No. Date

$\boxed{1.A}$ $\hat{\theta}_{ML} = ;$ αν $p(x;\theta) = \begin{cases} \frac{1}{\theta} e^{(-x/\theta)} & x \geq 0 \\ 0 & \text{αλλού} \end{cases}$

Λόγω της εκθετικής συνάρτησης και της μονοτονίας της μπορώ να υπολογίσω τη συνάρτηση λογαριθμικής πιθανοφάνειας

$$L(x;\theta) = L(\theta) = \ln \prod_{k=1}^{N} p(x_k;\theta) \overset{\log}{=} \sum_{n=1}^{N} \ln p(x_k,\theta)$$

$$= \sum_{1}^{N} \ln\left(\frac{1}{\theta}\right) \cdot e^{-x/\theta} \overset{(\log)}{=} \sum_{1}^{N} \ln \frac{1}{\theta} + \sum_{1}^{N} \ln e^{-x/\theta} \overset{\lambda = 1/\theta \; ③}{=}$$

$e^{\ln(x)} = x$

$\ln e^{x} = x$

$$= \sum_{1}^{N} \ln \lambda + \sum_{1}^{N} \ln e^{-x\lambda} = N \cdot \ln \lambda + \sum_{1}^{N} \cdot (-x\lambda) = N \cdot \ln \lambda - \lambda \sum_{1}^{N} x_i$$

$$\boxed{\frac{\partial L}{\partial \lambda} = N \cdot \frac{1}{\lambda} - \sum_{1}^{N} x_i} \quad ①$$

Αν το θέσουμε ίσο με μηδέν.

$$\frac{N}{\lambda} = \sum_{1}^{N} x_i \Rightarrow \boxed{\hat{\lambda}_{ML} = \frac{N}{\sum_{1}^{N} x_i}} \quad ②$$

Εμείς όμως ψάχνουμε το θ, οπότε

$$② \overset{③}{\Rightarrow} \frac{1}{\hat{\theta}_{ML}} = \frac{N}{\sum_{1}^{N} x_i} \Rightarrow \boxed{\hat{\theta}_{ML} = \sum_{1}^{N} x_i \cdot \frac{1}{N}}$$

1.B $\hat{\theta}_{ML}$ για $p(x;\theta) = \begin{cases} 1/2\theta & -\theta \le x \le \theta \\ 0 & \text{αλλού} \end{cases}$

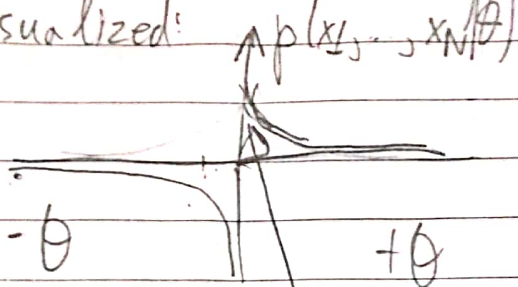
I: indicator function
και ισούται με 1 στο $[-\theta,\theta]$
αλλιώς 0

$[-\theta \le x_n \le \theta]$

$$\prod_{1}^{N} p(x_n;\theta) = \prod_{1}^{N} \frac{1}{2\theta} \cdot I(\cdot)$$

$$= \left(\frac{1}{2\theta}\right)^N \cdot I\left(\theta \ge \max_{n=1,\dots,N}\{x_n\}\right) \cdot I\left(\min_{n=1,\dots,N}\{x_n\} \ge -\theta\right)$$

Visualized:



$$\boxed{\hat{\theta}_{ML} = \max\{x_n\}}$$

No. Date

2 $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3$ $\ell = 3$, $M = 3$

$P(\omega_1) = P(\omega_2) = P(\omega_3) = 1/3$; $P(x; \omega_i) \sim N(\mu_i, I)$

$\mu_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$, $\mu_2 = \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}$, $\mu_3 = \begin{bmatrix} 0 \\ -1 \\ -1 \end{bmatrix}$ Bayesian Classifier ισοκατανεμημένες κλάσεις

Για κάθε κλάση i, το discriminant function είναι

$g_i(x) = \text{constant} - \frac{1}{2}(x-\mu_i)^T(x-\mu_i) =$

ίδιο για κάθε κλάση

$= -\frac{1}{2}x^T x + \frac{1}{2}x^T \mu_i + \frac{1}{2}\mu_i^T x - \frac{1}{2}\mu_i^T \mu + c$

$= \text{constant} - \frac{1}{2}x^T x + \mu_i^T x - \frac{1}{2}\mu_i^T \mu$

Class 1: $g_1(x) = \left[c - \frac{1}{2}x^T x\right] + x_1 + x_2 - x_3 - \frac{1}{2}\begin{bmatrix} 1 & 1 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$

$= \boxed{\left[c - \frac{1}{2}x^T x\right] + x_1 + x_2 - x_3 - 3/2}$

Class 2: $g_2(x) = \left[c - \frac{1}{2}x^T x\right] - x_1 + x_2 - x_3 - \frac{1}{2}\begin{bmatrix} -1 & 1 & -1 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}$

$= \boxed{\left[c - \frac{1}{2}x^T x\right] - x_1 + x_2 - x_3 + 1/2}$

Class 3: $g_3(x) = \boxed{-x_2 - x_3 - 1 + \left(c - \frac{1}{2}x^T x\right)}$

Σύμφωνα με το Bayes, τα ακόλουθα πρέπει να ισχύουν για τα decision boundaries:

$g_1(x) > g_2(x) \Leftrightarrow x_1 + \cancel{x_2} - \cancel{x_3} + 1/2 > -x_1 + \cancel{x_2} - \cancel{x_3} + 1/2 \Leftrightarrow \boxed{x_1 > 0}$ &

