

Άσκηση 1^η (3 μονάδες)

Δίνονται τα εξής πρότυπα: $x_1 = \begin{bmatrix} 1 \\ -0.5 \\ -1 \end{bmatrix}$, $x_2 = \begin{bmatrix} 3 \\ 1.25 \\ 2.5 \end{bmatrix}$, $x_3 = \begin{bmatrix} 3.5 \\ 1.75 \\ 3.5 \end{bmatrix}$, $x_4 = \begin{bmatrix} 0.5 \\ 0.25 \\ 0.5 \end{bmatrix}$

Χρησιμοποιώντας το κριτήριο της ομοιότητας να απορριφθεί ένα χαρακτηριστικό με βάσει το συντελεστή συσχέτισης. Γράψτε τα πρότυπα που προκύπτουν μετά την απόρριψη χαρακτηριστικού.

Όπως μπορεί κανείς να παρατηρήσει, το δεύτερο χαρακτηριστικό (δεύτερο στοιχεία σε καθένα από τα πρότυπα) είναι το μισό του τρίτου χαρακτηριστικού. Εφόσον γνωρίζουμε ότι η ανεξαρτησία των χαρακτηριστικών είναι αυτή που είναι σημαντική για να έχουμε κάνει καλή επιλογή χαρακτηριστικών θα πρέπει να απορριφθεί ένα από τα δύο αυτά χαρακτηριστικά που είναι απολύτως ισοδύναμα. Αυτό ποσοτικοποιείται υπολογίζοντας τους συντελεστές συσχέτισης ως εξής:

Αρχικά εκτιμούμε τις μέσες τιμές:

$$\mu_1 = \frac{1 + 3 + 3.5 + 0.5}{4} \Rightarrow \mu_1 = 2$$

$$\mu_2 = \frac{-0.5 + 1.25 + 1.75 + 0.25}{4} \Rightarrow \mu_2 = 0.688$$

$$\mu_3 = \frac{-1 + 2.5 + 3.5 + 0.5}{4} \Rightarrow \mu_3 = 1.375$$

Στη συνέχεια υπολογίζουμε τις τυπικές αποκλίσεις-συνδιασπορές:

$$\sigma_1^2 = \frac{1}{4} [(1-2)^2 + (3-2)^2 + (3.5-2)^2 + (0.5-2)^2] = 1.625 \Rightarrow \sigma_1 = 1.275$$

$$\sigma_2^2 = \frac{1}{4} [(-0.5-0.688)^2 + (1.25-0.688)^2 + (1.75-0.688)^2 + (0.25-0.688)^2] = 0.762$$

$$\Rightarrow \sigma_2 = 0.873$$

$$\sigma_3^2 = \frac{1}{4} [(-1-1.375)^2 + (2.5-1.375)^2 + (3.5-1.375)^2 + (0.5-1.375)^2] = 3.047$$

$$\Rightarrow \sigma_3 = 1.746$$

$$\sigma_{12} = \frac{1}{4} [(1-2)(-0.5-0.688) + (3-2)(1.25-0.688) + (3.5-2)(1.75-0.688) + (0.5-2)(0.25-0.688)] \Rightarrow \sigma_{12} = 1$$

$$\sigma_{13} = \frac{1}{4} [(1-2)(-1-1.375) + (3-2)(2.5-1.375) + (3.5-2)(3.5-1.375) + (0.5-2)(0.5-1.375)] \Rightarrow \sigma_{13} = 2$$

$$\sigma_{23} = \frac{1}{4} [(-0.5-0.688)(-1-1.375) + (1.25-0.688)(2.5-1.375) + (1.75-0.688)(3.5-1.375) + (0.25-0.688)(0.5-1.375)] \Rightarrow \sigma_{23} = 1.524$$

Κανονικοποιούμε τις συνδιασπορές ως εξής

$$\rho_{12} = \frac{\sigma_{12}}{\sigma_1 \sigma_2} = 0.898, \rho_{13} = \frac{\sigma_{13}}{\sigma_1 \sigma_3} = 0.898, \rho_{23} = \frac{\sigma_{23}}{\sigma_2 \sigma_3} = 1$$

Γνωρίζουμε ότι οι κανονικοποιημένοι συντελεστές συσχέτισης ρ που υπολογίσαμε λαμβάνουν τιμές στο διάστημα $[-1, 1]$ με τιμές κοντά στο μηδέν να ανταποκρίνονται σε χαρακτηριστικά που δε σχετίζονται ενώ με τιμές κοντά στο ± 1 να φανερώνουν ισχυρή συσχέτιση που είναι

