

Μάθημα 7

ΕΞΟΥΥΞΗ ΔΕΔΟΜΕΝΩΝ ΑΠΟ ΨΗΦΙΑΚΟ ΚΑΙ ΔΙΑΔΙΚΤΥΑΚΟ ΠΕΡΙΕΧΟΜΕΝΟ

Συσταδοποίηση Δεδομένων (Συνέχεια)

Ο Αλγόριθμος c-Means

Β. Η χρήση των Κέντρων των Συστάδων στον Προσδιορισμό των Συναρτήσεων Συμμετοχής

Στην περίπτωση αυτή είναι γνωστά τα κέντρα των συστάδων και θέλουμε να δούμε τα δεδομένα σε ποια συστάδα ανήκει το καθένα από αυτά

	\mathbf{x}	
	x_1	x_2
\mathbf{x}_1	10	9
\mathbf{x}_2	4	3.5
\mathbf{x}_3	9.5	11
\mathbf{x}_4	2	3
\mathbf{x}_5	12	9.5
\mathbf{x}_6	1	5.5

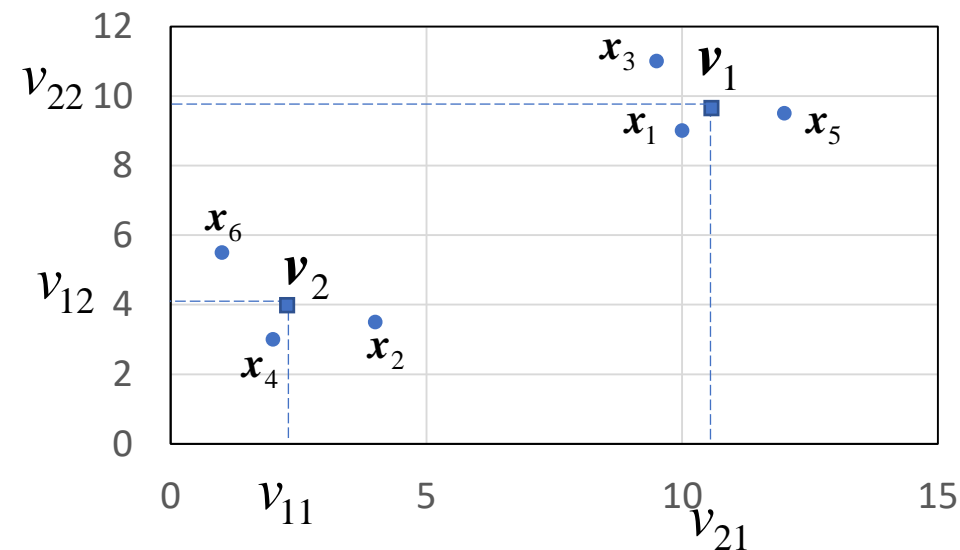
	\mathbf{v}	
	x_1	x_2
\mathbf{v}_1	10.5	9.83
\mathbf{v}_2	2.3	4

$$N = 6, \quad c = 2, \quad p = 2$$

$$k = 1, 2, \dots, N$$

$$i = 1, 2, \dots, c$$

$$j = 1, 2, \dots, p$$



Ο Αλγόριθμος c-Means

B. Η χρήση των Κέντρων των Συστάδων στον Προσδιορισμό των Συναρτήσεων Συμμετοχής

Στην περίπτωση αυτή είναι γνωστά τα κέντρα των συστάδων και θέλουμε να δούμε τα δεδομένα σε ποια συστάδα ανήκει το καθένα από αυτά

\xrightarrow{j}

\mathbf{x}

	x_1	x_2
\mathbf{x}_1	10	9
\mathbf{x}_2	4	3.5
\mathbf{x}_3	9.5	11
\mathbf{x}_4	2	3
\mathbf{x}_5	12	9.5
\mathbf{x}_6	1	5.5

