Aprendizaje no supervisado

VC04: Agrupamiento espectral – Conocimientos básicos

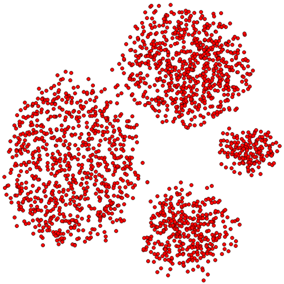
Félix José Fuentes Hurtado
felixjose.fuentes@campusviu.es

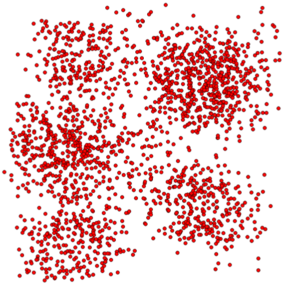
Universidad Internacional de Valencia

Agrupamiento

Tipos de algoritmos de agrupamiento

- Basados en particiones
- Jerárquicos
- Espectrales
- Basados en densidad
- Probabilísticos

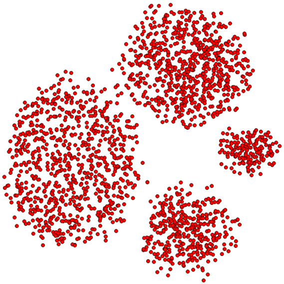


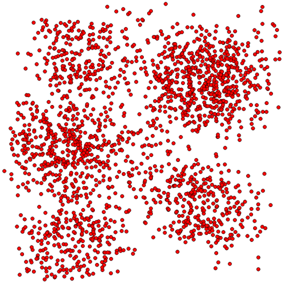


Agrupamiento

Tipos de algoritmos de agrupamiento

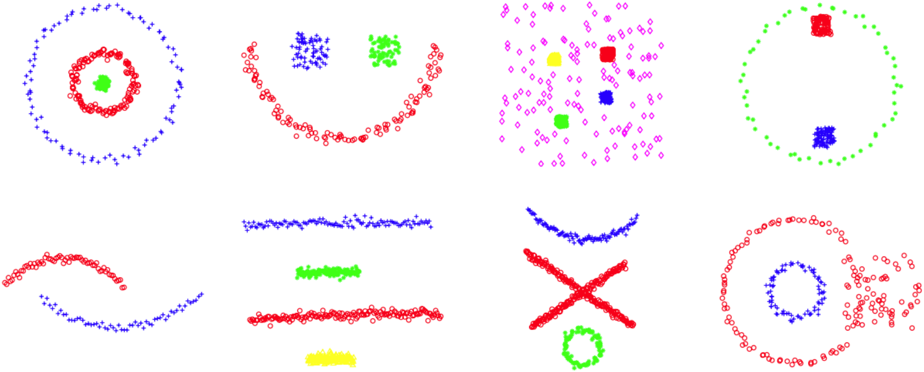
- Basados en particiones
- Jerárquicos
- **Espectrales**
- Basados en densidad
- Probabilísticos