ανεπιθύμητη. Είναι φανερό ότι ο συντελεστής συσχέτισης που δείχνει ποια χαρακτηριστικά σχετίζονται περισσότερο είναι ο ρ_{23} που σημαίνει ότι ένα από τα δύο χαρακτηριστικά (2 ή 3) μπορεί να απορριφθεί. Για να επιλέξουμε πιο από τα δύο θα απορριφθεί εξετάζουμε το καθένα ξεχωριστά ως προς τη συσχέτισή του με το τρίτο χαρακτηριστικό (το 1). Όποιο έχει ισχυρότερη συσχέτιση με το τρίτο χαρακτηριστικό (δηλαδή όποιο από τα ρ_{12} και ρ_{13} είναι μεγαλύτερο) είναι αυτό που επιλέγεται προς απόρριψη. Στη συγκεκριμένη περίπτωση είναι ίσα και συνεπώς όποιο από τα δύο (2 ή 3) απορριφθεί είναι το ίδιο σωστό αποτέλεσμα.

Αν απορριφθεί το 2^ο χαρακτηριστικό, τα πρότυπα γίνονται:

$$x_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, x_2 = \begin{bmatrix} 3 \\ 2.5 \end{bmatrix}, x_3 = \begin{bmatrix} 3.5 \\ 3.5 \end{bmatrix}, x_4 = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$$

Αν απορριφθεί το 3^ο χαρακτηριστικό, τα πρότυπα γίνονται:

$$x_1 = \begin{bmatrix} 1 \\ -0.5 \end{bmatrix}, x_2 = \begin{bmatrix} 3 \\ 1.25 \end{bmatrix}, x_3 = \begin{bmatrix} 3.5 \\ 1.75 \end{bmatrix}, x_4 = \begin{bmatrix} 0.5 \\ 0.25 \end{bmatrix}$$

Άσκηση 2^η (2.5 μονάδες)

Δίνονται τα εξής πρότυπα: $x_1 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, x_2 = \begin{bmatrix} 1 \\ 1.5 \end{bmatrix}, x_3 = \begin{bmatrix} -3 \\ -2 \end{bmatrix}, x_4 = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$

Με βάση την απόσταση city block, διαχωρίστε τα πρότυπα σε κλάσεις με τη μέθοδο της αλυσίδας.

Υπολογίζουμε αρχικά όλες τις αποστάσεις city block ως εξής:

$$d_{12} = |-1 - 1| + |-1 - 1.5| = 4.5$$

$$d_{13} = |-1 + 3| + |-1 + 2| = 3$$

$$d_{14} = |-1 - 3| + |-1 - 6| = 11$$

$$d_{23} = |1 + 3| + |1.5 + 2| = 7.5$$

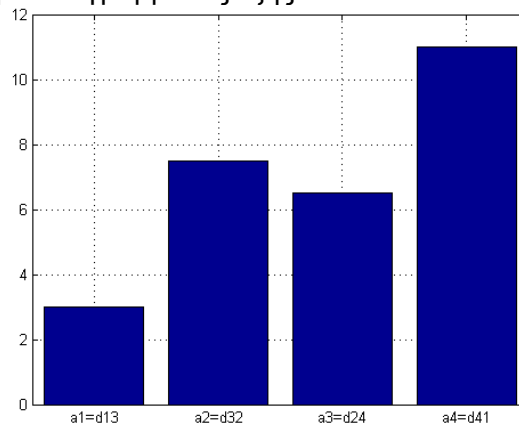
$$d_{24} = |1 - 3| + |1.5 - 6| = 6.5$$

$$d_{34} = |-3 - 3| + |-2 - 6| = 14$$

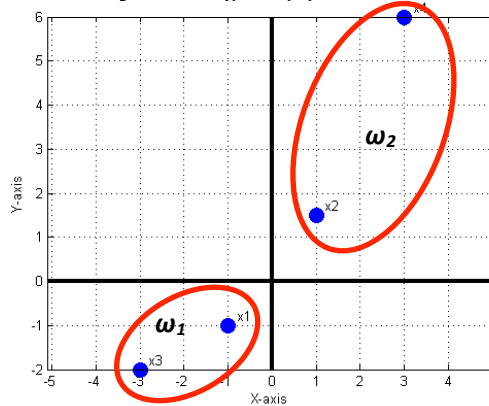
Μετά τον υπολογισμό των αποστάσεων μπορούμε να εκκινήσουμε τον αλγόριθμο της αλυσίδας.