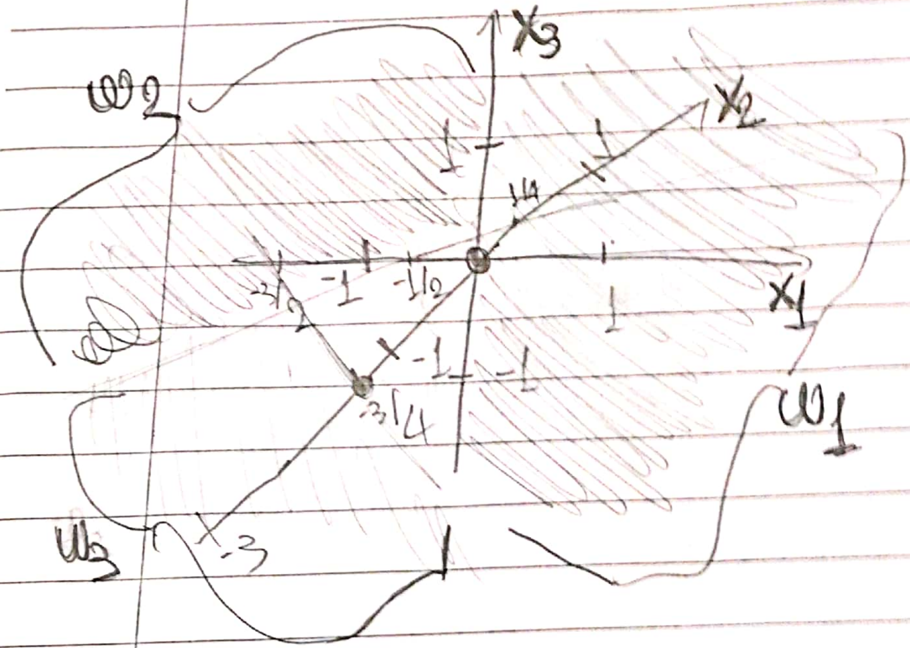
$g_1(x) > g_3(x) \Leftrightarrow x_1 + x_2 - \cancel{x_3} + 1/2 > -x_2 - \cancel{x_3} - 1 \Leftrightarrow \boxed{x_1 + 2x_2 + 3/2 > 0}$

Για την ω_2:

- $g_2(x) > g_1(x) \Leftrightarrow x_1 < 0$ &
- $g_2(x) > g_3(x) \Leftrightarrow -x_1 + x_2 \cancel{- x_3} + 1/2 > -x_2 \cancel{- x_3} + 1 \Leftrightarrow -x_1 + 2x_2 > 1/2$

Για την ω_3:

- $g_3(x) > g_1(x) \Leftrightarrow x_1 + 2x_2 + 3/2 < 0$ &
- $g_3(x) > g_2(x) \Leftrightarrow -x_1 + 2x_2 - 1/2 < 0$



ω_3 graphs

$x_1 = 0$ $\quad 2x_2 = -3/2$

$x_2 = 0$ $\quad x_1 = -3/2$

$x_1 = 0$ $\quad 2x_2 = 1/2 \Rightarrow x_2 = 1/4$

$x_2 = 0$ $\quad -x_1 = 1/2 \Rightarrow x_1 = -1/2$

ω_2 graphs

- $x_1 < 0$

- $x_1 = 0$ $\quad x_2 = 1/4$

$x_2 = 0$ $\quad x_1 = -1/2$

ω_1 graphs

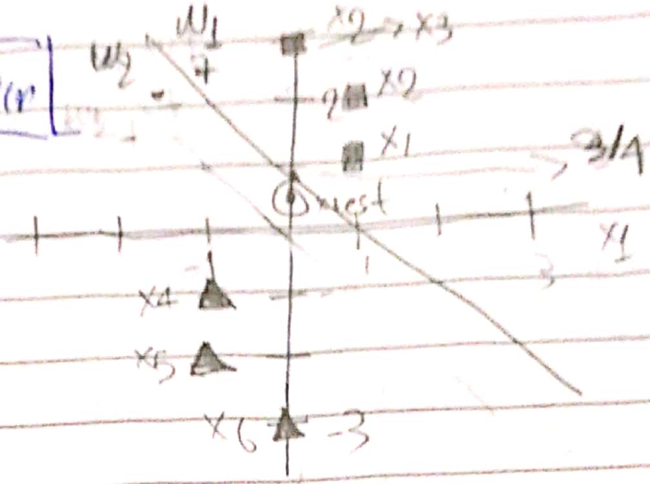
- $x_1 > 0$

- $x_1 = 0$ & $x_2 = -3/4$

$x_2 = 0$ $\quad x_1 = -3/2$

3 w_1: $\begin{bmatrix}1\\1\end{bmatrix}$, $\begin{bmatrix}1\\2\end{bmatrix}$, $\begin{bmatrix}0\\3\end{bmatrix}$ και w_2: $\begin{bmatrix}-1\\-1\end{bmatrix}$, $\begin{bmatrix}-1\\-2\end{bmatrix}$, $\begin{bmatrix}0\\-3\end{bmatrix}$, ενώ x_{test}: $\begin{bmatrix}0\\1/2\end{bmatrix}$

3.A Visualization



- Οι δύο κλάσεις (▲ και ■) είναι προφανώς γραμμικά διαχωρίσιμες
- Η ευθεία $x_2 + x_1 = 3/4$ είναι μία τέτοια ευθεία, π.χ.

$$g(x) = x_1 + x_2 - 3/4$$ της μορφής $g(x) = w_1 x_1 + w_2 x_2 + w_0$

όπου $w_1 = 1$, $w_2 = 1$, $w_0 = -3/4$

(δηλαδή $w = \begin{bmatrix}1\\1\\-3/4\end{bmatrix}$)

... $g(x_{test}) = g([0 \ 1/2]^T) = 0 + 1/2 - 3/4 = -1/4 < 0 \Rightarrow x_{test} \in w_2$

3.B **1-NN kNN(k=1)** Χρησιμοποιώντας το Manhattan Distance, υπολογίζω την απόσταση ως προς το x_1 pattern το οποίο φαίνεται πιο κοντά.

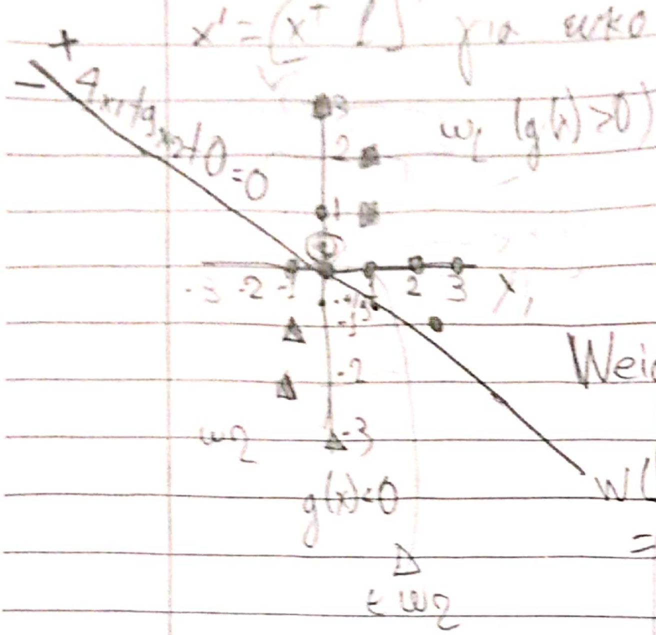
$$\text{Manhattan}(x_{test}, x_1) = |1-0| + |1-1/2| = 1.5$$

$x_1 \in w_1$, άρα $x_{test} \in w_1$ επίσης, καθώς ο πλησιέστερος γείτονας είναι το x_1.