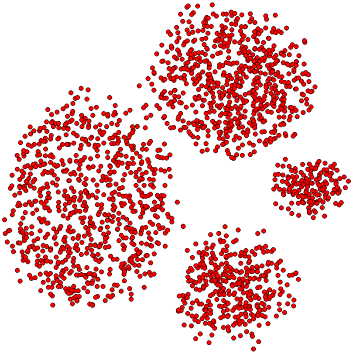
Agrupamiento

Clústeres de formas diversas

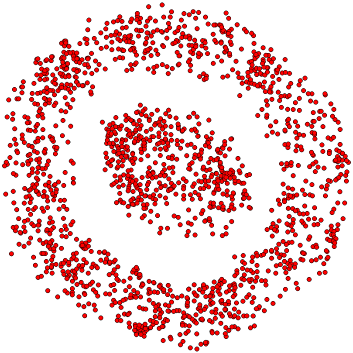


Agrupamiento

Tipos de clústeres



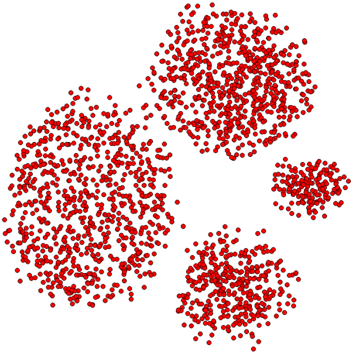
Grupos compactos



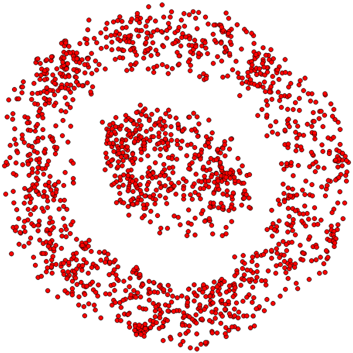
Grupos conexos

Agrupamiento

Tipos de clústeres



Grupos compactos
K-**means**



Grupos conexos
Espectral

Definición

Dado un conjunto de datos, el agrupamiento trata de identificar **subgrupos homogéneos** de ejemplos que manifiestan **diferencias relevantes con los otros subgrupos** que se formen.

Definición

Dado un conjunto de datos, el agrupamiento trata de identificar **subgrupos homogéneos** de ejemplos que manifiestan **diferencias relevantes con los otros subgrupos** que se formen.

Buscar el agrupamiento que maximiza la **dispersión interclúster**:

y minimiza la **dispersión intraclúster**:

Definición

Dado un conjunto de datos, el agrupamiento trata de identificar **subgrupos homogéneos** de ejemplos que manifiestan **diferencias relevantes con los otros subgrupos** que se formen.

Buscar el agrupamiento que maximiza la **dispersión interclúster**:

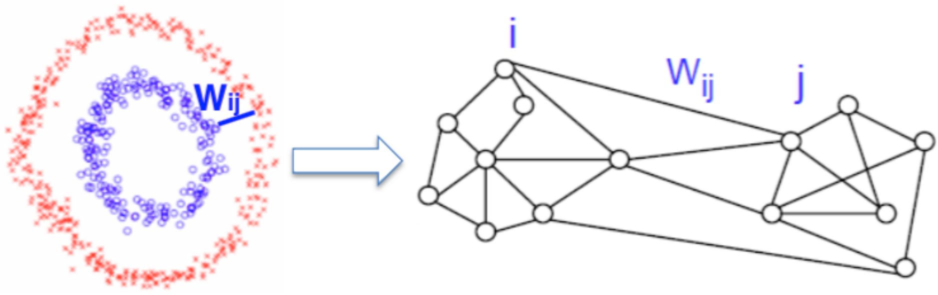
$$O(C) = \frac{1}{2} \sum_{k=1}^{K} \sum_{i:C(x_i)=k} \sum_{i':C(x_{i'})\neq k} d(x_i, x_{i'})$$

y minimiza la **dispersión intraclúster**:

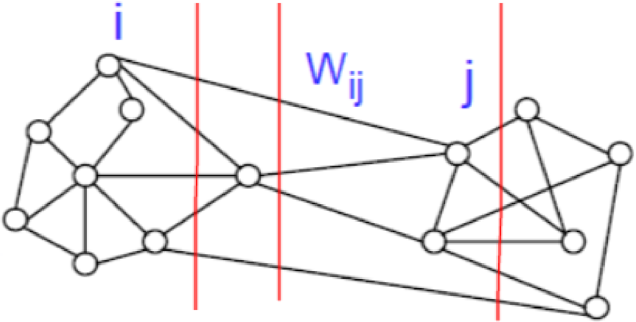
$$I(C) = \frac{1}{2} \sum_{k=1}^{K} \sum_{i:C(x_i)=k} \sum_{i':C(x_{i'})=k} d(x_i, x_{i'})$$