1. Επιλέγουμε εκκίνηση (τυχαία) από το **πρότυπο 1**, και υπολογίζουμε
 $\alpha_1 = \min\{d_{12}, d_{13}, d_{14}\} = \{4.5, 3, 11\} = 3$ άρα πιο κοντά στο 1 είναι το **πρότυπο 3**
2. Υπολογίζουμε την ελάχιστη απόσταση από το **πρότυπο 3**
 $\alpha_2 = \min\{d_{23}, d_{34}\} = \{7.5, 14\} = 7.5$ άρα πιο κοντά στο 3 είναι το **πρότυπο 2**
3. Υπολογίζουμε την ελάχιστη απόσταση από το **πρότυπο 2**
 $\alpha_3 = \min\{d_{24}\} = 6.5$ άρα πιο κοντά στο 2 είναι το **πρότυπο 4**
4. Κλείνουμε «κυκλικά» από το **πρότυπο 4 στο πρότυπο 1** που ξεκινήσαμε
 $\alpha_4 = d_{41} = 11$

Από τα $\alpha_{1...4}$ κατασκευάζουμε ιστόγραμμα ως εξής:



Από το ιστόγραμμα αναγνωρίζουμε τις κλάσεις όπως ορίζονται από τις τιμές που βρίσκονται μεταξύ τοπικών μεγίστων. Τα τοπικά μέγιστα ορίζονται ως η αριστερή ακμή του ιστογράμματος, η μπάρα a_2 και η μπάρα a_4 . Ανάμεσα στα μέγιστα αυτά βρίσκονται τα a_1 και a_3 που αποτελούνται αντίστοιχα από τα πρότυπα $\{1,3\}$ και $\{2,4\}$, όπως φανερώνουν οι αντίστοιχες αποστάσεις d_{13} και d_{24} . Άρα οι κλάσεις που δημιουργούνται είναι οι $\omega_1=\{x_1, x_3\}$ και $\omega_2=\{x_2, x_4\}$



Άσκηση 3^η (3 μονάδες)

Δίνονται τα εξής πρότυπα: $x_1 = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$, $x_2 = \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$, $x_3 = \begin{bmatrix} 1.5 \\ 2 \\ 1 \end{bmatrix}$, $x_4 = \begin{bmatrix} -3 \\ -2 \\ 1 \end{bmatrix}$

Τα πρότυπα διαχωρίζονται γραμμικά στις κλάσεις $\omega_1=\{x_1, x_4\}$ (+) και $\omega_2=\{x_2, x_3\}$ (-). Σχεδιάστε νευρωνικό δίκτυο-ταξινομητή (perceptron) που να μπορεί να εκτελεί την ταξινόμηση των κλάσεων αυτών, με αρχικές συνθήκες: $w_1 = [0 \ 0 \ 0]^T$ και συντελεστή μάθησης $\rho=0.5$.

Για την ορθή λειτουργία του αλγορίθμου εργαζόμαστε στο χώρο των επαυξημένων διανυσμάτων, δηλαδή επαυξανούμε τη διάσταση των διανυσμάτων με μία ακόμη διάσταση προσθέτοντας σε όλα ένα τρίτο στοιχείο με τιμή μονάδα:

$$x_1 = \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix}, x_2 = \begin{bmatrix} -2 \\ 2 \\ 1 \\ 1 \end{bmatrix}, x_3 = \begin{bmatrix} 1.5 \\ 2 \\ 1 \\ 1 \end{bmatrix}, x_4 = \begin{bmatrix} -3 \\ -2 \\ 1 \\ 1 \end{bmatrix}$$

Αρχικά το διάνυσμα βαρών είναι:

$$w_1 = [0 \ 0 \ 0]^T$$

Η κλάση ω_1 έχει θετικό ενώ η κλάση ω_2 αρνητικό πρόσημο. Έτσι όταν προκύπτει γινόμενο αρνητικό για την ω_1 ή θετικό πρόσημο για την ω_2 απαιτείται αλλαγή προσήμου. Επίσης έχουμε συντελεστή μάθησης $\rho=0.5$.