3.c PERCEPTRON (batch-training) με $w(0) = [0, 0, 0]^T$

ρ = learning rate = arbitrary

$x' = (x^T \; 1)^T$ για ευκολία



Η αρχική ευθεία είναι η $g(x) = w^T(0) \cdot x_i' = 0$ η οποία ταξινομεί όλα τα πρότυπα λάθος

Weights Update

$$w(1) = w(0) + \rho \cdot (x_1' + x_2' + x_3') - \rho (x_4' + x_5' + x_6')$$

$$= \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix} - \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} - \begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ -3 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 4 \\ 9 \\ 0 \end{bmatrix} = w(1) \rightarrow w^T(1)\, x_i' = \begin{cases} \rightarrow 13 > 0 \\ \rightarrow 22 > 0 \quad \in \omega_1 \\ \rightarrow 27 > 0 \\ -13 < 0 \\ -22 < 0 \\ -27 < 0 \quad \in \omega_2 \end{cases}$$

Όλα τα πρότυπα ταξινομούνται σωστά με ένα update και $\rho = 1$

Τέλος $w^T(1) \cdot x_{test} = 4 \cdot 0 + 9 \cdot 1/2 + 0 \cdot 1 = \frac{9}{2} > 0$ → το οποίο είναι λάθος → misclassification on test set ∇

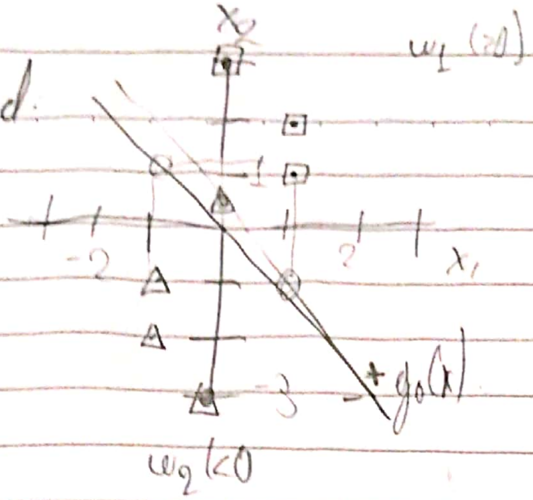
Το decision boundary είναι $g(x) = 4x_1 + 9x_2 + 0 = 0$

σημεία: Για $x_1 = 0$, $x_2 = 0$

Για $x_1 = 1$, $x_2 = -4/9 = -0.444$ (βλ. σχήμα πάνω)

3.D SVM

Visualized.



$x_1=1,\ x_2=-1$

$x_1=-1,\ x_2=1$

Προφανώς, μία ευθεία σαν τη $g_0(x)$ επιτυγχάνει το μέγιστο περιθώριο διαχωρισμού

$\lambda = \frac{1+1}{-1-1} = -1$, $x_2 - 1 = -1(x_1+1) \Rightarrow x_2 = -x_1 \cancel{-1} + 1 = x_1 + x_2 = 0$

οπότε θέλω $w_D^T \cdot x_1' = 1$ κ' $w_D^T \cdot x_4' = -1$.

Αυτό επιτυγχάνεται για $A = B = \frac{1}{2}$, συνεπώς $g(x) = \frac{1}{2}x_1 + \frac{1}{2}x_2$

Δηλαδή, $w_D = \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \end{bmatrix}$ ✓

Συνολικά, 2 σημεία εκπαίδευσης βρίσκονται στο όριο του περιθωρίου, το x1 και το x4.

$\underline{w}_D^T x_1' = 1$ και $\underline{w}^T x_4' = -1$

Τώρα πρέπει να εκφράσω τα w_{D1} και w_{D2} ως γραμμικό συνδυασμό των 2 παραπάνω σημείων, x1 και x4.

Συνεπώς $\begin{bmatrix} 0.5 \\ 0.5 \\ 0 \end{bmatrix} = \lambda_1 \cdot \underset{+1}{y_1} x_1 + \lambda_4 \cdot \underset{-1}{y_4} x_4 \Leftrightarrow$

$\Leftrightarrow \begin{bmatrix} 0.5 \\ 0.5 \\ 0 \end{bmatrix} = \lambda_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \lambda_4 \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \Leftrightarrow$

$\lambda_1 + \lambda_4 = 0,5$ (1) κ' $\lambda_1 - \lambda_4 = 0 \Rightarrow \lambda_1 = \lambda_4$

Συνεπώς, $2\lambda_1 = 0.5 \Rightarrow \boxed{\lambda_1 = \lambda_4 = 0,25}$

margin $= \frac{2}{\|w_D\|} = \frac{2}{\sqrt{(\frac{1}{2})^2 + (\frac{1}{2})^2}} = \frac{2}{\sqrt{\frac{1}{4}}} = 2.82$ (Προφανώς, $w_0 = 0$)

$g(x_{test}) = 0 \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} > 0$

$x_{test} \in \omega_1$

3.

1.

3.E. L.S.E

Σχηματίζω τον extended πίνακα X και y

$$X=\begin{bmatrix}1&1&1\\1&2&1\\0&3&1\\-1&-1&1\\-1&-2&1\\0&-3&1\end{bmatrix},\quad y=\begin{bmatrix}1\\1\\1\\-1\\-1\\-1\end{bmatrix}$$

Βρίσκω το $X^TX=\begin{bmatrix}1&1&0&-1&-1&0\\1&2&3&-1&-2&-3\\1&1&1&1&1&1\end{bmatrix}\begin{bmatrix}1&1&1\\1&2&1\\0&3&1\\-1&-1&1\\-1&-2&1\\0&-3&1\end{bmatrix}$

3 6   6 3

$$\Rightarrow X^TX=\begin{bmatrix}4&6&0\\6&28&0\\0&0&6\end{bmatrix}$$

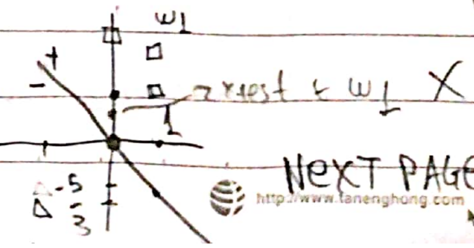
Βρίσκω $X^Ty=\begin{bmatrix}1&1&0&-1&-1&0\\1&2&3&-1&-2&-3\\1&1&1&1&1&1\end{bmatrix}\begin{bmatrix}1\\1\\1\\-1\\-1\\-1\end{bmatrix}=\begin{bmatrix}4\\12\\0\end{bmatrix}$

