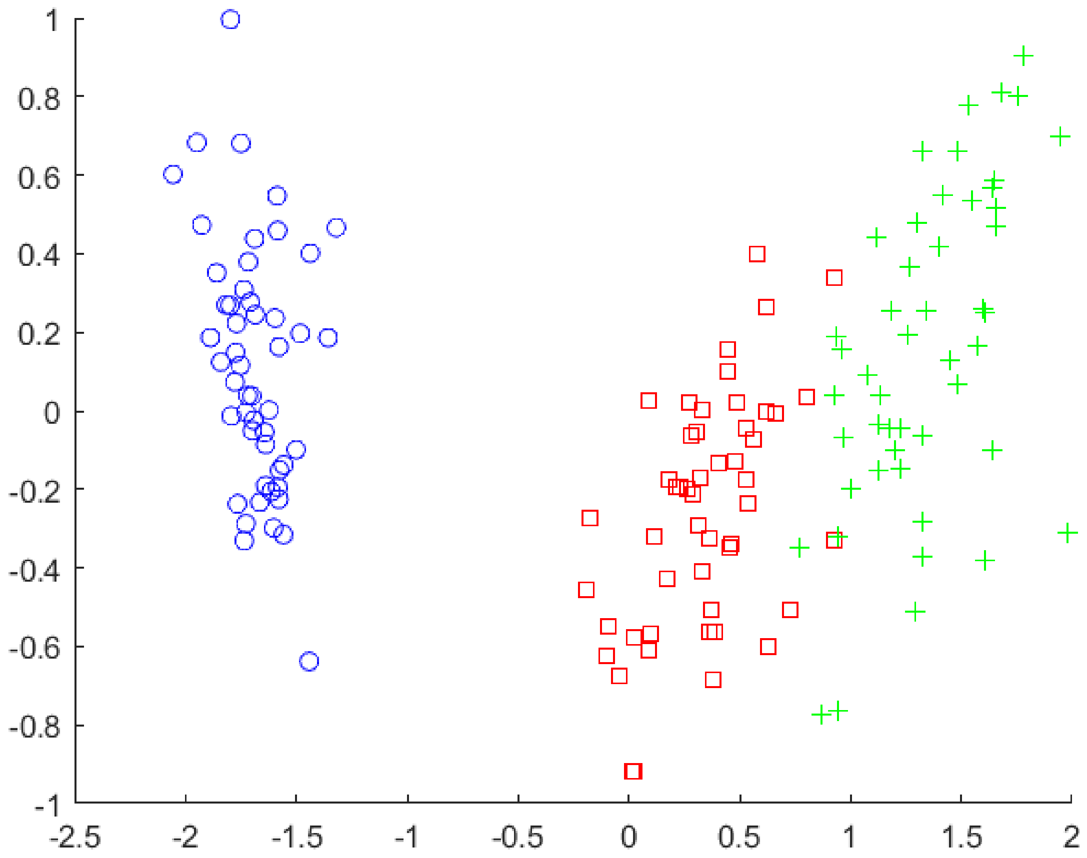
<u>Μέρος 2ο:</u>

Στο 2ο μέρος χρησιμοποιήθηκε ως dataset η βάση δεδομένων Fisher's Iris η οποία περιέχει 50 δείγματα απο τρία διαφορετικά είδη της οικογένειας λουλουδιών Iris και κάθε δείγμα αποτελείται από τέσσερα χαρακτηριστικά των λουλουδιών και συγκεκριμένα το μήκος και το πλάτος (σε cm) των σέπαλων και των ανθόφυλλων. Τα δεδομένα φορτώνονται μέσω της εντολής load , κανονικοποιούνται με μέση τιμή μηδέν και διασπορά ίση με 1 και σχεδιάζονται τα 2 πρώτα χαρακτηριστικά στο σχήμα 13 χρησιμοποιώντας ξεχωριστά χρώματα για κάθε κλάση.



Σχήμα 13 : Αρχικά δεδομένα από τη βάση δεδομένων iris

Στη συνέχεια υλοποιήθηκε η συνάρτηση myLDA μειώνοντας τον χώρο των χαρακτηριστικών στον δισδιάστατο χώρο και εφαρμόστηκαν τα διανύσματα προβολής που είχαν υπολογιστει με την myLDA πάνω στα αρχικά δείγματα με τη συνάρτηση projectDataLDA ώστε να μειωθεί η διάσταση του σε 2. Τα ιδιοδιανύσματα υπολογίζονται με χρήση της μεθόδου eigendecomposition σε αντίθεση με τον υπολογισμό του στο προηγούμενο ερώτημα.



Σχήμα 14 : Ανακτόμενα δείγματα με τη χρήση της μεθόδου LDA

Στο σχήμα 14 παρατηρείτα ότι έχει επιτύχει ο πλήρης διαχωρισμός των κλάσεων, το οποίο είναι και ο λόγος που χρησιμοποιούμε τη μέθοδο LDA έναντι της μεθόδου PCA.