Agrupamiento espectral

Dataset a grafo



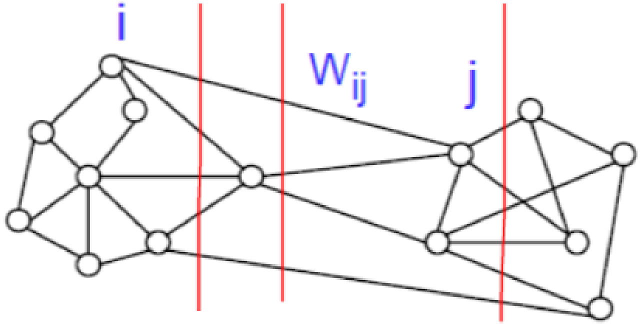
El peso W_{ij} será mayor cuanto más parecidos sean dos elementos



Separar en dos el grafo de tal manera que se eliminen el mínimo número de aristas

Agrupamiento espectral

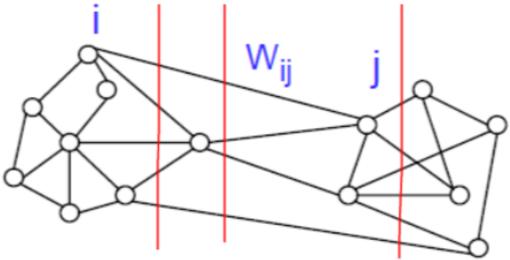
Corte mínimo de un grafo



Separar en dos el grafo de tal manera que la suma de los pesos de las aristas eliminadas sea mínima

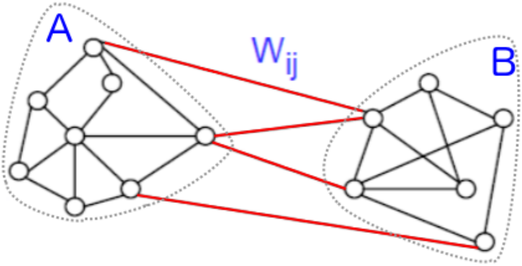
Separar en dos el grafo de tal manera que la suma de los pesos de las aristas eliminadas sea mínima

$$\arg \min_{\{A,B\}} corte(A,B) = \arg \min_{\{A,B\}} \sum_{i \in A} \sum_{i \in B} W_{ij}$$



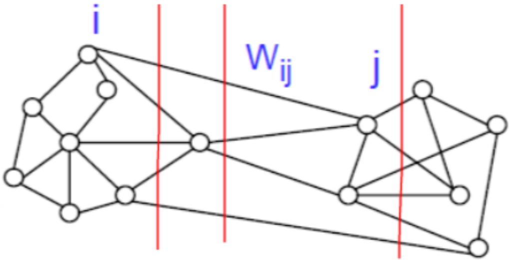
Separar en dos el grafo de tal manera que la suma de los pesos de las aristas eliminadas sea mínima

$$\arg \min_{\{A,B\}} corte(A, B) = \arg \min_{\{A,B\}} \sum_{i \in A} \sum_{i \in B} W_{ij}$$



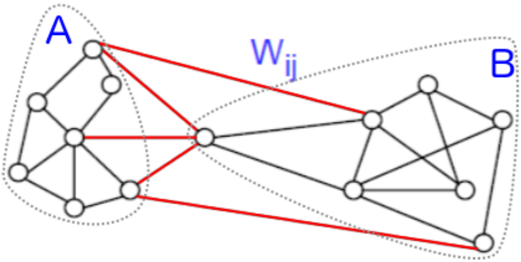
Separar en dos el grafo de tal manera que la suma de los pesos de las aristas eliminadas sea mínima

$$\arg \min_{\{A,B\}} corte(A,B) = \arg \min_{\{A,B\}} \sum_{i \in A} \sum_{i \in B} W_{ij}$$



Separar en dos el grafo de tal manera que la suma de los pesos de las aristas eliminadas sea mínima

$$\arg \min_{\{A,B\}} corte(A, B) = \arg \min_{\{A,B\}} \sum_{i \in A} \sum_{i \in B} W_{ij}$$