3x6   6x1

$M^{-1}=\frac{1}{\det(M)}\times Adj(M)$ // $\det(X^TX)=(28\cdot6-0)\cdot4+6(36-0)+0=885$

πράξεις

$$w_E=(X^TX)^{-1}X^Ty=\frac{1}{19}\begin{bmatrix}10\\6\\0\end{bmatrix}$$

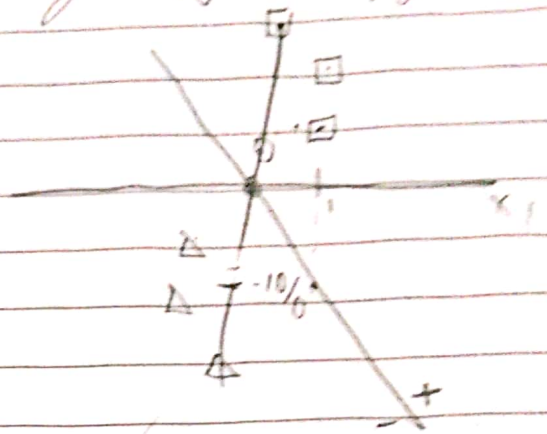


NEXT PAGE

Συνεπώς η ευθεία είναι η $g(x) = \frac{10}{19} x_1 + \frac{6}{19} x_2 + 0$

$x_1 = 0 \Rightarrow x_2 = 0$

$x_1 = 1 \Rightarrow \frac{6}{19} x_2 = -\frac{10}{19} \Rightarrow x_2 = \frac{-10}{6}$ $\Rightarrow$



Παρατηρούμε πως το $g(x_{test}) > 0$, δηλαδή το ξ ταξινομείται στην κλάση 1

$$\text{Total error} = \sum_{i=1}^{N=6} (y_i - x_i^T w_E)^2 = 2\left(1 - \frac{1}{19}(10+6)\right)^2 + 2\left(1 - \frac{1}{19}(10+12)\right)^2 + 2\left(1 + \frac{18}{19}\right)^2 \simeq 7.68$$

3.f PCA+1NN

Παρατηρώ ότι η μέση τιμή είναι 0 καθώς

$$\frac{1}{6}\sum_{i=1}^{N=6} x_i = \frac{1}{6}\left(\begin{bmatrix}1\\1\end{bmatrix}+\begin{bmatrix}-1\\-1\end{bmatrix}+\begin{bmatrix}1\\2\end{bmatrix}+\begin{bmatrix}-1\\-2\end{bmatrix}+\begin{bmatrix}0\\3\end{bmatrix}+\begin{bmatrix}0\\-3\end{bmatrix}\right)=\frac{1}{6}\cdot 0=0$$

$$X=\begin{bmatrix}1 & 1 & 0 & -1 & -1 & 0\\ 1 & 2 & 3 & -1 & -2 & -3\end{bmatrix}_{2\times 6}$$

ενώ $R=\frac{1}{6}XX^T=\frac{1}{6}\begin{bmatrix}4 & 6\\ 6 & 28\end{bmatrix}$

Βρίσκω τις ιδιοτιμές $(\lambda-4)(\lambda-28)-36=0 \Rightarrow$
$\Rightarrow \lambda_1=29.41$ και $\lambda_2=2.58$.

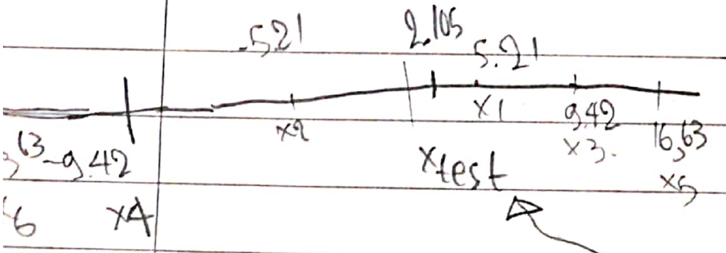
Επιλέγω τη μεγαλύτερη δηλαδή τη λ_1

$$\begin{bmatrix}4 & 6\\ 6 & 28\end{bmatrix}\cdot\begin{bmatrix}x_1\\x_2\end{bmatrix}=29.41\begin{bmatrix}x_1\\x_2\end{bmatrix} \Leftrightarrow \begin{matrix}4x_1+6x_2=29.41x_1\\ 6x_1+28x_2=29.41x_2\end{matrix}$$

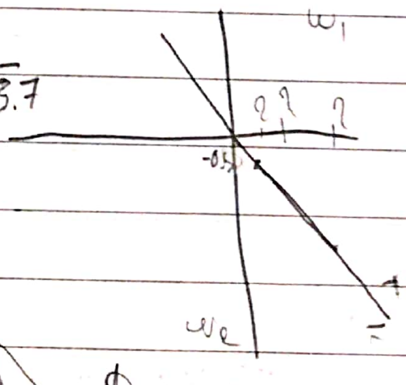
$\Leftrightarrow 4{,}21x_1-x_2=0$ δηλαδή $w_{PCA}=\begin{bmatrix}4{,}2\\-1\\0\end{bmatrix}$ ευθεία

κατά μήκος δηλαδή του διανύσματος $w_{proj}=\begin{bmatrix}1\\4{,}21\\0\end{bmatrix}$

Οι προβολές των σημείων θα είναι $\frac{w^T_{proj}x_i}{\|w_{proj}\|}$



$/\sqrt{18.7}$



Προφανώς χρησιμοποιώντας 1-NN, θα πάει στην κλάση του x_1, δηλαδή ω_1.