Θέμα 4: Bayes

Για την εύρεση του συνόρου απόφασης θα χρησιμοποιήσω τη συνάρτηση διάκρισης

$$g_i(x) = P(\omega_i|x) = \frac{p(x|\omega_i)\cdot p(\omega_i)}{\sum_{j=1}^{C} p(x|\omega_j)\cdot p(\omega_j)}$$

Ειδικότερα για $g_i(x) > g_j(x)\, \mu\varepsilon\, i \neq j$ αποφασίζω ω_i αλλιώς ω_j

Ακόμη η συνάρτηση διάκρισης μπορεί να μετασχηματιστεί με τη χρήση της συνάρτησης ln διότι η ln είναι γνησίως άυξουσα συνάρτηση, κατά συνέπεια ο διαχωρισμός μεταξύ των κλάσεων ειναι αμετάβλητος. Επίσης ο πολλαπλασιασμός σταθερών με όλες τις συναρτήσεις διάκρισης δεν θα επηρεάσει το αποτέλεσμα του classification. Άρα η παρακάτω συνάρτηση $g_i(x)$ είναι ισοδύναμη με την αρχική συνάρτηση.

$$g_i(x) = P(\omega_i|x) = \ln(p(x|\omega_i)) + \ln(p(\omega_i))$$

Γνωρίζουμε όμως ότι οι κλάσεις περιγράφονται απο τις ακόλουθες κανονικές κατανομές,

$$p(x|\omega_i)=N(\mu_i,\Sigma_i)$$

με μέσες τιμες ίσες με

$$\mu_1=\begin{bmatrix}3\\3\end{bmatrix} \ , \ \mu_2=\begin{bmatrix}6\\6\end{bmatrix}$$

με πίνακες συνδιασποράς και τις ορίζουσες ίσες με»

$$\Sigma_1=\begin{pmatrix}1.2 & -0.4\\-0.4 & 1.2\end{pmatrix} \ , \ \Sigma_2=\begin{pmatrix}1.2 & 0.4\\0.4 & 1.2\end{pmatrix} \ , \ |\Sigma_1|=|\Sigma_2|=\frac{1}{1.28}$$

Άρα προκύπτει οτι η discriminant function $g_i(x)$ μπορεί να αναλυθεί ως εξής:

$$g_i(x)=-\frac{1}{2}(x-\mu_i)^T\Sigma^{-1}(x-\mu_i)-\frac{d}{2}\ln(2\pi)-\frac{1}{2}\ln|\Sigma_i|+\ln P(\omega_i)$$

Πραματοποιώντας ορισμένες προσαιτεριστικές ιδιοτητες στην παραπάνω συνάρτηση , η $g_i(x)$ μπορεί να πάρει τη μορφή μίας τετραγωνικής συνάρτησης

$$g_i(x)=x^T W_i x+w_i^T x+w_{i0}$$

με

$$W_i=-\frac{1}{2}\Sigma_i^{-1}$$

$$W_i=\Sigma_i^{-1}\mu_i$$

$$w_{i0}=-\frac{1}{2}\mu_i^T\Sigma^{-1}\mu_i-\frac{1}{2}\ln|\Sigma_i|+\ln P(\omega_i)$$

Οι αντιστροφοι πίνακες των πινακων συνδιασπορας υπολογίζονται με βάση τον τύπο που αναπτυχθηκε στο θέμα 2ο και είνα ίσοι με

$$\Sigma_1^{-1}=\frac{1}{1.28}\begin{pmatrix}1.2 & 0.4\\0.4 & 1.2\end{pmatrix} \ , \ \Sigma_2^{-1}=\frac{1}{1.28}\begin{pmatrix}1.2 & -0.4\\-0.4 & 1.2\end{pmatrix}$$

Στη συνέχεια

$$W_1=-\frac{1}{2}\Sigma_1^{-1}=-\frac{1}{2.56}\begin{pmatrix}1.2 & 0.4\\0.4 & 1.2\end{pmatrix}$$

$$W_2 = -\frac{1}{2}\Sigma_2^{-1} = -\frac{1}{2.56}\begin{pmatrix} 1.2 & -0.4 \\ -0.4 & 1.2 \end{pmatrix}$$

$$w_1 = \Sigma_1^{-1}\mu_1 = \frac{1}{1.28}\begin{pmatrix} 1.2 & 0.4 \\ 0.4 & 1.2 \end{pmatrix}\begin{bmatrix} 3 \\ 3 \end{bmatrix} = \begin{bmatrix} \frac{3.6}{1.28} + \frac{1.2}{1.28} \\ \frac{1.2}{1.28} + \frac{3.6}{1.28} \end{bmatrix} = \begin{bmatrix} \frac{4.8}{1.28} \\ \frac{4.8}{1.28} \end{bmatrix}$$

$$w_2 = \Sigma_2^{-1}\mu_2 = \frac{1}{1.28}\begin{pmatrix} 1.2 & -0.4 \\ -0.4 & 1.2 \end{pmatrix}\begin{bmatrix} 6 \\ 6 \end{bmatrix} = \begin{bmatrix} \frac{7.2}{1.28} - \frac{2.4}{1.28} \\ -\frac{2.4}{1.28} + \frac{7.2}{1.28} \end{bmatrix} = \begin{bmatrix} \frac{4.8}{1.28} \\ \frac{4.8}{1.28} \end{bmatrix}$$

Με χρήση του συνόρου απόφασης που αναπτύχθηκε παραπάνω για $g_1(x) > g_2(x)$, αποφασίζω ω_1 αλλιώς ω_2 . Άρα,

$$g_1(x) > g_2(x) \Leftrightarrow x^T W_1 x + w_1^T x + w_{10} > x^T W_i x + w_2^T x + w_{20} \Leftrightarrow$$
$$x^T W_1 x - \frac{1}{2}\mu_1{}^T \Sigma^{-1}\mu_1 - \frac{1}{2}\ln|\Sigma_1| + \ln P(\omega_1) > x^T W_2 x - \frac{1}{2}\mu_2{}^T \Sigma^{-1}\mu_2 - \frac{1}{2}\ln|\Sigma_2| + \ln P(\omega_2) \Leftrightarrow$$
$$x^T(W_1 - W_2)x - \frac{1}{2}(\mu_1{}^T \Sigma^{-1}\mu_1 - \mu_2{}^T \Sigma^{-1}\mu_2) + \ln\frac{P(\omega_1)}{P(\omega_2)} > 0 \quad (I)$$