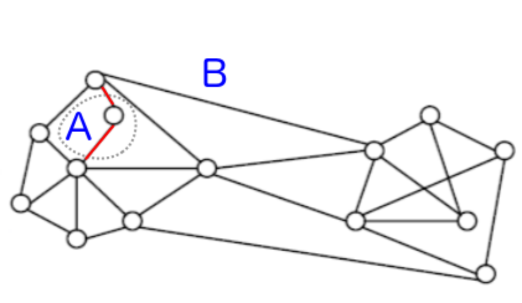


Características

- Fácil de obtener en tiempo razonable
- Cortes mínimos, no esperados

Características

- Fácil de obtener en tiempo razonable
- Cortes mínimos, no esperados



Separar en dos el grafo de tal manera que la suma de los pesos de las aristas eliminadas sea mínima

y los grupos resultantes sean de tamaño similar

$$\arg\min_{\{A,B\}} \left(\frac{1}{vol(A)} + \frac{1}{vol(B)} \right) corte(A, B)$$

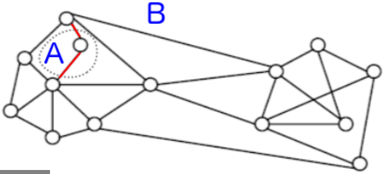
$$vol(A) = \sum_{i \in A} grado_i = \sum_{i \in A} \sum_{j \neq i} W_{ij}$$

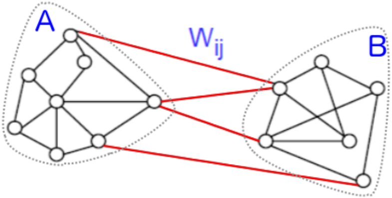
Separar en dos el grafo de tal manera que la suma de los pesos de las aristas eliminadas sea mínima

y los grupos resultantes sean de tamaño similar

$$\arg\min_{\{A,B\}} \left(\frac{1}{vol(A)} + \frac{1}{vol(B)} \right) corte(A, B)$$

$$vol(A) = \sum_{i \in A} grado_i = \sum_{i \in A} \sum_{j \neq i} W_{ij}$$





Problema:

$$\arg\min_{\{A,B\}} \left(\frac{1}{\sum_{i\in A}\sum_{j\neq i} W_{ij}} + \frac{1}{\sum_{i\in B}\sum_{j\neq i} W_{ij}} \right) \sum_{i\in A}\sum_{i\in B} W_{ij}$$

Características

- Los casos aislados no serán detectados como cortes mínimos
- No se puede resolver en un tiempo razonable

Problema:

$$\arg \min_{\{A,B\}} \left(\frac{1}{\sum_{i \in A} \sum_{j \neq i} W_{ij}} + \frac{1}{\sum_{i \in B} \sum_{j \neq i} W_{ij}} \right) \sum_{i \in A} \sum_{i \in B} W_{ij}$$

Características

- Los casos aislados no serán detectados como cortes mínimos
- No se puede resolver en un tiempo razonable

El agrupamiento espectral aproxima esta optimización mediante una transformación de los datos a partir de la matriz de adyacencias del grafo

Idea

Probabilidad de alcanzar un nodo del grafo transitando por él de manera aleatoria.
En cada momento (nodo i), se selecciona el siguiente nodo j de manera aleatoria según el peso de las aristas de i

$$P_{ij} = \begin{cases} \frac{W_{ij}}{\sum_{j' \neq i} W_{ij'}} & , \quad \text{si el nodo } j' \text{ está conectado con el } i \\ 0 & , \quad \text{si no} \end{cases}$$

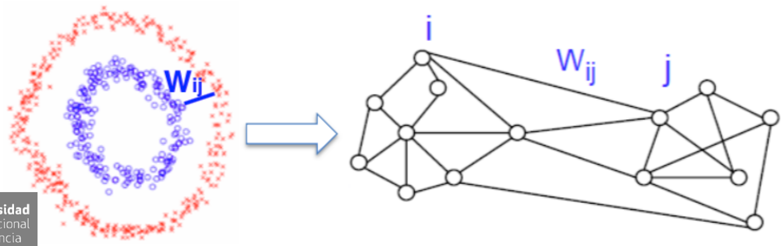
Agrupamiento espectral

Interpretación de camino aleatorio en un grafo

Idea

Probabilidad de alcanzar un nodo del grafo transitando por él de manera aleatoria.
En cada momento (nodo i), se selecciona el siguiente nodo j de manera aleatoria según el peso de las aristas de i