$\downarrow k$

\xrightarrow{j}

\mathbf{v}

	x_1	x_2
\mathbf{v}_1	10.5	9.83
\mathbf{v}_2	2.3	4

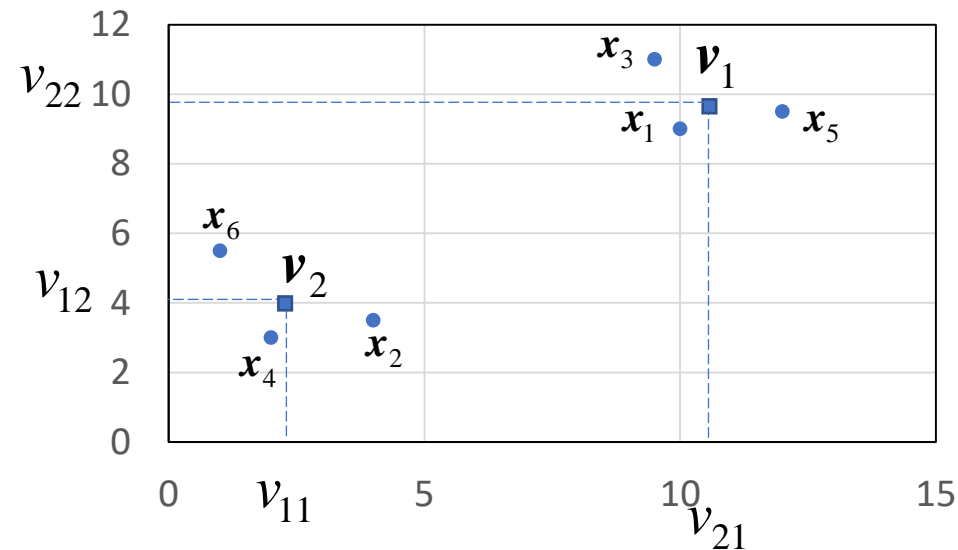
$\downarrow i$

$$N = 6, \quad c = 2, \quad p = 2$$

$$k = 1, 2, \dots, N$$

$$i = 1, 2, \dots, c$$

$$j = 1, 2, \dots, p$$



function D=Distance(N,c,p,x,v)

D=zeros(N,c);

```

for k=1:N
    for i=1:c
        s=0;
        for j=1:p
            s=s+(x(k,j)-v(i,j))^2;
        endfor
        D(k,i)=sqrt(s);
    endfor
endfor
    
```

\xrightarrow{i}

D

	v_{11}	v_{21}
\mathbf{x}_1	0.96897	9.181
\mathbf{x}_2	9.073	1.772
\mathbf{x}_3	1.5391	10.042
\mathbf{x}_4	10.904	1.044
\mathbf{x}_5	1.5359	11.151
\mathbf{x}_6	10.44	1.9849