$$W_1 - W_2 = -\frac{1}{2.56}\begin{pmatrix} 1.2 & 0.4 \\ 0.4 & 1.2 \end{pmatrix} - \begin{pmatrix} 1.2 & -0.4 \\ -0.4 & 1.2 \end{pmatrix} = -\frac{1}{2.56}\begin{pmatrix} 0 & 0.8 \\ 0.8 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -\frac{0.8}{2.56} \\ -\frac{0.8}{2.56} & 0 \end{pmatrix}$$

$$\mu_1{}^T \Sigma^{-1}\mu_1 = (3 \quad 3)(\frac{1}{1.28})\begin{pmatrix} 1.2 & 0.4 \\ 0.4 & 1.2 \end{pmatrix}\begin{bmatrix} 3 \\ 3 \end{bmatrix} = (3 \quad 3)\begin{bmatrix} \frac{4.8}{1.28} \\ \frac{4.8}{1.28} \end{bmatrix} = \frac{6 \cdot 4.8}{1.28} = \frac{28.8}{1.28}$$

$$\mu_1{}^T \Sigma^{-1}\mu_1 = (6 \quad 6)(\frac{1}{1.28})\begin{pmatrix} 1.2 & -0.4 \\ -0.4 & 1.2 \end{pmatrix}\begin{bmatrix} 6 \\ 6 \end{bmatrix} = (6 \quad 6)\begin{bmatrix} \frac{4.8}{1.28} \\ \frac{4.8}{1.28} \end{bmatrix} = \frac{12 \cdot 4.8}{1.28} = \frac{57.6}{1.28}$$

$$(I) \Rightarrow x^T \begin{pmatrix} 0 & -\frac{0.8}{2.56} \\ -\frac{0.8}{2.56} & 0 \end{pmatrix} x - \frac{1}{2} \cdot \frac{28.8}{1.28} + \ln \frac{P(\omega_1)}{P(\omega_2)} > 0 \Leftrightarrow (x_1 \quad x_2) \begin{pmatrix} 0 & -\frac{0.8}{2.56} \\ -\frac{0.8}{2.56} & 0 \end{pmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \frac{1}{2} \cdot \frac{28.8}{1.28} + \ln \frac{P(\omega_1)}{P(\omega_2)} > 0 \Leftrightarrow$$

$$(x_1 \quad x_2) \begin{bmatrix} (-\frac{0.8}{2.56})\, x_2 \\ (-\frac{0.8}{2.56})\, x_1 \end{bmatrix} - 11.25 + \ln \frac{P(\omega_1)}{P(\omega_2)} > 0 \Leftrightarrow (-\frac{0.8}{1.28})\, x_1 x_2 - 11.25 + \ln \frac{P(\omega_1)}{P(\omega_2)} > 0 \Leftrightarrow$$

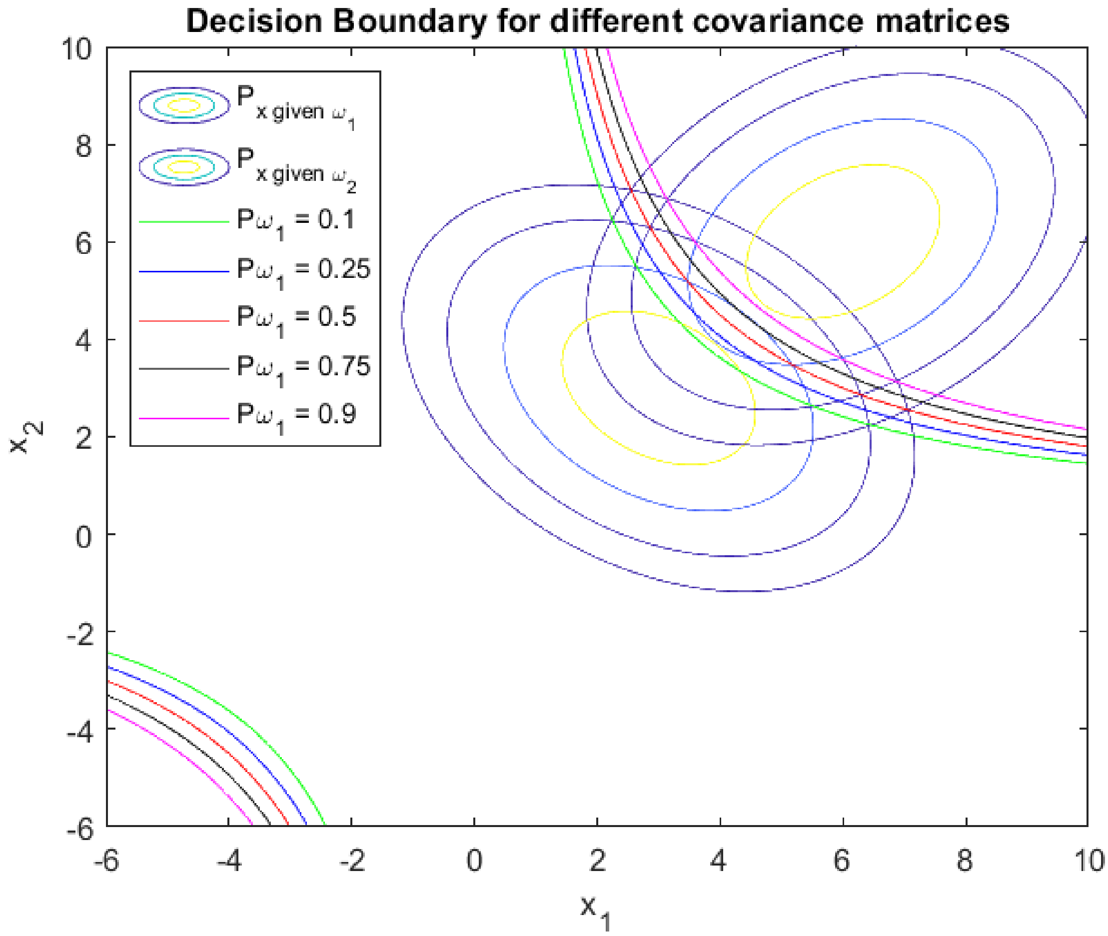
$$x_1 x_2 < 18 + 1.6 \cdot \ln \frac{P(\omega_1)}{P(\omega_2)}$$

Άρα το σύνορο απόφασης είναι ότι για $x_1 x_2 < 18 + 1.6 \cdot \ln \frac{P(\omega_1)}{P(\omega_2)}$ αποφασίζω ω_1 αλλιώς ω_2 .