$\boxed{3.g}$ **LDA + 1NN**

Οι κλάσεις είναι ισοπίθανες άρα $P(\omega_1)=P(\omega_2)=1/2$

Βρίσκω τις μέσες τιμές των κλάσεων

$$\mu_1 = \frac{1}{3}\sum_{1}^{3} x_i = \frac{1}{3}\left(\begin{bmatrix}1\\1\end{bmatrix}+\begin{bmatrix}1\\2\end{bmatrix}+\begin{bmatrix}0\\3\end{bmatrix}\right) = \begin{bmatrix}2\\6\end{bmatrix}\cdot\frac{1}{3} = \begin{bmatrix}2/3\\2\end{bmatrix} \quad \text{κ} \quad \sum_{1}^{N} x_i = \begin{bmatrix}0\\0\end{bmatrix}$$

$$\mu_2 = \frac{1}{3}\sum_{4}^{6} x_i = \frac{1}{3}\left(\begin{bmatrix}-1\\-1\end{bmatrix}+\begin{bmatrix}-1\\-2\end{bmatrix}+\begin{bmatrix}0\\-3\end{bmatrix}\right) = \begin{bmatrix}-2\\-6\end{bmatrix}\cdot\frac{1}{3} = \begin{bmatrix}-2/3\\-2\end{bmatrix}$$

Για να βρω το Σ:

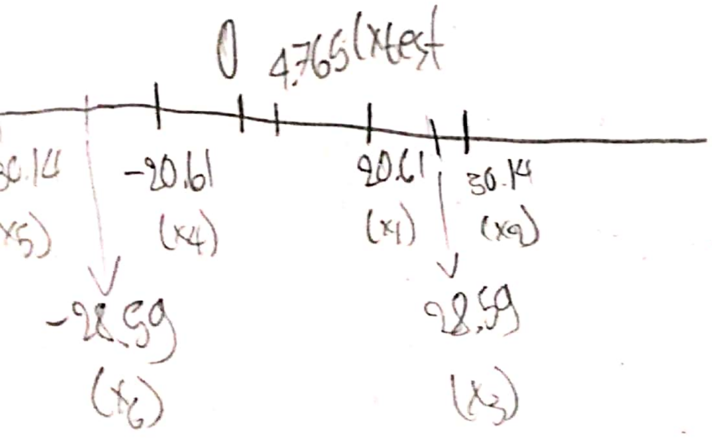
$$\begin{bmatrix}x_1-\mu_1, & x_2-\mu_1, & x_3-\mu_1\end{bmatrix} = \begin{bmatrix}1/3 & 1/3 & -2/3\\ -1 & 0 & 1\end{bmatrix} \Rightarrow \Sigma_1 = \frac{1}{2}\underbrace{\begin{bmatrix}1/3 & 1/3 & -2/3\\ -1 & 0 & 1\end{bmatrix}}_{2\times3}\underbrace{\begin{bmatrix}1/3 & -1\\ 1/3 & 0\\ -2/3 & 1\end{bmatrix}}_{3\times2}$$

$$= \frac{1}{2}\begin{bmatrix}1.11 & -1\\ -1 & 2\end{bmatrix} = \Sigma_2 \Rightarrow \Sigma_1=\Sigma_2=\begin{bmatrix}0.55 & -0.5\\ -0.5 & 1\end{bmatrix}$$

Έχουμε $\Sigma_2=\Sigma_1$, άρα $S_w = P(\omega_1)\Sigma_1 + P(\omega_2)\Sigma_2 = \frac{1}{2}(2\Sigma_1) = \begin{bmatrix}0.55 & -0.5\\ -0.5 & 1\end{bmatrix}$

Συνεπώς $w_{proj}^{(LDA)} = S_w^{-1}(\mu_1-\mu_2) = \begin{bmatrix}3.33 & 1.66\\ 1.66 & 1.83\end{bmatrix}\begin{bmatrix}4/3\\4\end{bmatrix} = \begin{bmatrix}11.08\\9.53\end{bmatrix}$

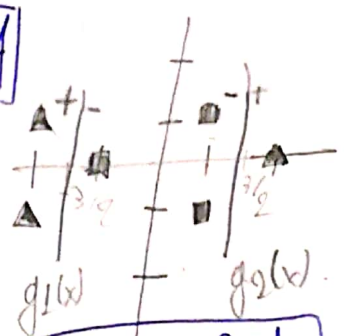
Βρίσκω τις προβολές σε 1-axis $(w^T_{proj}x_i/\|w_{proj}\|)$



114:6

Βλέπουμε πως $d(x_{test},x_1) < d(x_{te\ldots}$
σύμφωνα με Manhattan distan...
άρα $\underline{x_{test} \in \omega_1}$

4.A



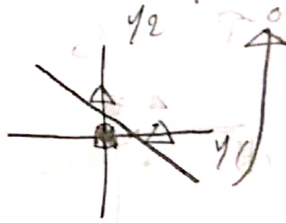
Προφανώς οι 2 κλάσεις ω_1 και ω_2 δεν είναι γραμμικά διαχωρίσιμες.

4.B PERCEPTRON WITH 2 LAYERS

Πρέπει να κάνουμε το εξής mapping:

x_1	x_2	y_1	y_2	
1	±1	0	0	ω_1
-1	0	0	0	ω_1
2	0	0	1	ω_2
-2	±1	1	0	ω_2

visual →



[illegible] που τα outputs

Θεωρώ τις 2 ευθείες (1ο layer) $x_2 = -\frac{3}{2} \Rightarrow x_1 + 0 \cdot x_2 + \frac{3}{2} = 0 = g_1(x)$

& $x_1 + 0 \cdot x_2 - \frac{3}{2} = 0 = g_2(x)$

... Βλέπουμε στο πάνω δεξιά γράφημα ότι τα στοιχεία στο νέο mapped χώρο μπορούν να διαχωριστούν από την ευθεία

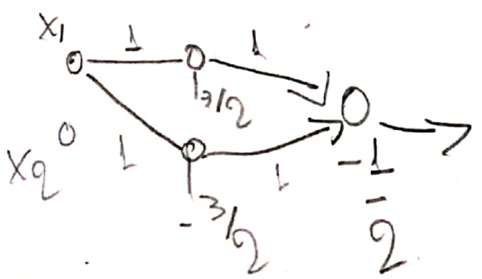
$y_1 = \frac{1}{2} \quad y_2 = 0$

$y_1 = 0 \quad y_2 = \frac{1}{2}$

$$\Rightarrow y_2 - 0 = \frac{1/2 - 0}{0 - 1/2}\left(y_1 - \frac{1}{2}\right) \Rightarrow y_2 = -y_1 + \frac{1}{2} \Rightarrow \underbrace{y_1 + y_2 - \frac{1}{2} = 0}_{g(y)}$$

(2ο layer)

Διάγραμμα του 2-layer perceptron



Το decision plane είναι όπως στο 4.A

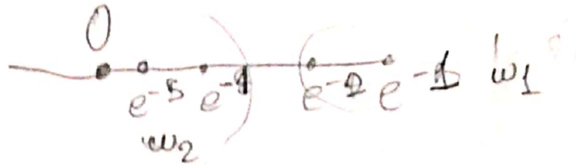
5 $x_1 = \begin{bmatrix}0\\0\end{bmatrix}$ ...