$\downarrow k$

Ελάχιστο στοιχείο γραμμής 1 αλλιώς 0

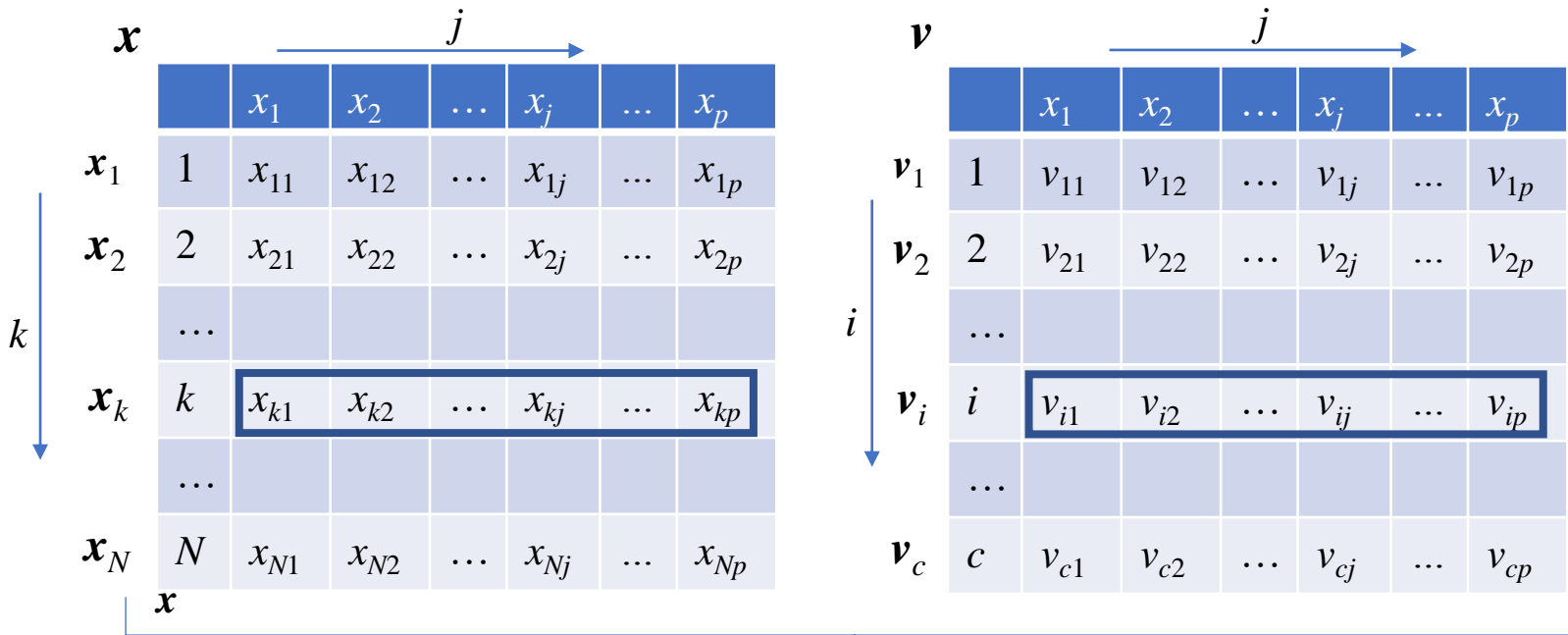


U

1	0
0	1
1	0
0	1
1	0
0	1

Ο Αλγόριθμος c-Means

Β. Η χρήση των Κέντρων των Συστάδων στον Προσδιορισμό των Συναρτήσεων Συμμετοχής



```
function D=Distance(N,c,p,x,v)
```

```
D=zeros(N,c);
```

```
for k=1:N
```

```
    for i=1:c
```

```
        s=0;
```

```
        for j=1:p
```

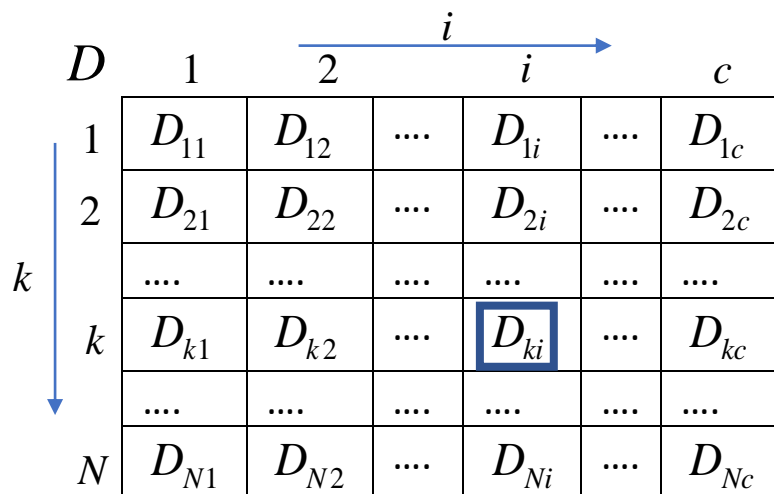
```
            s=s+(x(k,j)-v(i,j))^2;
```

```
        endfor
```