Στο σχήμα 16 έχουνν σχεδιαστεί μερικές ισοϋψείς καμπύλες των δεσμευμένων πιθανοτήτων $p(x|\omega_i) \in \Re$ για κάθε κλάση i μαζί με τα σύνορα απόφασης για τις ακόλουθες τιμες των a-priori πιθανοτήτων $P(\omega_1) = 0.1,\ 0.25,\ 0.5,\ 0.75,\ 0.9$ και $P(\omega_2) = 1 - P(\omega_1)$.Τα σύνορα απόφασης που προκύπτουν είναι ορθογώνιες υπερβολές, το οποίο επαληθέυεται τόσο από τον θεωρητικό τύπο όσο και απο τη σχεδίαση τους. Επίσης το σύνορο απόφασης μπορεί να γραφεί ως:

$$x_1 x_2 < 18 + 1.6 \cdot \ln \frac{P(\omega_1)}{1 - P(\omega_1)} \quad (a)$$

Παρατηρείται στο σχήμα 16 ότι αυξάνωντας την εκ των προτέρων πιθανοτητα $P(\omega_1)$ μειώνεται προφανώς η εκ των προτέρων πιθανοτητα $P(\omega_2)$ και το σύνορο απόφασης μετατίθεται σε πιο υψηλές τιμές αφόυ αυξάνεται το δεξιό μέλος της ανισότητας α. Επίσης όσο αυξάνεται η $P(\omega_1)$ το σύνορο απόφασης πλησιάζει την λιγότερη πιθανή κλάση, άρα ο κανόνας αποφασίζει περισσότερο υπέρ της πιθανότερης κλάσης.



Σχήμα 16 : Σύνορο απόφασης για τους διαφορετικούς πίνακες συνδιασποράς,με τις ισοϋψείς καμπύλες πάνω δεξιά να ανήκουν στο ω_2 και τις ισοϋψείς καμπύλες στο κέντρο να ανήκουν στο ω_1

Υποθέτωντας ότι οι πίνακες συνδιασποράς είναι ίδιοι τότε ομοίως προκύπτει ότι :

$$\Sigma=\Sigma_1=\Sigma_2=\begin{pmatrix}1.2 & 0.4\\ 0.4 & 1.2\end{pmatrix}$$

$$\mu_1=\begin{bmatrix}3\\3\end{bmatrix} \ , \quad \mu_2=\begin{bmatrix}6\\6\end{bmatrix}$$

$$g_i(x)=-\frac{1}{2}(x-\mu_i)^T\Sigma^{-1}(x-\mu_i)+\ln P(\omega_i)$$

$$g_i(x)=w_i^T x+w_{i0}$$
$$w_i=\Sigma^{-1}\mu_i$$

$$w_{i0}=-\frac{1}{2}\mu_i^T\Sigma^{-1}\mu_i+\ln P(\omega_i)$$

$g_1(x)=w_1^T x+w_{10}$

$$w_1=\Sigma^{-1}\mu_1=\frac{1}{1.28}\begin{pmatrix}1.2 & -0.4\\ -0.4 & 1.2\end{pmatrix}\begin{bmatrix}3\\3\end{bmatrix}=\begin{pmatrix}\frac{1.2}{1.28} & \frac{-0.4}{1.28}\\ \frac{-0.4}{1.28} & \frac{1.2}{1.28}\end{pmatrix}\begin{bmatrix}3\\3\end{bmatrix}=\begin{bmatrix}\frac{3.6-1.2}{1.28}\\ \frac{-1.2+3.6}{1.28}\end{bmatrix}=\begin{bmatrix}\frac{2.4}{1.28}\\ \frac{2.4}{1.28}\end{bmatrix}$$

$$w_{10}=-\frac{1}{2}\begin{bmatrix}3\\3\end{bmatrix}^T\frac{1}{1.28}\begin{pmatrix}1.2 & -0.4\\ -0.4 & 1.2\end{pmatrix}\begin{bmatrix}3\\3\end{bmatrix}+\ln P(\omega_1)=-\frac{1}{2}(3\quad 3)\begin{bmatrix}\frac{2.4}{1.28}\\ \frac{2.4}{1.28}\end{bmatrix}+\ln P(\omega_1)=-\frac{7.2}{1.28}\ln P(\omega_1)$$