4.G RBF: Ας υποθέσουμε κέντρο της 'γκαουσιανής' το $c = \begin{bmatrix}0\\0\end{bmatrix}$ κ' $\sigma^2 = 1/2$

Άρα $y = \exp(-\|x\|)^2 = e^{(-x_1^2 - x_2^2)}$

Έχουμε λοιπόν το mapping:

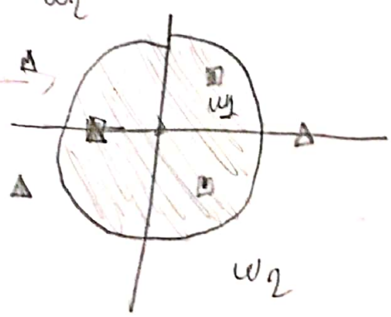
x_1	x_2	$\to y$	ΚΛΑΣΗ
1	± 1	e^{-2}	ω_1
-1	0	e^{-1}	ω_1
-2	± 1	e^{-5}	ω_2
-2	0	e^{-4}	ω_2



Από το δεξί διάγραμμα βλέπουμε ότι το threshold e^{-3} θα διαχωρίσει επιτυχώς τις 2 κλάσεις

Άρα $y \underset{\omega_2}{\overset{\omega_1}{\gtrless}} e^{-3} \Rightarrow e^{(-x_1^2 - x_2^2)} \underset{\omega_2}{\overset{\omega_1}{\gtrless}} e^{-3} \Rightarrow x_1^2 + x_2^2 \underset{\omega_1}{\overset{\omega_2}{\gtrless}} 3$

Συνεπώς:

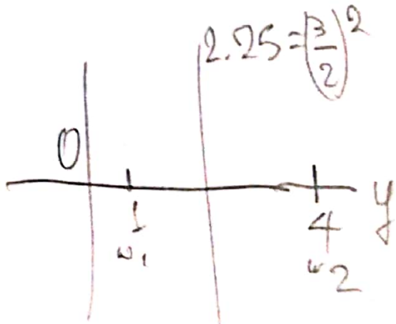


FD-POLYNOMIAL Δοκιμάζω mapping $y = x_1^2$

Έτσι έχω

x_1	x_2	$\to y$	ΚΛΑΣΗ
1	± 1	1	ω_1
-1	0	1	ω_1
2	0	4	ω_2
-2	± 1	4	ω_2

Συνεπώς $\Rightarrow$



threshold $2.25 = \left(\frac{3}{2}\right)^2$ μπορεί να διαχωρίσει τις κλάσεις

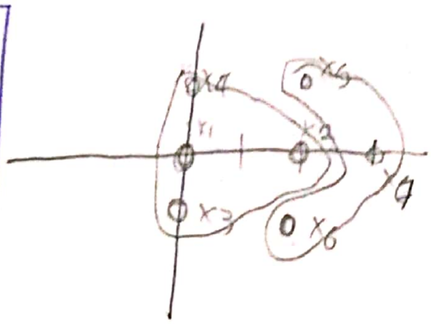
$y \underset{\omega_1}{\overset{\omega_2}{\gtrless}} \left(\frac{3}{2}\right)^2 \Rightarrow x_1^2 \underset{\omega_1}{\overset{\omega_2}{\gtrless}} \left(\frac{3}{2}\right)^2 \Leftrightarrow -\frac{3}{2} < x_1 < \frac{3}{2}$ για το ω_1

και $x_1 < -3/2$ κ' $x_1 > 3/2$ γ...

Scanned by CamScanner

6 $\boxed{5}$ $x_1=\begin{bmatrix}0\\0\end{bmatrix}, x_2=\begin{bmatrix}2\\0\end{bmatrix}, x_3=\begin{bmatrix}0\\-1\end{bmatrix}, x_4=\begin{bmatrix}0\\1\end{bmatrix}, x_5=\begin{bmatrix}2\\1\end{bmatrix}, x_6=\begin{bmatrix}2\\-1\end{bmatrix}, x_7=\begin{bmatrix}3\\0\end{bmatrix}$

$\boxed{5.A}$ BSAS, $\theta=2.1$
L_1 dist (Manhattan),
centroids
K_{max} = inf = [illegible]



STEP	CLUSTER
1	$\{\{x_1\}\}$

2 $d(x_2,x_1)=2-0+0-0=2<\theta$

Μπαίνει στο cluster και έχουν centroid $c_1=\frac{x_1+x_2}{2}=\begin{bmatrix}1\\0\end{bmatrix}$

$\{\{x_1,x_2\}\}$

---

3 $d(x_3,c_1)=|1-0|+|0-1|=2<\theta$

Μπαίνει στο cluster και έχει centroid $c_1'=\left(\begin{bmatrix}0\\-1\end{bmatrix}+\begin{bmatrix}1\\0\end{bmatrix}\right)/2$

Οπότε έχουμε $\{\{x_1,x_2,x_3\}\}$ $=\begin{bmatrix}0.5\\-0.5\end{bmatrix}=c_1'$

---

4 $d(x_4,c_1')=|0-0.5|+|1-(-0.5)|=2<\theta$

Μπαίνει στο cluster και έχουν centroid $c_1''=\left(\begin{bmatrix}0.5\\-0.5\end{bmatrix}+\begin{bmatrix}0\\1\end{bmatrix}\right)/2=\begin{bmatrix}0.25\\0.25\end{bmatrix}$

Οπότε έχουμε $\{\{x_1,x_2,x_3,x_4\}\}$

---

5 $d(x_5,c_1'')=|2-0.25|+|1-0.25|=1.75+0.75>\theta$

Οπότε δημιουργεί νέο cluster με $c_2=x_5$

Άρα έχουμε $\{\{x_1,x_2,x_3,x_4\},\{x_5\}\}$

---

6 $d(x_6,c_1'')=|2-0.25|+|-1-0.25|=1.75+1.25=3>\theta$

$d(x_6,c_2)=|2-2|+|-1-1|=2<\theta$

Άρα μπαίνει στο 2ο cluster και έχουν $c_2'=\left(\begin{bmatrix}2\\1\end{bmatrix}+\begin{bmatrix}2\\-1\end{bmatrix}\right)/2=\begin{bmatrix}2\\0\end{bmatrix}$