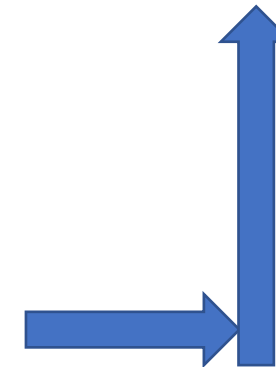
```
        D(k,i)=sqrt(s);
```

```
    endfor
```

```
endfor
```

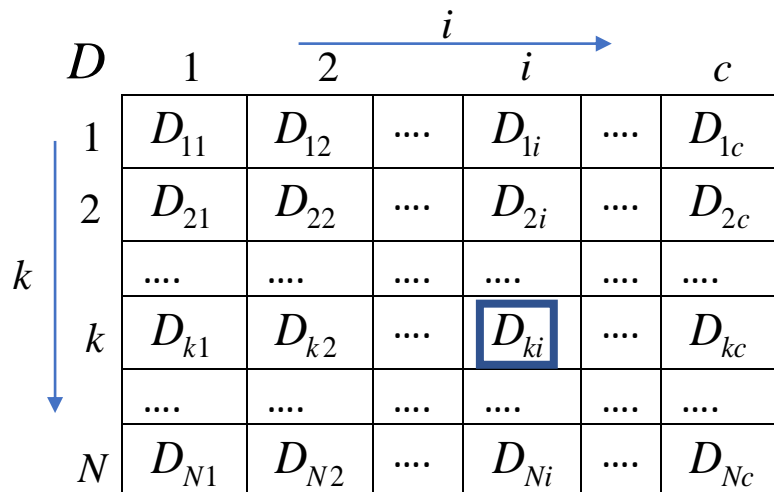
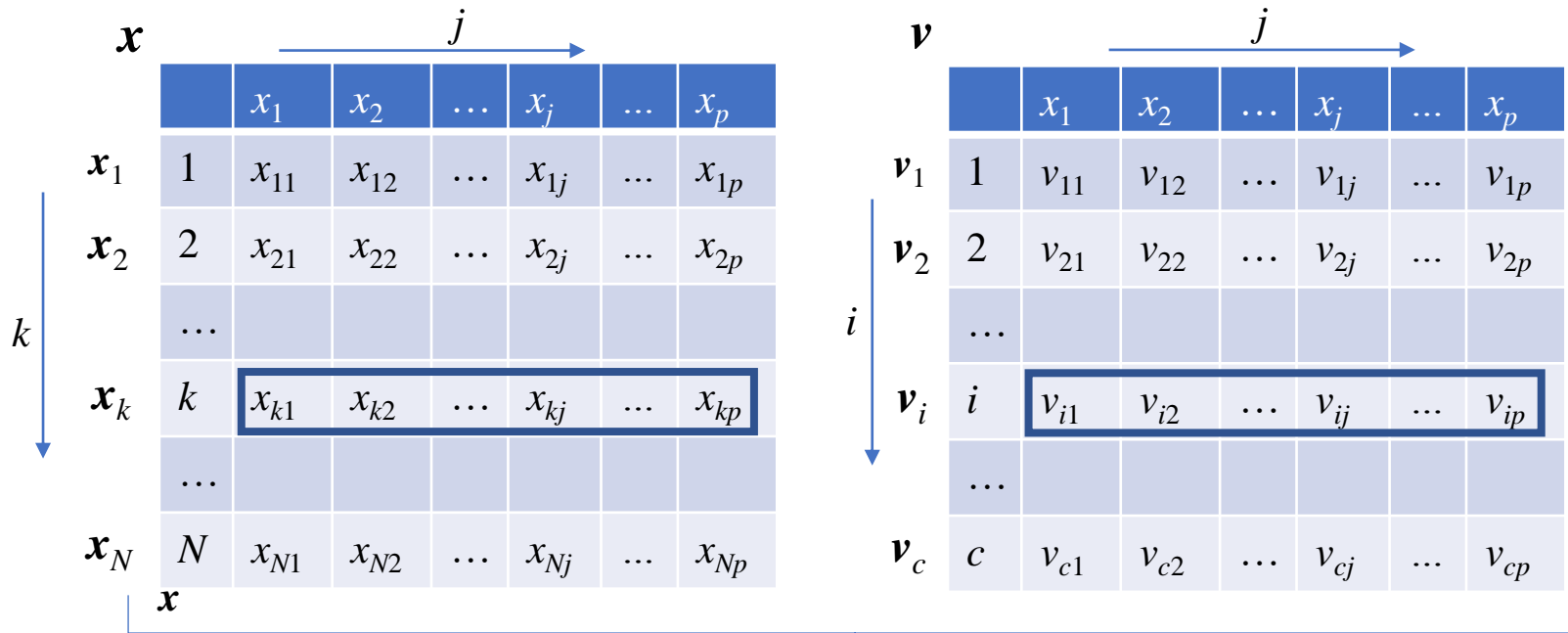