$g_2(x)=w_2^T x+w_{20}$

$$w_2=\Sigma^{-1}\mu_2=\frac{1}{1.28}\begin{pmatrix}1.2 & -0.4\\ -0.4 & 1.2\end{pmatrix}\begin{bmatrix}6\\6\end{bmatrix}=\begin{pmatrix}\frac{1.2}{1.28} & \frac{-0.4}{1.28}\\ \frac{-0.4}{1.28} & \frac{1.2}{1.28}\end{pmatrix}\begin{bmatrix}6\\6\end{bmatrix}=\begin{bmatrix}\frac{7.2-2.4}{1.28}\\ \frac{-2.4+7.2}{1.28}\end{bmatrix}=\begin{bmatrix}\frac{4.8}{1.28}\\ \frac{4.8}{1.28}\end{bmatrix}$$

$$w_{20}=-\frac{1}{2}\begin{bmatrix}6\\6\end{bmatrix}^T\frac{1}{1.28}\begin{pmatrix}1.2 & -0.4\\ -0.4 & 1.2\end{pmatrix}\begin{bmatrix}6\\6\end{bmatrix}+\ln P(\omega_2)=-\frac{1}{2}(6\quad 6)\begin{bmatrix}\frac{4.8}{1.28}\\ \frac{4.8}{1.28}\end{bmatrix}+\ln P(\omega_2)=-\frac{28.8}{1.28}\ln P(\omega_2)$$

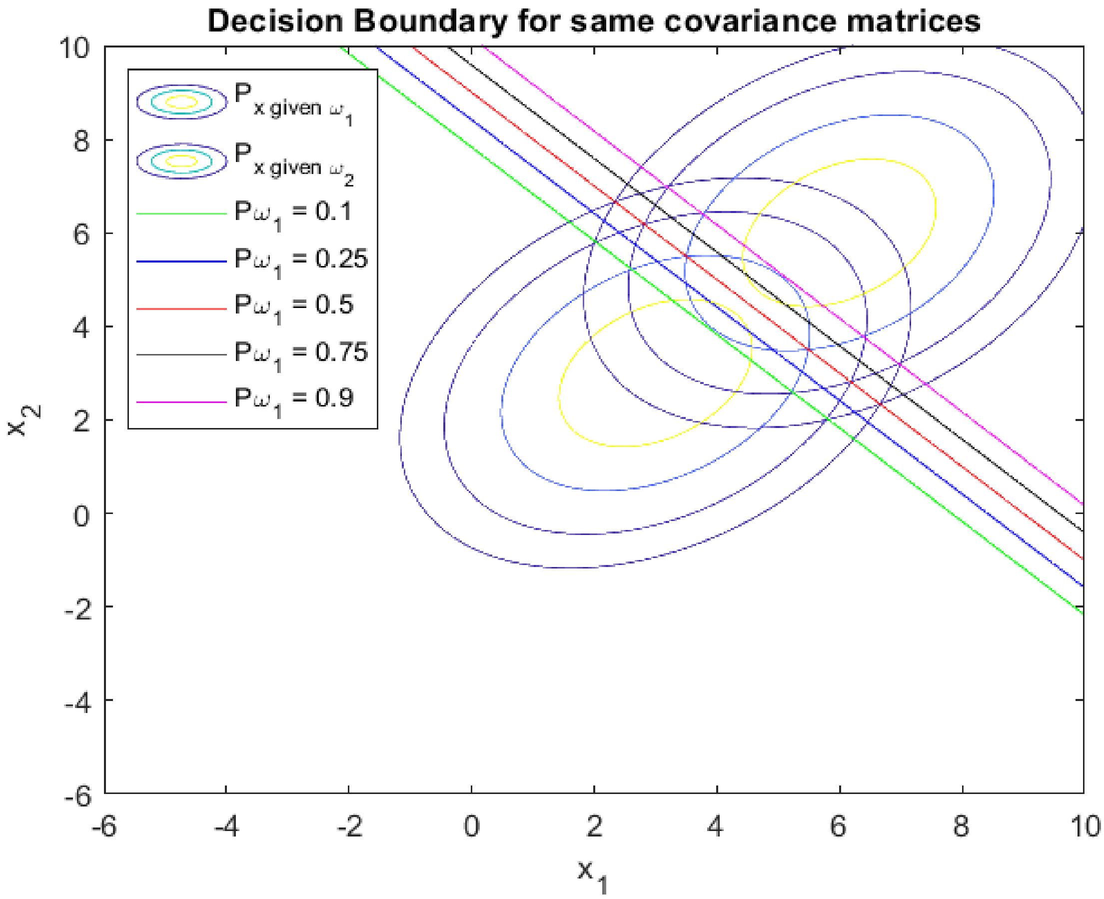
$$g_1(x)>g_2(x)\Leftrightarrow w_1^T x+w_{10}>w_2^T x+w_{20}\Leftrightarrow(w_1^T-w_2^T)x+w_{10}-w_{20}>0\Leftrightarrow$$
$$\begin{pmatrix}-\frac{2.4}{1.28} & -\frac{2.4}{1.28}\end{pmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}+16.875+\ln(\frac{P(\omega_1)}{P(\omega_2)})>0\Leftrightarrow(-\frac{2.4}{1.28})x_1+(-\frac{2.4}{1.28})x_2+16.875+\ln(\frac{P(\omega_1)}{P(\omega_2)})>0\Leftrightarrow$$
$$x_1+x_2-9-0.533\ln(\frac{P(\omega_1)}{P(\omega_2)})<0\Leftrightarrow x_1+x_2<9+0.533\ln(\frac{P(\omega_1)}{P(\omega_2)})$$

Άρα το σύνορο απόφασης είναι $x_1+x_2<9+0.533\ln(\frac{P(\omega_1)}{P(\omega_2)})$.