Δηλαδή έχουμε πλέον $\{\{x_1,x_2,x_3,x_4\}\{x_5,x_6\}\}$

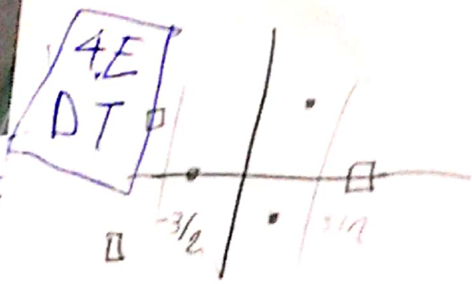
---

7 $d(x_7,c_1'')=|3-0.25|+|0-0.25|>\theta$

$d(x_7,c_2')=|3-2|+|0-0|=1<\theta$ → Άρα μπαίνει στο 2ο cluster με centroid $c_2''=\begin{bmatrix}2.5\\0\end{bmatrix}$

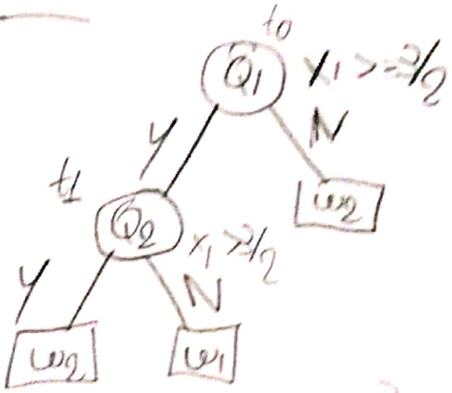
κ' έχουμε (x_1,x_2,x_3,x_4) και (x_5,x_6,x_7)

4.E DT



$Q_1: x_1 > -3/2$ και $Q_2: x_1 > 3/2$

Ένα πιθανό decision tree είναι το εξής:



και 4 δείγματα στην κλάση 1

$I_t = -\frac{2}{6}\log\frac{2}{6} \ldots = 0.26$

0.15 0.11

$I_{Q_1}(t_Y) = \ldots \frac{1}{\ldots} = 0.24$

$I_{Q_1}(t_N) = 0$

$I_{Q_2}(t_Y) = 0$

0.15

$I_{Q_2}(t_N) = \ldots = 0.28$

0.13

Scanned by CamScanner

5.B Init 5.A
K-Means
L_2 dist (Euclidean)

$d = \sqrt{(x_2-x_1)^2+(y_2-y_1)^2}$

Οι ομάδες που προέκυψαν στο 5.A είναι οι εξής: $\{\{x_1, x_2, x_3, x_4\}, \{x_5, x_6, x_7\}\}$

Τα κέντρα αυτών είναι $c_1 = \begin{bmatrix} 0.25 \\ 0.25 \end{bmatrix}$ και $c_2 = \begin{bmatrix} 2.5 \\ 0 \end{bmatrix}$

RE-CLUSTERING:

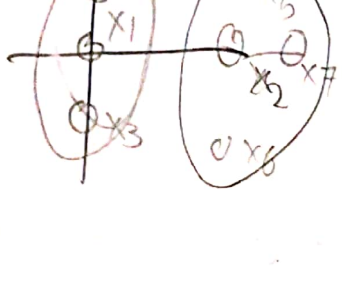
$d(x_1, c_1) = \sqrt{(0-0.25)^2+(0-0.25)^2} = 0.35$, $d(x_1, c_2) = \ldots = 2.5$ → $\boxed{x_1 \to c_1}$	
$d(x_2, c_1) = 1.76$, $d(x_2, c_2) = 0.5$ → $\boxed{x_2 \to c_2}$	$d(x_3, c_1) = 1.27$, $d(x_3, c_2) = 2.69$ → $\boxed{x_3 \to c_1}$
$d(x_4, c_1) = 3.75$, $d(x_4, c_2) = 4.71$ → $\boxed{x_4 \to c_1}$	$d(x_5, c_1) = 1.90$, $d(x_5, c_2) = 1.11$ → $\boxed{x_5 \to c_2}$
$d(x_6, c_1) = 2.15$, $d(x_6, c_2) = 1.11$ → $\boxed{x_6 \to c_2}$	$d(x_7, c_1) = 2.76$, $d(x_7, c_2) = 0.5$ → $\boxed{x_7 \to c_2}$

Άρα, τελικά έχουμε τα clusters
$\{x_1, x_3, x_4\}$ και $\{x_2, x_5, x_6, x_7\}$

με centroids

$$c_1' = \frac{x_1 + x_3 + x_4}{3} = \frac{\begin{bmatrix} 0 \\ -1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}}{3} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \kappa' \quad c_2' = \frac{x_2+x_5+x_6+x_7}{4}$$

$$\Rightarrow c_2' = \frac{\begin{bmatrix} 2 \\ 0 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ -1 \end{bmatrix} + \begin{bmatrix} 3 \\ 0 \end{bmatrix}}{4} = \frac{\begin{bmatrix} 9 \\ 0 \end{bmatrix}}{4} = \begin{bmatrix} 2.25 \\ 0 \end{bmatrix}$$



5.c Medoids for a) & b)

ως medoid ορίζεται το σημείο στην ομάδα (cluster) που παρουσιάζει το ελάχιστο average dissimilarity. ως dissimilarity εδώ παίρνουμε να έχουμε το L_1 ή L_2 distance μεταξύ όλων των προτύπων ενός cluster.

Προφανώς, για τις ομάδες του 5.a) έχουμε:

$medoid_{c1} = x_1$ και $medoid_{c2} = x_7$.

Ομοίως, για τις ομάδες του 5.b) έχουμε:

$medoid_{c1} = x_1$ και $medoid_{c2} = x_2$

5.D. Agglomerative clustering with L_1 (Manhattan) & Single vs complete link

Υπολογίζω τον πίνακα ομοιότητας για τα 7 διανύσματα.

