#### Esercizio 1

Si consideri la seguente matrice di distanza

u1 u2 u3 u4 u5 u6

u1 0 4 6 12 13 7

u2 4 0 2 8 9 5

u3 6 2 0 6 7 5

u4 12 8 6 0 1 9

u5 13 9 7 1 0 8

u6 7 5 5 9 8 0

1. si determini la sequenza delle partizioni identificata secondo il metodo del legame singolo;
2. si disegni il dendrogramma corrispondente identificando in base ad esso quando è opportuno arrestare la procedura di formazione dei gruppi motivando la risposta fornita
3. Si riporti la definizione di distanza di Mahalanobis tra due righe della matrice dei dati X e si dimostri che questa è invariante rispetto ad una traslazione dei due vettori rispetto allo stesso vettore c

*SVOLGIMENTO*

1. *Distanza legame singolo tra dove*

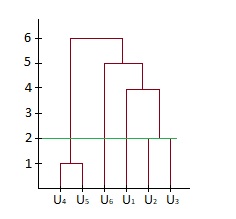
*A.*

*B.*

*C.*

*D.*

1. *Si ritiene opportuno tagliare il dendogramma a distanza 2, per evitare di comporre cluster troppo eterogenei al loro interno*



1. Distanza di Mahalanobis

Traslazione di un vettore C =>

=

#### Esercizio 2

Sia assegnata la matrice dei dati relativa alle unità statistiche u1,…,u5 per cui sono state rilevate le variabili x e y

id u1 u2 u3 u4 u5

x -2 -1 0 1 2

y -2 0 0 0 1

1. Convenendo di considerare come outlier un punto che dista dal centroide del campione più di 2 secondo la metrica di Mahalanobis, si dica se u1 è un outlier.
2. Si determini la sequenza delle partizioni ottenuta tramite il metodo del legame singolo ricorrendo alla metrica euclidea.

Si approssimino i calcoli a due decimali

*SVOLGIMENTO*

1. *Baricentro:*

*= – 0 = 2*

*= – = 0.96*

*= = 1.2*

*=>*

1. Distanza Euclidea:

2.24

2.83

3.61

5

1

2

3.16

1

2.24

1.41

**Esercizio 4**

Data la matrice di correlazione con |*r*| ≤ 1 e *r* ≠ 0 se ne calcolino gli autovalori e si mostri che l’autovettore normato di R associato al primo autovalore non dipende da *r.*

*SVOLGIMENTO*

*Polinomio caratteristico:*

*det(R-λI)==(1-λ)=(1-λ)=(1-λ)*

*da cui, senza perdita di generalità:*

*poiché per definizione l’autocoppia (a,*

*= da cui:*

*=> Inserendo la condizione di normalità per gli autovalori:*

#### *=>=>=>=>Scegliendo la soluzione positiva*

#### *a=(0.707 , 0 , 0.707)’ Non dipende da r*

#### Esercizio 3

Si consideri la matrice dei dati relativa a cinque unità statistiche

x y z

1 -2 1.0 -1

2 2 -0.5 1

3 0 -1.0 0

4 -2 -0.5 -1

5 2 1.0 1

1. si calcoli il baricentro del campione;
2. si calcoli la matrice di varianza e covarianza utilizzando la formula non corretta e la varianza totale;
3. si calcoli l’indice di variabilità relativo di Wilks e si interpreti il risultato ottenuto.

SVOLGIMENTO

1. (0,0,0)
2. S= V.T.=Tr(S)=4.7

=> Perfetta collinearità tra X, Y e Z cioè una di esse è combinazione lineare delle altre infatti