$$D_{ki} = \sqrt{\sum_{j=1}^p (x_{kj} - v_{ij})^2}$$



Ο Αλγόριθμος c-Means

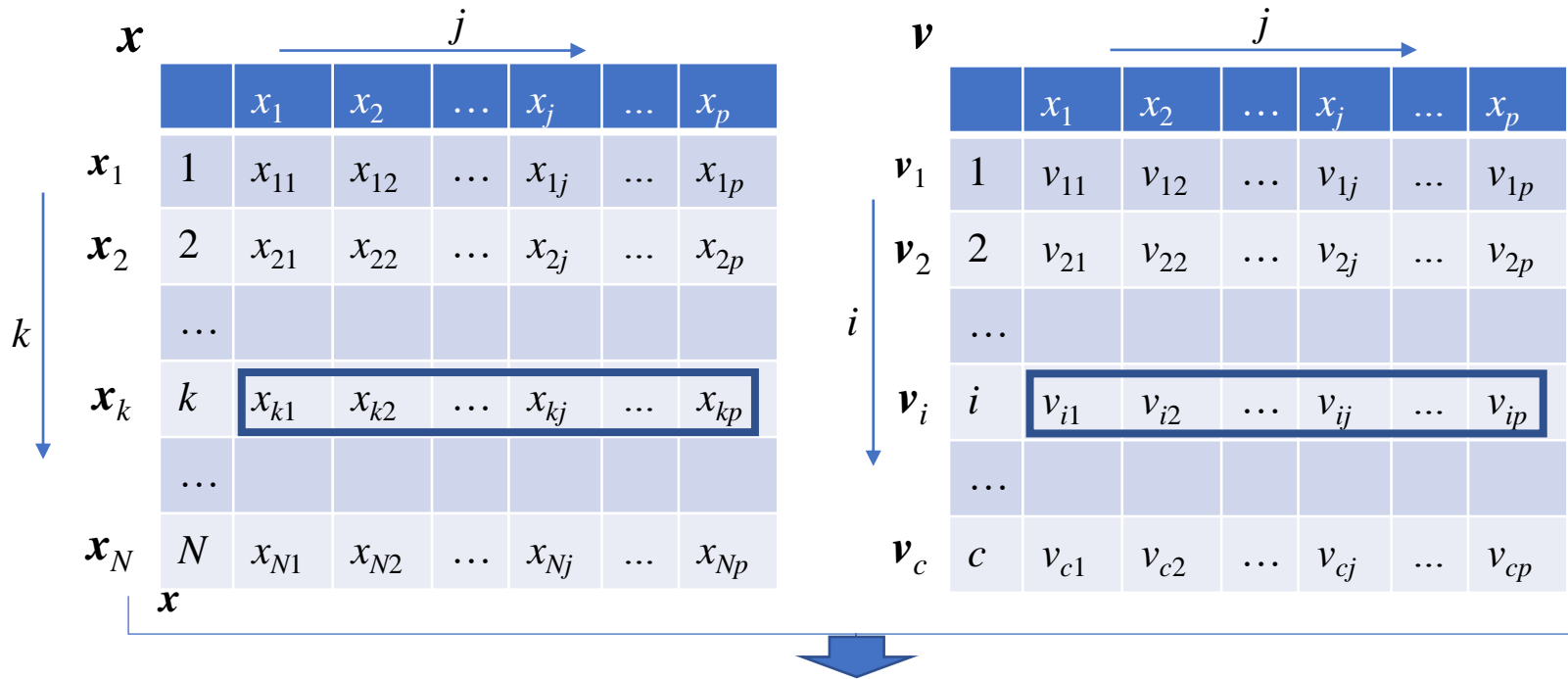
Β. Η χρήση των Κέντρων των Συστάδων στον Προσδιορισμό των Συναρτήσεων Συμμετοχής



