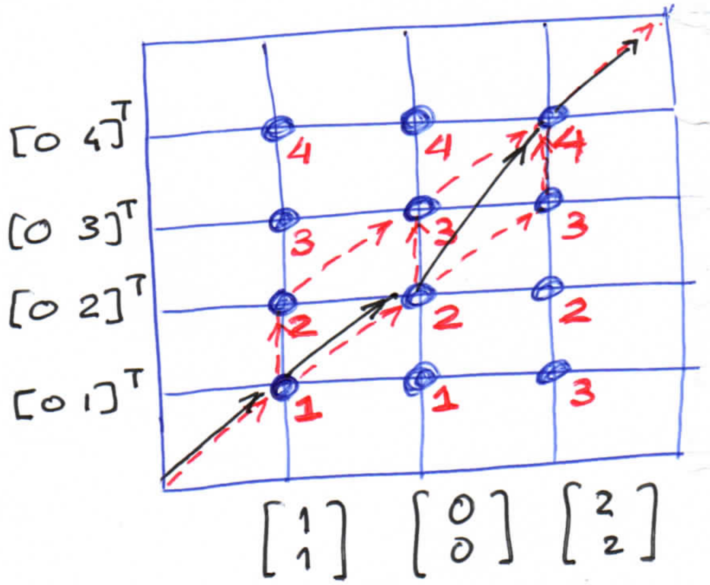
3 a

Υπολογίζουμε το πλέγμα αποστάσεων ως εξής:



- Τα 3 κόκκινα μονοπάτια έχουν όλα κόστος

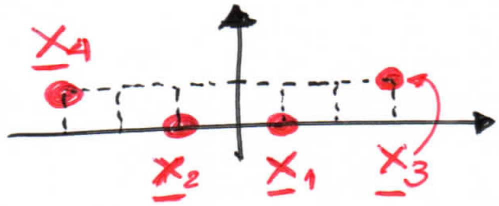
  $1 + 2 + 3 + 4 = 10$ και είναι

  σύμφωνα με τους περιορισμούς Sakoe-Chiba.

- Το μαύρο μονοπάτι τηρεί τους περιορισμούς Itakura, και το κόστος του είναι:

  $1 + 2 + 4 = 7$

3.b.



- Σχηματίζουμε τον πίνακα ανομοιότητας των δεδομένων:

$$\begin{array}{c|cccc} & \underline{x}_1 & \underline{x}_2 & \underline{x}_3 & \underline{x}_4 \\ \hline \underline{x}_1 & 0 & 2 & 3 & 5 \\ \underline{x}_2 & 2 & 0 & 5 & 3 \\ \underline{x}_3 & 3 & 5 & 0 & 6 \\ \underline{x}_4 & 5 & 3 & 6 & 0 \end{array}$$

- Η μικρότερη απόσταση είναι μεταξύ των $\underline{x}_1, \underline{x}_2$ (2)

  άρα ο καινούργιος πίνακας είναι ο (με χρήση αλγορίθμου υπολογισμού αποστάσεων SINGLE LINK)

$$\begin{array}{c|ccc} & \{\underline{x}_1, \underline{x}_2\} & \underline{x}_3 & \underline{x}_4 \\ \hline \{\underline{x}_1, \underline{x}_2\} & 0 & 3 & 3 \\ \underline{x}_3 & 3 & 0 & 6 \\ \underline{x}_4 & 3 & 6 & 0 \end{array}$$

πχ

$$d(\{\underline{x}_1, \underline{x}_2\}, \{\underline{x}_4\}) = \min\{d(\underline{x}_1, \underline{x}_4), d(\underline{x}_2, \underline{x}_4)\} = \min\{5, 3\} = 3$$

- Η μικρότερη απόσταση είναι μεταξύ του $\{\underline{x}_1, \underline{x}_2\}$ κ του $\underline{x}_3$ ή $\underline{x}_4$ (3)

3. b. ΣΥΝΕΧΕΙΑ

- Ας πούμε ότι ομαδοποιούμε το $\underline{x}_3$ με το $\{\underline{x}_1, \underline{x}_2\}$ οπότε ο καινούργιος πίνακας ανομοιότητας γίνεται ο:

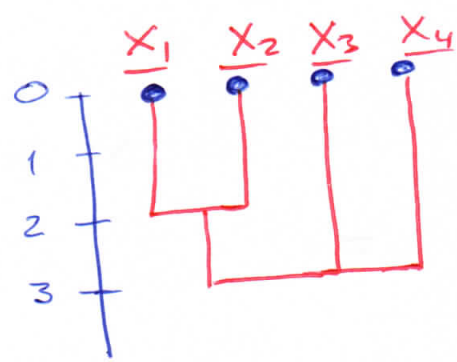
$$\begin{array}{c} \\ \{\underline{x}_1, \underline{x}_2, \underline{x}_3\} \\ \underline{x}_4 \end{array} \begin{array}{c} \begin{array}{cc} \{\underline{x}_1, \underline{x}_2, \underline{x}_3\} & \underline{x}_4 \end{array} \\ \begin{bmatrix} 0 & 3 \\ 3 & 0 \end{bmatrix} \end{array}$$

$$\pi x = \min\{d(\underline{x}_4, \{\underline{x}_1, \underline{x}_2\}), d(\underline{x}_4, \underline{x}_3)\}$$

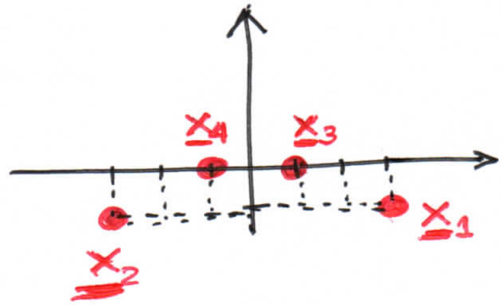
$$= \min\{3, 3\} = 3$$

- Τέλος, το $\underline{x}_4$ ομαδοποιείται με το $\{\underline{x}_1, \underline{x}_2, \underline{x}_3\}$ με απόσταση 3.

- Το δεντρόγραμμα είναι:



3.c.



Σειρά: $\underline{x}_1, \underline{x}_2, \underline{x}_3, \underline{x}_4$

1^{η} επανάληψη: $\{\{x_1\}\}$ (αρχικοποιηση)

2^{η} επανάληψη: $d(\underline{x}_1, \underline{x}_2) = \|\underline{x}_1 - \underline{x}_2\|_1 = 6 > 3.5$

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

3^{η} επανάληψη: $d(\underline{x}_3, \underline{x}_1) = \|\underline{x}_3 - \underline{x}_1\|_1 = 3$ $\downarrow$ CLUSTER 1

$d(\underline{x}_3, \underline{x}_2) = \|\underline{x}_3 - \underline{x}_2\|_1 = 4$ $\hookrightarrow$ CLUSTER 2

$\xrightarrow{MIN}$ $3 < 3.5$

$\rightsquigarrow \{\{\underline{x}_1, \underline{x}_3\}, \{\underline{x}_2\}\}$

CLUSTER 1 — CLUSTER 2

ΚΕΝΤΡΟ: $\frac{1}{2}(\underline{x}_1 + \underline{x}_3) = [2, -0.5]^T = \underline{c}_1$

4^{η} επανάληψη:

$d(\underline{x}_4, \underline{c}_1) = \|\underline{x}_4 - \underline{c}_1\|_1 = 3.5$

$d(\underline{x}_4, \underline{x}_2) = \|\underline{x}_4 - \underline{x}_2\|_1 = 3$ $\Rightarrow$

3 c - συνέχεια

$\Longrightarrow \min\{3, 3.5\} = 3 < 3.5$

$$\left\{ \{\underline{X}_1, \underline{X}_3\}, \{\underline{X}_2, \underline{X}_4\} \right\} \longleftarrow \text{ΤΕΛΙΚΗ ΟΜΑΔΟΠΟΙΗΣΗ}$$

ΣΕΙΡΑ: $\underline{X}_4, \underline{X}_3, \underline{X}_2, \underline{X}_1$

Ακολουθώντας παρόμοια συλλογιστική, καταλήγουμε στην ομαδοποίηση:

$$\left\{ \{\underline{X}_4, \underline{X}_3\}, \{\underline{X}_2\}, \{\underline{X}_1\} \right\}$$

που όπως βλέπουμε διαφέρει από την προηγούμενη.
