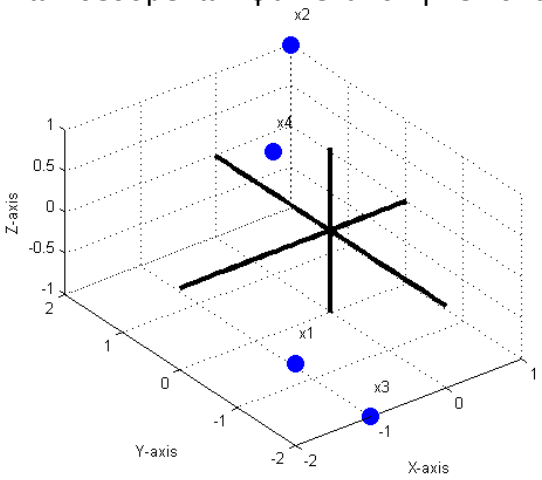
ΤΕΙ ΣΕΡΡΩΝ – ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΕΠΙΚΟΙΝΩΝΙΩΝ
ΕΞΕΤΑΣΗ ΣΤΟ ΜΑΘΗΜΑ «ΑΝΑΓΝΩΡΙΣΗ ΠΡΟΤΥΠΩΝ – ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ»
ΙΟΥΝΙΟΣ 2013

Δίνονται τα εξής πρότυπα: $x_1 = \begin{bmatrix} -2 \\ -2 \\ 0 \end{bmatrix}, x_2 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, x_3 = \begin{bmatrix} -1 \\ -2 \\ -1 \end{bmatrix}, x_4 = \begin{bmatrix} 0 \\ 1 \\ 0.5 \end{bmatrix}$

Άσκηση 1η (3 μονάδες)
Χρησιμοποιώντας το κριτήριο της ομοιότητας να απορριφθεί ένα χαρακτηριστικό με βάση το συντελεστή συσχέτισης. Γράψτε εδώ το χαρακτηριστικό που απορρίψατε και τη μορφή των προτύπων μετά την απόρριψη.

Μια απεικόνιση των αρχικών δεδομένων φαίνεται στην εικόνα που ακολουθεί.



Υπολογίζουμε τους συντελεστές συσχέτισης ως εξής:

Αρχικά εκτιμούμε τις μέσες τιμές:

$$\mu_1 = \frac{-2+1-1+0}{4} \Rightarrow \boldsymbol{\mu_1 = -0.5}$$

$$\mu_2 = \frac{-2+2-2+1}{4} \Rightarrow \boldsymbol{\mu_2 = -0.25}$$

$$\mu_3 = \frac{0+1-1+0.25}{4} \Rightarrow \boldsymbol{\mu_3 = 0.125}$$

Στη συνέχεια υπολογίζουμε τις τυπικές αποκλίσεις-συνδιασπορές:

$$\sigma_1^2 = \frac{1}{4}[(-2+0.5)^2 + (1+0.5)^2 + (-1+0.5)^2 + (0+0.5)^2] \Rightarrow \boldsymbol{\sigma_1 = 1.1180}$$

$$\sigma_2^2 = \frac{1}{4}[(-2+0.25)^2 + (2+0.25)^2 + (-2+0.25)^2 + (1+0.25)^2] \Rightarrow \boldsymbol{\sigma_2 = 1.7854}$$

$$\sigma_3^2 = \frac{1}{4}[(0-0.125)^2 + (1-0.125)^2 + (-1-0.125)^2 + (0.5-0.125)^2] \Rightarrow \boldsymbol{\sigma_3 = 0.7395}$$

$$\sigma_{12} = \frac{1}{4}[(-2+0.5)(-2+0.25) + (1+0.5)(2+0.25) + (-1+0.5)(-2+0.25) + (0+0.5)(1+0.25)] \Rightarrow \boldsymbol{\sigma_{12} = 1.875}$$

$$\sigma_{13} = \frac{1}{4}[(-2+0.5)(0-0.125) + (1+0.5)(1-0.125) + (-1+0.75)(-1-0.125) + (0+0.75)(0.5-0.125)] \Rightarrow \boldsymbol{\sigma_{13} = 0.5624}$$

$$\sigma_{23} = \frac{1}{4}[(-2+0.25)(0-0.125)+(2+0.25)(1-0.125)+(-2+0.25)(-1-0.125)$$
$$+(1+0.25)(0.5-0.125)] \Rightarrow \boldsymbol{\sigma_{23} = 1.1563}$$