ΕΠΟΧΗ 1:

$$\begin{aligned} w_1^T x_1 &= [0 \ 0 \ 0] \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix} = 0 \stackrel{=0}{\Rightarrow} w_2 = w_1 + \rho x_1 \Rightarrow w_2 = \begin{bmatrix} -0.5 \\ -0.5 \\ 0.5 \end{bmatrix} \\ w_2^T x_2 &= [-0.5 \ -0.5 \ 0.5] \begin{bmatrix} -2 \\ 2 \\ 1 \\ 1 \end{bmatrix} = 0.5 \stackrel{>0}{\Rightarrow} w_3 = w_2 - \rho x_2 \Rightarrow w_3 = \begin{bmatrix} 0.5 \\ -1.5 \\ 0 \end{bmatrix} \\ w_3^T x_3 &= [0.5 \ -1.5 \ 0] \begin{bmatrix} 1.5 \\ 2 \\ 1 \\ 1 \end{bmatrix} = -2.25 \stackrel{<0}{\Rightarrow} w_4 = w_3 \\ w_4^T x_4 &= [0.5 \ -1.5 \ 0] \begin{bmatrix} -3 \\ -2 \\ 1 \\ 1 \end{bmatrix} = 1.5 \stackrel{>0}{\Rightarrow} w_5 = w_4 \end{aligned}$$

Στην Εποχή 1 χρειάστηκαν ενημερώσεις βαρών οπότε απαιτείται και άλλη εποχή.

ΕΠΟΧΗ 2:

$$w_5^T x_1 = [0.5 \quad -1.5 \quad 0] \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} = 1 \stackrel{>0}{\Rightarrow} w_6 = w_5$$

$$w_6^T x_2 = [0.5 \quad -1.5 \quad 0] \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix} = -4 \stackrel{\leq 0}{\Rightarrow} w_7 = w_6$$

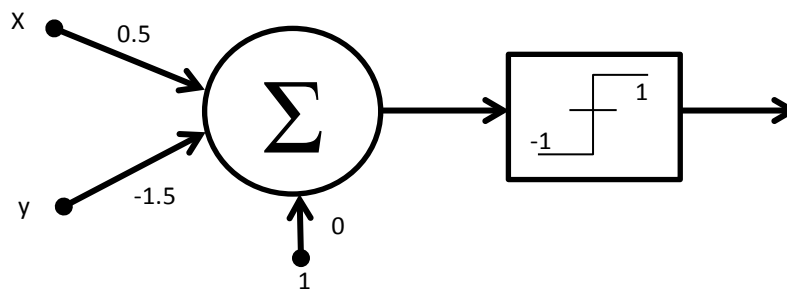
$$w_7^T x_3 = [0.5 \quad -1.5 \quad 0] \begin{bmatrix} 1.5 \\ 2 \\ 1 \end{bmatrix} = -2.25 \stackrel{\leq 0}{\Rightarrow} w_8 = w_7$$

$$w_8^T x_4 = [0.5 \quad -1.5 \quad 0] \begin{bmatrix} -3 \\ -2 \\ 1 \end{bmatrix} = 1.5 \stackrel{>0}{\Rightarrow} w_9 = w_8$$

Στην Εποχή 2 δεν πραγματοποιήθηκε καμία ενημέρωση βάρους οπότε ο αλγόριθμος καταλήγει στα εξής βάρη:

$$w = \begin{bmatrix} 0.5 \\ -1.5 \\ 0 \end{bmatrix}$$

Συνεπώς ο ταξινομητής είναι ένας νευρώνας, όπως φαίνεται στο σχήμα που ακολουθεί.



Άσκηση 4^η (1.5 μονάδα)

Έστω ότι μετά την εκπαίδευση ενός γραμμικού ταξινομητή-νευρώνα με βάρη $w = \begin{bmatrix} -0.4 \\ -0.1 \\ 0 \end{bmatrix}$ παρουσιάζεται το άγνωστο πρότυπο $x = [0 \quad 1 \quad 1]^T$. Ποια είναι η έξοδος του ταξινομητή και που ταξινομείται το εν λόγω πρότυπο αν οι κλάσεις ταξινόμησης είναι οι $\omega_1 (+)$ και $\omega_2 (-)$;

Για τον έλεγχο σε ποια κλάση ταξινομείται ένα άγνωστο πρότυπο αρκεί να γίνει ο πολλαπλασιασμός του διανύσματος του με το διάνυσμα των βαρών του ταξινομητή. Από το γινόμενο αυτό και το πρόσημο που προκύπτει το διάνυσμα ταξινομείται στη μία ή στην άλλη κλάση:

$$w^T x = [-0.4 \quad -0.1 \quad 0] \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = -0.1 \stackrel{\leq 0}{\Rightarrow} \text{Το πρότυπο ταξινομείται στην κλάση } \omega_2$$