Το διάνυσμα \mathbf{x}_k ανήκει στην συστάδα το κέντρο της οποίας απέχει λιγότερο από το \mathbf{x}_k

Ο Αλγόριθμος c-Means

Β. Η χρήση των Κέντρων των Συστάδων στον Προσδιορισμό των Συναρτήσεων Συμμετοχής

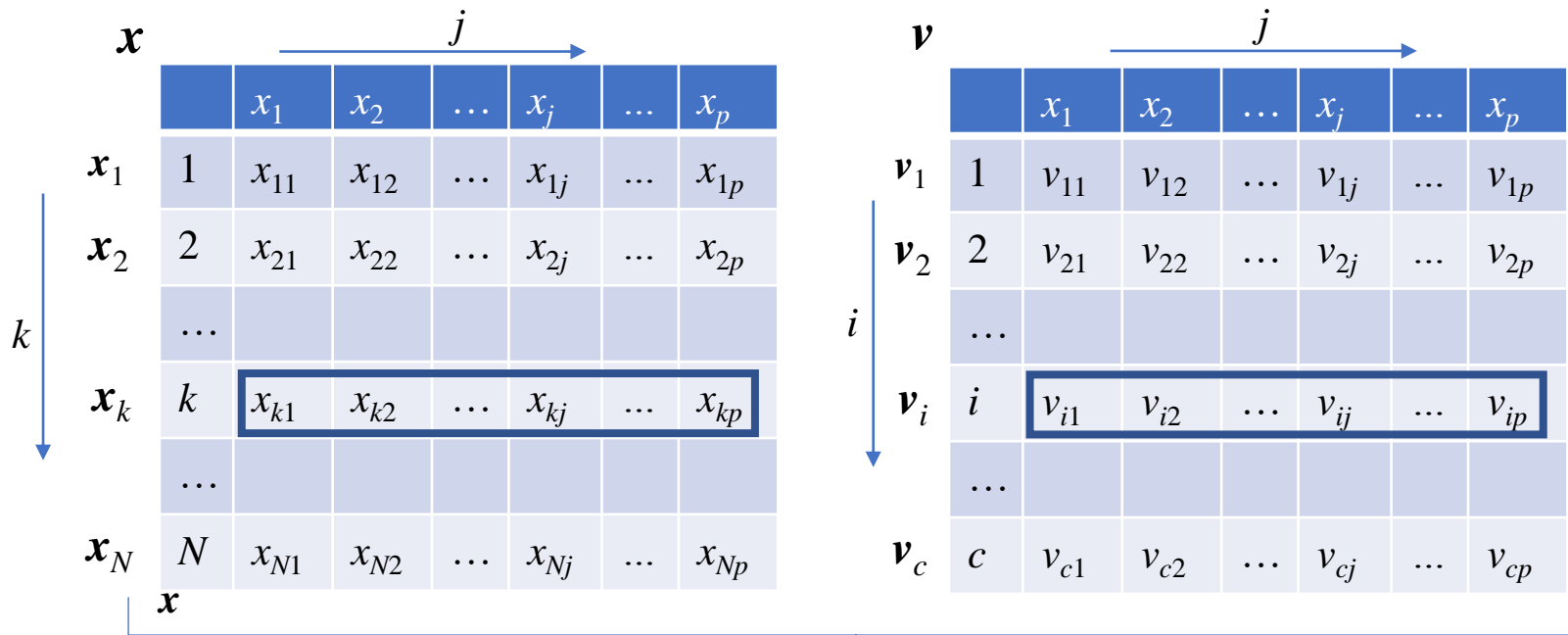


D	1	2	\dots	i	\dots	c
1	D_{11}	D_{12}	\dots	D_{1i}	\dots	D_{1c}
2	D_{21}	D_{22}	\dots	D_{2i}	\dots	D_{2c}
\dots	\dots	\dots	\dots	\dots	\dots	\dots
k	D_{k1}	D_{k2}	\dots	D_{ki}	\dots	D_{kc}
\dots	\dots	\dots	\dots	\dots	\dots	\dots
N	D_{N1}	D_{N2}	\dots	D_{Ni}	\dots	D_{Nc}

Το διάνυσμα \mathbf{x}_k ανήκει στην συστάδα το κέντρο της οποίας απέχει λιγότερο από το \mathbf{x}_k . Άρα, για να δούμε που ανήκει το \mathbf{x}_k υπολογίζουμε την θέση της ελάχιστης απόστασης της γραμμής k του πίνακα D .

Ο Αλγόριθμος c-Means

B. Η χρήση των Κέντρων των Συστάδων στον Προσδιορισμό των Συναρτήσεων Συμμετοχής



```
function [u]=Membership(N,c,D)
```

```
u=zeros(N,c);
```

```
for k=1:N,
```

```
z=D(k,:);
```

```
elaxisto=min(z);
```

```
for i=1:c
```

```
if elaxisto==D(k,i);
```

```
u(k,i)=1;
```

```
else
```

```
u(k,i)=0;
```

```
endif
```

```
endfor
```

```
endfor
```

	1	2	i	c		
1	D_{11}	D_{12}	\dots	D_{1i}	\dots	D_{1c}