Κανονικοποιούμε τις συνδιασπορές ως εξής

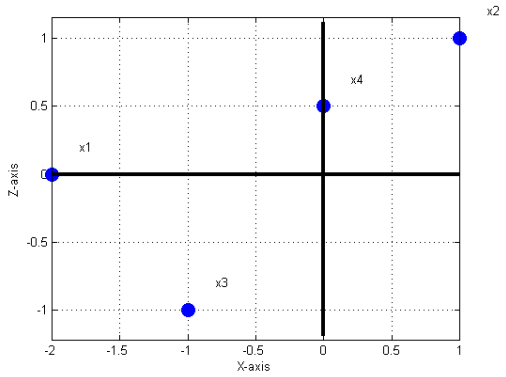
$$\rho_{12} = \frac{\sigma_{12}}{\sigma_1 \cdot \sigma_2} = 0.9393,\ \rho_{13} = \frac{\sigma_{13}}{\sigma_1 \cdot \sigma_3} = 0.6803,\ \rho_{23} = \frac{\sigma_{23}}{\sigma_2 \cdot \sigma_3} = 0.8758$$

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 συσχετίζονται ενώ με τιμές κοντά στο +/-1 να φανερώνουν ισχυρή συσχέτιση που είναι ανεπιθύμητη. Προς απόρριψη επιλέγεται το χαρακτηριστικό που υπάρχει ως δείκτης στους δύο από τους τρεις συντελεστές με τη μεγαλύτερη απόλυτη τιμή. Οι συντελεστές με τη μεγαλύτερη απόλυτη τιμή είναι οι ρ_{12} και ρ_{23}. Στους συντελεστές αυτούς κοινό χαρακτηριστικό είναι το 2 και συνεπώς **απορρίπτεται το 2ο χαρακτηριστικό**.

Τα πρότυπα πλέον είναι τα εξής:

$$x_1 = \begin{bmatrix} -2 \\ 0 \end{bmatrix}, x_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, x_3 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, x_4 = \begin{bmatrix} 0 \\ 0.5 \end{bmatrix}$$

Και η "εικόνα" τους στο χώρο των δύο πλέον διαστάσεων είναι η εξής:



Άσκηση 2η (2.5 μονάδες)

Με βάση την απόσταση city block, διαχωρίστε τα πρότυπα σε κλάσεις με τη μέθοδο της αλυσίδας. Γράψτε εδώ τις κλάσεις που προέκυψαν.

Υπολογίζουμε αρχικά όλες τις αποστάσεις city block ως εξής:

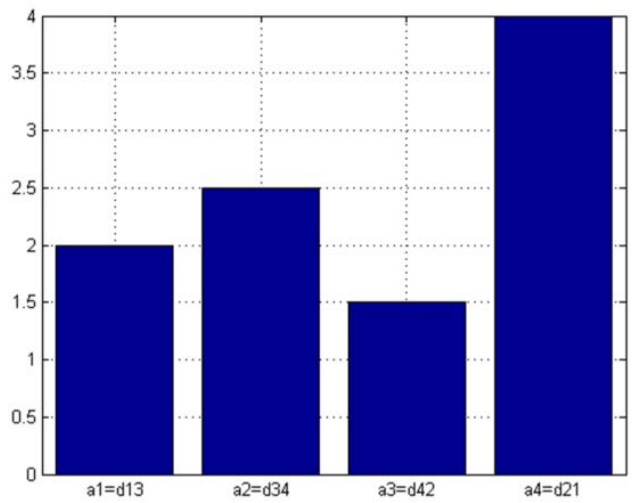
$$d_{12} = |-2-1| + |0-1| = 4$$
$$d_{13} = |-2+1| + |0+1| = 2$$
$$d_{14} = |-2+0| + |0-0.5| = 2.5$$
$$d_{23} = |1+1| + |1+1| = 4$$
$$d_{24} = |1+0| + |1-0.5| = 1.5$$
$$d_{34} = |-1+0| + |-1-0.5| = 2.5$$