Στο σχήμα 17 έχουνν σχεδιαστεί μερικές ισοϋψείς καμπύλες των δεσμευμένων πιθανοτήτων $p(x|\omega_i)\in\Re$ για κάθε κλάση i μαζί με τα σύνορα απόφασης για τις ακόλουθες τιμες των a-priori πιθανοτήτων $P(\omega_1)=0.1, 0.25, 0.5, 0.75, 0.9$ και $P(\omega_2)=1-P(\omega_1)$. Τα σύνορα απόφασης που προκύπτουν είναι ευθείες γραμμές με αρνητική κλίση, το οποίο επαληθέυεται τόσο από τον θεωρητικό τύπο όσο και απο τη σχεδίαση τους. Επίσης το σύνορο απόφασης μπορεί να γραφεί ως:

$$x_1+x_2<9+0.533\ln\frac{P(\omega_1)}{1-P(\omega_1)}\quad (b)$$

Παρατηρείται στο σχήμα 17 ότι αυξάνωντας την εκ των προτέρων πιθανοτητα $P(\omega_1)$ μειώνεται προφανώς η εκ των προτέρων πιθανοτητα $P(\omega_2)$ και το σύνορο απόφασης μετατίθεται σε πιο υψηλές τιμές αφόυ αυξάνεται το δεξιό μέλος της ανισότητας b. Επίσης όσο αυξάνεται η $P(\omega_1)$ το σύνορο απόφασης πλησιάζει την λιγότερη πιθανή κλάση, άρα ο κανόνας αποφασίζει περισσότερο υπέρ της πιθανότερης κλάσης.

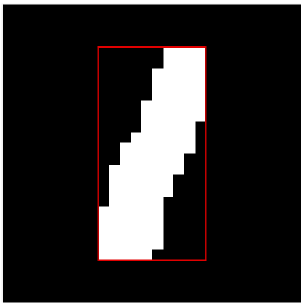
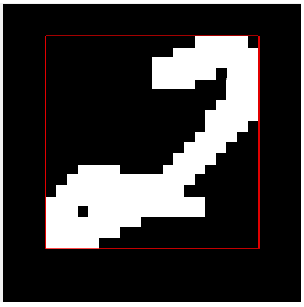


Σχήμα 17 : Σύνορο απόφασης για ίδιους πίνακες συνδιασποράςμε τις ισοϋψείς καμπύλες πάνω δεξιά να ανήκουν στο ω2 και τις ισοϋψείς καμπύλες στο κέντρο να ανήκουν στο ω1

Θέμα 5: Εξαγωγή χαρακτηριστικών και Bayes Classification

Στην 5η άσκηση υλοποιήθηκε ο υπολογισμός ενός χαρακτηριστικού για ορισμένες χειρόγραφες εικόνες οι οποίες αναπαριστούν ψηφία απο το 0 μεχρι το 9 και λαμβάνονται από ένα dataset. Συγκεκριμένα πραγματοποιήθηκε ο υπολογισμος του aspect ratio για τα ψηφία 1 και 2 , ο οποίος ορίζεται ως εξής:
$aspect\_ratio = width / height$ όπου το $width$ και $height$ είναι το μήκος και το ύψος του ελάχιστου ορθογωνίου που περικλείει ένα ψηφίο αντίστοιχα.

Αρχικά γίνεται η φόρτωση των δεδομένων και υπολογίζεται για τις εικόνες των ψηφίων 1,2 των δεδομένων εκπαίδευσης το aspect ratio κάθε μιας με τη χρήση της συνάρτησης compute_aspect_ratio. Η συνάρτηση υπολογισμού του aspect ratio εχει υλοποιήθει με τη λογική του ελάχιστου ορθογωνίου το οποίο περιβάλλει την αντίστοιχη εικόνα.Στη συνέχεια εκτυπώνο-νται ενδεικτικά δύο δείγματα από τα δύο ψηφία ένα από το καθένα τα οποία περιβάλλει το ελάχιστο ορθογώνιο(Bounding Box).

Από όλα τα δεδομένα εκπαίδευσης των ψηφίων 1,2 υπολογίζονται τα minAspectRatio και maxAspectRatio τα οποία είναι αντίστοιχα η ελάχιστη και η μέγιστη τιμή του λόγου όψεως από όλες τις εικόνες.

Υλοποιώντας έναν Naive Bayes ταξινομητή ο οποίος βασίζεται στην εφαρμογή του θεωρήματος του Bayes πραγματοποιείται η εκτίμηση της κλάσης που ανήκουν τα δείγματα ελέγχου. Ο Naive Bayes ταξινομητής χρησιμοποιεί τις εκ των υστέρων (a-posteriori) πιθανότητες, οι οποίες υπολογίστηκαν απο τα δεδομένα εκπαίδευσης(a-priori) και ελέγχου(Likelihood), για να εκτιμήσει σε ποιά κλάση ανήκει το εκάστοτε δείγμα ελέγχου. Συγκεκριμένα ο ταξινομητής επιλέγει ότι το δείγμα ανήκει στη κλάση του ψηφίου 1 αν η εκ των υστέρων πιθανότητα του ψηφίου 1 ειναι μεγαλύτερη απο την εκ των υστέρων πιθανότητα του ψηφίου 2, αλλιώς επιλέγει ότι το δείγμα ανήκει στη κλάση του ψηφίου 2. Τέλος υπολογίστηκε το σφάλμα εκτίμησης ως το ποσοστό των λανθασμένων αποφάσεων του ταξινομητή στο σύνολο των δειγμάτων ελέγχου το οποίο είναι 10.93%.