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
x_1	0	2	1	1	3	3	3
x_2	2	0	3	3	2	1	1
x_3	1	3	0	2	4	3	4
x_4	1	3	2	0	2	4	4
x_5	3	2	4	2	0	2	2
x_6	3	1	3	4	2	0	2
x_7	3	1	4	4	2	2	0

SINGLE LINK ($d = d_{min}$)

ΟΜΑΔΟΠΟΙΗΣΗ x_1, x_3 ($d=1$)

$\{x_1, x_3\}$	x_2	x_4	x_5	x_6	x_7
0	2	1	3	3	3
2	0	3	2	1	1
1	3	0	2	4	4
3	2	2	0	2	2
3	1	4	2	0	2
3	1	4	2	2	0

COMPLETE LINK ($d = d_{max}$)

ΟΜΑΔΟΠΟΙΗΣΗ x_1, x_3 ($d=1$)

	x_2	$\{x_1,x_3\}$	x_4	x_5	x_6	x_7
x_2	0	3	3	2	1	1
$\{x_1, x_3\}$	3	0	2	4	3	4
x_4	3	2	0	2	4	4
x_5	2	4	2	0	2	2
x_6	1	3	4	2	0	2
x_7	1	4	4	2	2	0

(συνέχεια 5.D)

ΟΜΑΔΟΠΟΙΗΣΗ x_4 με x_1, x_3 (d=1)

	$\{x_1, x_3, x_4\}$	x_2	x_5	x_6	x_7
$\{x_1, x_3, x_4\}$	0	2	3	3	3
x_2	2	0	2	1	1
x_5	3	2	0	2	2
x_6	3	1	2	0	2
x_7	3	1	2	2	0

ΟΜΑΔΟΠΟΙΗΣΗ x_2 με x_6 (d=1)

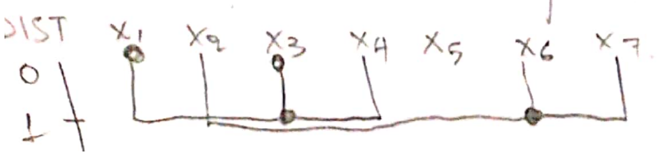
	x_1, x_3, x_4	x_2, x_6	x_5	x_7
x_1, x_3, x_4	0	2	3	3
x_2, x_6	2	0	2	1
x_5	3	2	0	2
x_7	3	1	2	0

ΟΜΑΔΟΠΟΙΗΣΗ x_7 με $\{x_2, x_6\}$ d=1

	x_1, x_3, x_4	x_2, x_6, x_7	x_5
x_1, x_3, x_4	0	2	3
x_2, x_6, x_7	2	0	2
x_5	3	2	0

ΟΜΑΔΟΠΟΙΗΣΗ x_2, x_6, x_7 με x_1, x_3, x_4 (d=2)

	$x_1, x_2, x_3, x_4, x_6, x_7$	x_5
$x_1, x_2, x_3, x_4, x_6, x_7$	0	2
x_5	2	0



ΟΜΑΔΟΠΟΙΗΣΗ x_2 με x_6 (d=1)

	x_1, x_3	x_4	x_5	x_2, x_6	x_7
x_1, x_3	0	2	4		
x_4	2	0	2	4	4
x_5	4	2	0	2	2
x_2, x_6	7	4	2	0	2
x_7	4	4	2	2	0

ΟΜΑΔΟΠΟΙΗΣΗ x_4 με x_1, x_3 (d=2)

	x_1, x_3, x_4	x_5	x_2, x_6	x_7
x_1, x_3, x_4	0	4	3	4
x_5	4	0	2	2
x_2, x_6	3	2	0	2
x_7	4	2	2	0

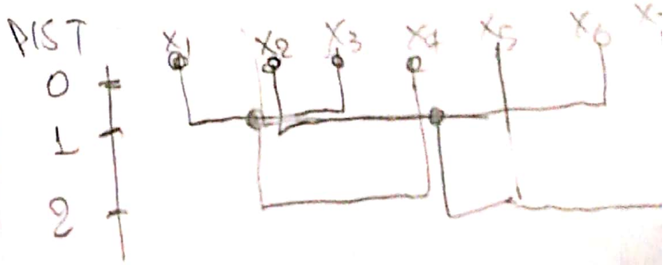
ΟΜΑΔΟΠΟΙΗΣΗ x_5 με x_2, x_6 d=2

	x_1, x_3, x_4	x_2, x_5, x_6	x_7
x_1, x_3, x_4	0	4	4
x_2, x_5, x_6	4	0	2
x_7	4	2	0

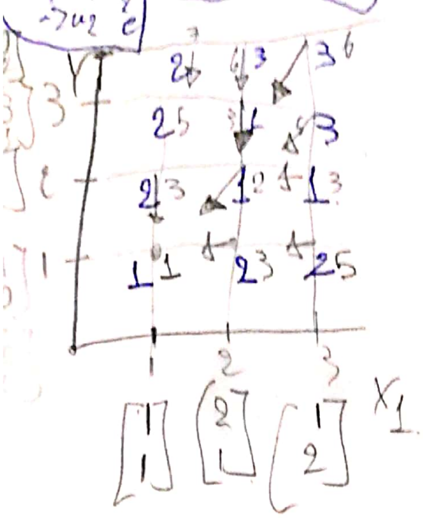
ΟΜΑΔΟΠΟΙΗΣΗ x_7 με x_2, x_5, x_6 (d=

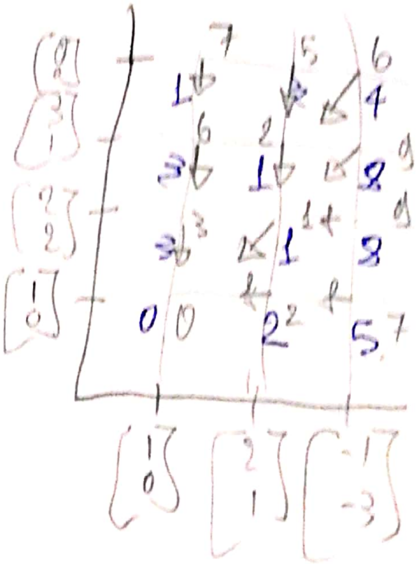
	x_1, x_3, x_4	x_2, x_5, x_6, x_7
x_1, x_3, x_4	0	4
x_2, x_5, x_6, x_7	4	0

ΔΕΝΤΡΟΓΡΑΜΜΑ

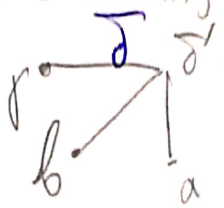


6. DTW TM L1 Sakoe-Chiba





Βρίσκω L_1 distance για κάθε xε x_1, x_2 με y (μικρές τιμές) σύμφωνα με Sakoe-Chiba:



$\delta' = \min\{\alpha, \beta, \gamma\} + \delta$

Σημειώνω το κόστος με το μολύβι καθώς και την κατεύθυνση το.

Το ολικό κόστος, σύμφωνα με Sakoe-Chiba, είναι το νούμερο με μολύβι που βρίσκεται στην πάνω δεξιά θέση

Δηλαδή $d(y, \omega_1) = 6 = d(y, \omega_2)$

Θα μπορούσε να γίνει classify σε οποιαδήποτε κλάση

Σημείωση: Το x_1 παρουσιάζει διπλότυπο το $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$, κάτι το οποίο μπορεί να απαλειφθεί, συνεπώς $X_1 = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$