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

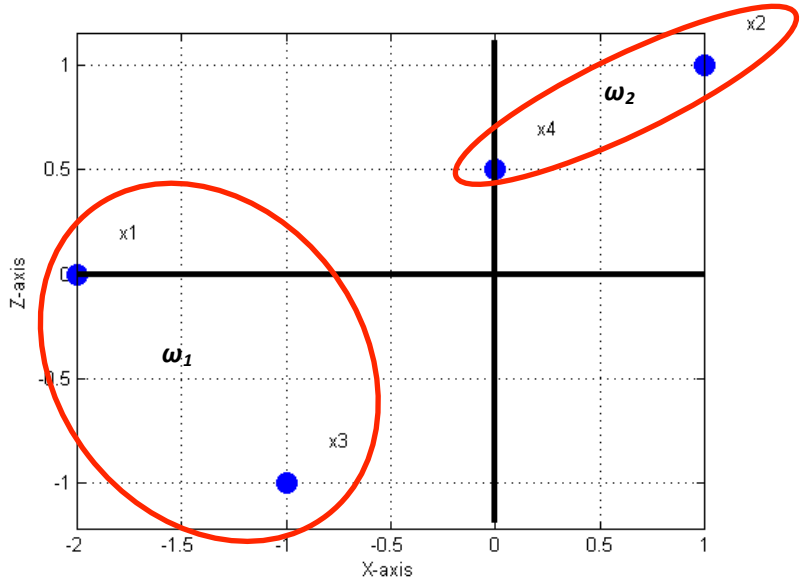
1. Επιλέγουμε εκκίνηση (τυχαία) από το **πρότυπο 1**, και υπολογίζουμε
   $\alpha_1 = min\{d_{12}, d_{13}, d_{14}\} = \{4, 2, 2.5\} = \mathbf{2}$ άρα πιο κοντά στο 1 είναι το **πρότυπο 3**
2. Υπολογίζουμε την ελάχιστη απόσταση από το **πρότυπο 3**
   $\alpha_2 = min\{d_{23}, d_{34}\} = \{4, 2.5\} = \mathbf{2.5}$ άρα πιο κοντά στο 3 είναι το **πρότυπο 4**

3. Υπολογίζουμε την ελάχιστη απόσταση από το **πρότυπο 4**
   $\alpha_3 = min\{d_{24}\} = \mathbf{1.5}$ άρα πιο κοντά στο 4 είναι το **πρότυπο 2**
4. Κλείνουμε «κυκλικά» από το **πρότυπο 2 στο πρότυπο 1** που ξεκινήσαμε
   $\alpha_4 = d_{12} = \mathbf{4}$

Από τα $a_{1...4}$ 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 3η (3 μονάδες)
Σχεδιάστε νευρωνικό δίκτυο-ταξινομητή που να μπορεί να εκτελεί την ταξινόμηση που προκύπτει στην Άσκηση 2. Εξηγήστε γιατί επιλέγετε το συγκεκριμένο ταξινομητή και τη συγκεκριμένη μέθοδο εκπαίδευσης. Αν χρησιμοποιήσετε perceptron: $w=[0\ 0\ 0]^T$ και $\rho=0.2$. Σχεδιάστε εδώ τον ταξινομητή με όλα τα στοιχεία του.

Από την Άσκηση 2 βρέθηκε ότι υπάρχουν δύο κλάσεις (και μάλιστα γραμμικώς διαχωρίσιμες). Άρα μπορούμε να χρησιμοποιήσουμε τον αλγόριθμο perceptron για να σχεδιάσουμε ταξινομητή δύο κλάσεων (νευρωνικό δίκτυο).
Για την ορθή λειτουργία του αλγορίθμου εργαζόμαστε στο χώρο των επαυξημένων διανυσμάτων, δηλαδή επαυξάνουμε τη διάσταση των διανυσμάτων με μία ακόμη διάσταση προσθέτοντας σε όλα ένα τρίτο στοιχείο με τιμή μονάδα:

$$x_1 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}, x_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, x_3 = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}, x_4 = \begin{bmatrix} 0 \\ 0.5 \\ 1 \end{bmatrix}$$

Επίσης ορίζουμε αρχικά το διάνυσμα βαρών ως:

$$w_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Ορίζουμε επίσης ότι η κλάση $\omega_1=\{x_1,x_3\}$ έχει θετικό πρόσημο και η κλάση $\omega_2=\{x_2,x_4\}$ αρνητικό πρόσημο. Έτσι όταν προκύπτει γινόμενο αρνητικό για την ω_1 ή θετικό πρόσημο για την ω_2 απαιτείται αλλαγή πρόσημου. Επίσης έχουμε συντελεστή μάθησης $\boldsymbol{\rho=0.2}$.