Θέμα 6ο : Minimum risk

$$P(\omega_1)=P(\omega_2)$$

$$p(x|\omega_i)=\begin{cases}\dfrac{x}{\sigma_i^2}\cdot e^{\frac{-x^2}{2\cdot\sigma_i^2}}, x\geq 0\\ 0, x<0\end{cases}$$

Ισχύει ότι $\sigma_1=1, \sigma_2=2$ και ο πίνακας

$$L=\begin{pmatrix} 0 & 0.5 \\ 1.0 & 0 \end{pmatrix}$$

Με το κριτήριο ελαχιστοποιήσης μέσου ρίσκου για δύο κλάσεις η ταξινόμηση γίνεται μέ βάση τις ποσότητες :

$$l_1=\lambda_{11}\cdot P(\omega_1)\cdot P(x|\omega_1)+\lambda_{21}\cdot P(\omega_2)\cdot P(x|\omega_2)$$
$$l_2=\lambda_{12}\cdot P(\omega_1)\cdot P(x|\omega_1)+\lambda_{22}\cdot P(\omega_2)\cdot P(x|\omega_2)$$

Ταξινομούμε το x στην κλάση ω_1 αν το $l_1 < l_2$ και στην κλάση ω_2 αν το $l_1 > l_2$.Άρα το σημείο διαχωρισμού (το κατώφλι x_o) είναι εκεί που:

$$l_1=l_2 \Leftrightarrow \lambda_{11}\cdot P(\omega_1)\cdot P(x|\omega_1)+\lambda_{21}\cdot P(\omega_2)\cdot P(x|\omega_2)=\lambda_{12}\cdot P(\omega_1)\cdot P(x|\omega_1)+\lambda_{22}\cdot P(\omega_2)\cdot P(x|\omega_2)$$
$$\Leftrightarrow 0+\lambda_{21}\cdot P(\omega_2)\cdot P(x|\omega_2)=\lambda_{12}\cdot P(\omega_1)\cdot P(x|\omega_1)+0 \Leftrightarrow \frac{P(x|\omega_1)}{P(x|\omega_2)}=\frac{\lambda_{21}\cdot P(\omega_2)}{\lambda_{12}\cdot P(\omega_1)} \quad (3)$$

Ακομή είναι γνωστό ότι οι εκ των προτέρων(a priori) πιθανότητες των δύο κλάσεων είνα ίσες δηλαδή ισχύει ότι $P(\omega_1)=P(\omega_2)$.Άρα

$$(3) \Rightarrow \ln\left(\frac{P(x|\omega_1)}{P(x|\omega_2)}\right)=\ln\left(\frac{\lambda_{21}}{\lambda_{12}}\right) \Leftrightarrow \ln P(x|\omega_1)-\ln P(x|\omega_2)=\ln\left(\frac{\lambda_{21}}{\lambda_{12}}\right) \Leftrightarrow \ln\left(\frac{x}{\sigma_1^2}\cdot e^{\frac{-x^2}{2\cdot\sigma_1^2}}\right)-\ln\left(\frac{x}{\sigma_2^2}\cdot e^{\frac{-x^2}{2\cdot\sigma_2^2}}\right)=\ln\left(\frac{\lambda_{21}}{\lambda_{12}}\right)$$
$$\Leftrightarrow \ln\left(\frac{x}{2\cdot\sigma_1^2}\right)-\frac{x}{2\cdot\sigma_1^2}-\ln\left(\frac{x}{2\cdot\sigma_2^2}\right)+\frac{x}{2\cdot\sigma_2^2}=\ln\left(\frac{\lambda_{21}}{\lambda_{12}}\right) \Leftrightarrow \ln\left(\frac{\frac{x}{2\cdot\sigma_1^2}}{\frac{x}{2\cdot\sigma_2^2}}\right)-\frac{x}{2\cdot\sigma_1^2}+\frac{x}{2\cdot\sigma_2^2}=\ln\left(\frac{\lambda_{21}}{\lambda_{12}}\right)$$
$$\Leftrightarrow 2\cdot\ln\left(\frac{\sigma_2}{\sigma_1}\right)+x^2\left(\frac{1}{\sigma_2^2}-\frac{1}{\sigma_1^2}\right)=\ln\left(\frac{\lambda_{21}}{\lambda_{12}}\right) \Leftrightarrow 2\cdot\ln 2+x^2\left(\frac{1}{8}-\frac{1}{2}\right)=\ln 2 \Leftrightarrow x^2\cdot\left(\frac{-3}{8}\right)=-\ln 2$$
$$\Leftrightarrow x^2=\frac{8}{3}\cdot\ln 2 \Rightarrow x=\sqrt{\frac{8}{3}\cdot\ln 2}$$

Άρα το σύνορο απόφασης ισούται με $x_o=\sqrt{\frac{8}{3}\cdot\ln 2}$.

ΘΕΜΑ 7ο : Singular Value Decomposition (SVD)

$$\boldsymbol{X}=\begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 3 \end{bmatrix}$$

1)

Η ιδιοανάλυση του πίνακα $X^T\cdot X$:

$$A = X^T \cdot X = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 6 & 7 \\ 7 & 14 \end{bmatrix}$$

$$A - \lambda I = \begin{bmatrix} 6-\lambda & 7 \\ 7 & 14-\lambda \end{bmatrix}$$

Έυρεση των ιδιοτιμών του πίνακα

$$det(A - \lambda I) = 0 \Rightarrow \begin{bmatrix} 6-\lambda & 7 \\ 7 & 14-\lambda \end{bmatrix} = 0 \Rightarrow (6-\lambda)\cdot(14-\lambda) - 49 = 0 \Rightarrow (6-\lambda)\cdot(14-\lambda) = 49$$

$$6 \cdot 14 - 14 \cdot \lambda - 6 \cdot \lambda + \lambda^2 = 49 \Rightarrow \lambda^2 - 20 \cdot \lambda + 35 = 0 \Rightarrow \lambda = \frac{+20 \pm \sqrt{(260)}}{2} \Rightarrow \lambda = +10 \pm \sqrt{(65)} \Rightarrow$$

$$\lambda_{1,2} = +10 \pm \sqrt{(65)} \Rightarrow \lambda_1 = 18.062 \quad , \lambda_2 = 1.937$$