2	D_{21}	D_{22}	\dots	D_{2i}	\dots	D_{2c}
\dots	\dots	\dots	\dots	\dots	\dots	\dots
k	D_{k1}	D_{k2}	\dots	D_{ki}	\dots	D_{kc}
\dots	\dots	\dots	\dots	\dots	\dots	\dots
N	D_{N1}	D_{N2}	\dots	D_{Ni}	\dots	D_{Nc}

Το διάνυσμα x_k ανήκει στην συστάδα το κέντρο της οποίας απέχει λιγότερο από το x_k . Άρα, για να δούμε που ανήκει το x_k υπολογίζουμε την θέση της ελάχιστης απόστασης της γραμμής k του πίνακα D .

Ο Αλγόριθμος c-Means

B. Η χρήση των Κέντρων των Συστάδων στον Προσδιορισμό των Συναρτήσεων Συμμετοχής

```
x=[10 9;4 3.5;9.5 11;2 3;12 9.5;1 5.5];
v=[10.5 9.83; 2.3 4];

dim_x=size(x);
dim_v=size(v);
N=dim_x(1);
p=dim_x(2);
c=dim_v(1);

D=Distances(N,c,p,x,v);
u=Membership(N,c,D);

save('Memberships.txt','u');
```

```
function D=Distance(N,c,p,x,v)
D=zeros(N,c);
for k=1:N
    for i=1:c
        s=0;
        for j=1:p
            s=s+(x(k,j)-v(i,j))^2;
        endfor
        D(k,i)=sqrt(s);
    endfor
endfor
```

```
function [u]=Membership(N,c,D)
u=zeros(N,c);
for k=1:N,
    z=D(k,:);
    elaxisto=min(z);
    for i=1:c
        if elaxisto==D(k,i);
            u(k,i)=1;
        else
            u(k,i)=0;
        endif
    endfor
endfor
```


ΚΑΛΟ ΑΠΟΓΕΥΜΑ