<u>ΕΠΟΧΗ 1:</u>

$$w_1^T x_1 = [0 \quad 0 \quad 0] \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} = 0 \overset{=0}{\Rightarrow} w_2 = w_1 + \rho x_1 \Rightarrow w_2 = \begin{bmatrix} -0.4 \\ 0 \\ 0.2 \end{bmatrix}$$

$$w_2^T x_2 = [-0.4 \quad 0 \quad 0.2] \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = -0.2 \overset{<0}{\Rightarrow} w_3 = w_2$$

$$w_3^T x_3 = [-0.4 \quad 0 \quad 0.2] \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} = 0.6 \overset{>0}{\Rightarrow} w_4 = w_3$$

$$w_4^T x_4 = [-0.4 \quad 0 \quad 0.2] \begin{bmatrix} 0 \\ 0.5 \\ 1 \end{bmatrix} = 0.2 \overset{>0}{\Rightarrow} w_5 = w_4 - \rho x_2 \Rightarrow w_3 = \begin{bmatrix} -0.4 \\ -0.1 \\ 0 \end{bmatrix}$$

Στην Εποχή 1 χρειάστηκαν ενημερώσεις βαρών οπότε απαιτείται και άλλη εποχή.
<u>ΕΠΟΧΗ 2:</u>

$$w_5^T x_1 = [-0.4 \quad -0.1 \quad 0] \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} = 0.8 \overset{>0}{\Rightarrow} w_6 = w_5$$

$$w_6^T x_2 = [-0.4 \quad -0.1 \quad 0] \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = -0.5 \overset{<0}{\Rightarrow} w_7 = w_6$$

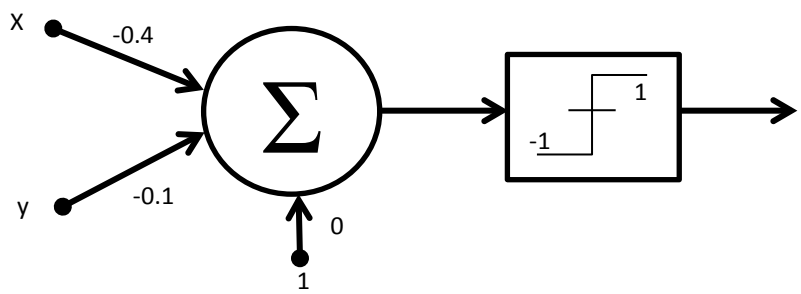
$$w_7^T x_3 = [-0.4 \quad -0.1 \quad 0] \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} = 0.5 \overset{>0}{\Rightarrow} w_8 = w_7$$

$$w_8^T x_4 = [-0.4 \quad -0.1 \quad 0] \begin{bmatrix} 0 \\ 0.5 \\ 1 \end{bmatrix} = -0.05 \overset{<0}{\Rightarrow} w_9 = w_8$$

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

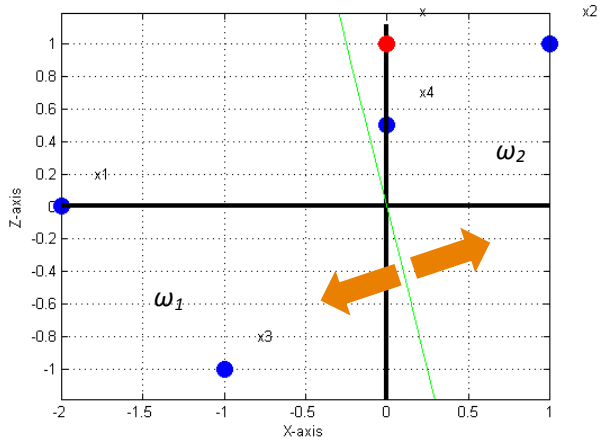
$$w = \begin{bmatrix} -0.4 \\ -0.1 \\ 0 \end{bmatrix}$$

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



Άσκηση 4η (1.5 μονάδα)
Έστω ότι μετά την εκπαίδευση του ταξινομητή παρουσιάζεται το πρότυπο $x = [0 \quad 0 \quad 1]^T$. Ποια είναι η έξοδος του ταξινομητή και που ταξινομείται το εν λόγω πρότυπο;

Για τον έλεγχο σε ποια κλάση ταξινομείται ένα άγνωστο πρότυπο αρκεί να γίνει ο πολλαπλασιασμός του διανύσματός 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.
Αρχικά απορρίπτουμε το 1° χαρακτηριστικό (0) όπως κάναμε για όλα τα πρότυπα στο σύστημά μας. Στη συνέχεια επαυξάνουμε το διάνυσμα προσθέτοντας μια μονάδα (1) ως τρίτη διάσταση και εκτελούμε το εσωτερικό γινόμενο με τα βάρη του ταξινομητή:

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

Στο σχήμα που ακολουθεί φαίνονται τα πρότυπα εκπαίδευσης x_1, x_2, x_3, x_4, η ευθεία *(ε)* που χωρίζει το χώρο στα ημιεπίπεδα των δύο κλάσεων (με πράσινο χρώμα) και η θέση του αγνώστου προτύπου *x* (με κόκκινο χρώμα).