Έυρεση των ιδιοδιανυσμάτων του πίνακα A

Για λ1 = 18.062:

$$(A - \lambda_1) \cdot X = 0 \Rightarrow \begin{bmatrix} 6-18.062 & 7 \\ 7 & 14-18.062 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0 \Rightarrow \begin{bmatrix} -12.062 & 7 \\ 7 & -4.062 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

$$-12.062 \cdot x_1 + 7 \cdot x_2 = 0 \Rightarrow 7 \cdot x_2 = 12.062 \cdot x_1 \Rightarrow x_2 = \frac{12.062}{7} \cdot x_1 \Rightarrow x_2 = 1.723 \cdot x_1$$

$$7 \cdot x_1 - 4.062 \cdot x_2 = 0 \Rightarrow 7 \cdot x_1 = 4.062 \cdot x_2 \Rightarrow x_2 = \frac{7}{4.062} \cdot x_1 \Rightarrow x_2 = 1.723 \cdot x_1$$

$$X_1 = \begin{bmatrix} 0.58 \\ 1 \end{bmatrix} \cdot t$$

Για λ2 = 1.937:

$$(A - \lambda_2) \cdot X = 0 \Rightarrow \begin{bmatrix} 6-1.937 & 7 \\ 7 & 14-1.937 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0 \Rightarrow \begin{bmatrix} 4.063 & 7 \\ 7 & 12.063 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0$$

$$4.063 \cdot x_1 + 7 \cdot x_2 = 0 \Rightarrow 7 \cdot x_2 = -4.063 \cdot x_1 \Rightarrow x_2 = \frac{-4.063}{7} \cdot x_1 \Rightarrow x_2 = -0.580 \cdot x_1$$

$$7 \cdot x_1 + 12.063 \cdot x_2 = 0 \Rightarrow 7 \cdot x_1 = -12.063 \cdot x_2 \Rightarrow x_2 = \frac{7}{-12.063} \cdot x_1 \Rightarrow x_2 = -0.580 \cdot x_1$$

$$X_2 = \begin{bmatrix} -1.723 \\ 1 \end{bmatrix} \cdot t$$

2) Έυρεση των singular values οι οποίες σχηματίζουν το πίνακα Σ

$$\sigma_1 = \sqrt{\lambda_1} = \sqrt{10 + \sqrt{65}} = 4.250$$

$$\sigma_2=\sqrt{\lambda_2}=\sqrt{10-\sqrt{65}}=1.392$$

3) Έυρεση των μη αρνητικών ιδιοδιανυσμάτων του πίνακα $X\cdot X^T$:

$$u_1=\frac{1}{\sqrt{\lambda_1}}\cdot X\cdot v_1=\frac{1}{4.25}\cdot\begin{bmatrix}1&2\\2&1\\1&3\end{bmatrix}\cdot\begin{bmatrix}0.58\\1\end{bmatrix}\Rightarrow u_1=\frac{1}{4.25}\cdot\begin{bmatrix}0.58+2\cdot1\\2\cdot0.58+1\\0.58+3\cdot1\end{bmatrix}\Rightarrow u_1=\frac{1}{4.25}\cdot\begin{bmatrix}2.58\\2.16\\3.58\end{bmatrix}\Rightarrow$$

$$u_1=\begin{bmatrix}0.607\\0.508\\0.842\end{bmatrix}$$

$$u_2=\frac{1}{\sqrt{\lambda_2}}\cdot X\cdot v_2=\frac{1}{1.392}\cdot\begin{bmatrix}1&2\\2&1\\1&3\end{bmatrix}\cdot\begin{bmatrix}-1.723\\1\end{bmatrix}\Rightarrow u_2=\frac{1}{1.392}\cdot\begin{bmatrix}-1.723+2\cdot1\\2\cdot(-1.723)+1\\-1.723+3\cdot1\end{bmatrix}\Rightarrow u_2=\frac{1}{1.392}\cdot\begin{bmatrix}0.277\\-2.446\\1.277\end{bmatrix}\Rightarrow$$

$$u_2=\begin{bmatrix}-0.199\\-1.757\\0.917\end{bmatrix}$$

4)

$$X=U\cdot\Sigma\cdot V^T=\begin{bmatrix}0.607&-0.199\\0.508&-1.757\\0.842&0.917\end{bmatrix}\cdot\begin{bmatrix}4.25&0\\0&1.392\end{bmatrix}\cdot\begin{bmatrix}0.58&1\\-1.723&1\end{bmatrix}$$

Επομένως η καλύτερη rank-1 προσέγγιση του πίνακα X είναι :

$$X_{rank-1}=4.25\cdot\begin{bmatrix}0.607\\0.508\\0.842\end{bmatrix}\cdot\begin{bmatrix}0.58&1\end{bmatrix}\Rightarrow$$

$$X_{rank-1}=\begin{bmatrix}1.496&2.579\\1.252&2.159\\2.075&3.578\end{bmatrix}$$

<u>Διευκρινίσεις πάνω στο κώδικα :</u>

Στους κώδικες που περιέχονται στο .zip αρχείο μαζί με την παρούσα αναφορά κρίνεται αναγκαίο να διευκρινιστουν κάποια κομμάτια κώδικα.Το πρώτο είναι στο 2ο μέρος του 1ου θέματος , στο οποίο ζητείται να επαναληφθεί η διαδικασία που προηγήθηκε για διαφορετικών αριθμό κύριων συνιστωσών(K =10 ,50,200).Αυτό έχει υλοποιήθει χειροκίνητα, συγκεκριμένα θέτοντας τις τιμές των σταθερών στις γραμμες 153 και 169 στο αρχείο ‘ex1_1_pca.m’ στις επιθυμητές τιμές.. Το άλλο είναι ότι για τη σχεδίαση των συνόρων απόφασης και των ισοϋψών καμπυλών των δεσμευμένων πιθανοτήτων υλοποιήθηκε κώδικας σε matlab ο οποίος βρισκεται στο φάκελο ‘exercise1_4 .m’.