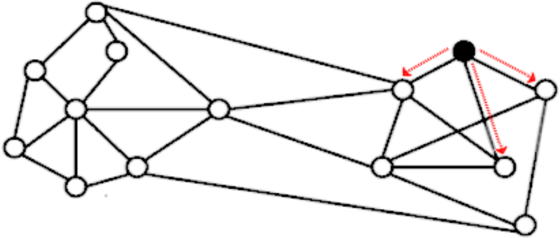
$$P_{ij} = \begin{cases} \frac{W_{ij}}{\sum_{j' \neq i} W_{ij'}} & , \quad \text{si el nodo } j' \text{ está conectado con el } i \\ 0 & , \quad \text{si no} \end{cases}$$



Idea

Probabilidad de alcanzar un nodo del grafo transitando por él de manera aleatoria.
En cada momento (nodo i), se selecciona el siguiente nodo j de manera aleatoria según el peso de las aristas de i

$$P_{ij} = \begin{cases} \frac{W_{ij}}{\sum_{j' \neq i} W_{ij'}} & , \quad \text{si el nodo } j' \text{ está conectado con el } i \\ 0 & , \quad \text{si no} \end{cases}$$



De manera natural, se construye una matriz de transiciones P_{ij}

La n-ésima potencia de una matriz de transiciones, P^n, recoge en cada celda P^n_{ij} la probabilidad de llegar del nodo i al nodo j en n pasos tomados aleatoriamente

Agrupamiento espectral

Interpretación de camino aleatorio en un grafo

De manera natural, se construye una matriz de transiciones P_{ij}

La n-ésima potencia de una matriz de transiciones, P^n, recoge en cada celda P^n_{ij} la probabilidad de llegar del nodo i al nodo j en n pasos tomados aleatoriamente

P	1	2	3	4	5	6
1	0.00	0.00	0.37	0.00	0.31	0.32
2	0.35	0.36	0.00	0.00	0.00	0.28
3	0.00	0.00	0.40	0.27	0.33	0.00
4	0.21	0.32	0.16	0.31	0.00	0.00
5	0.52	0.00	0.00	0.18	0.30	0.00
6	0.17	0.27	0.00	0.23	0.14	0.18

Agrupamiento espectral

Interpretación de camino aleatorio en un grafo

De manera natural, se construye una matriz de transiciones P_{ij}

La n-ésima potencia de una matriz de transiciones, P^n, recoge en cada celda P^n_{ij} la probabilidad de llegar del nodo i al nodo j en n pasos tomados aleatoriamente

P	1	2	3	4	5	6
1	0.00	0.00	0.37	0.00	0.31	0.32
2	0.35	0.36	0.00	0.00	0.00	0.28
3	0.00	0.00	0.40	0.27	0.33	0.00
4	0.21	0.32	0.16	0.31	0.00	0.00
5	0.52	0.00	0.00	0.18	0.30	0.00
6	0.17	0.27	0.00	0.23	0.14	0.18

P^2	1	2	3	4	5	6
1	0.21	0.08	0.15	0.23	0.26	0.06
2	0.18	0.21	0.13	0.07	0.15	0.27
3	0.23	0.09	0.20	0.25	0.23	0.00
4	0.18	0.22	0.19	0.14	0.12	0.16
5	0.19	0.06	0.22	0.11	0.25	0.16
6	0.25	0.22	0.10	0.14	0.12	0.16

Resultado

La probabilidad de transitar entre nodos de distintos clústeres, si están separados, será menor que la de transitar entre los nodos de un mismo clúster

Si los diferentes clústeres están conectados, la matriz es única

Agrupamiento espectral

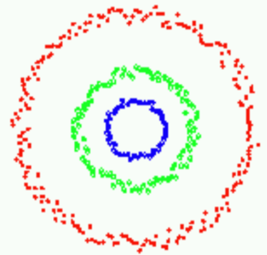
Interpretación de camino aleatorio en un grafo

Resultado

La probabilidad de transitar entre nodos de distintos clústeres, si están separados, será menor que la de transitar entre los nodos de un mismo clúster

Si los diferentes clústeres están conectados, la matriz es única

Si están desconectados, la matriz tiene una forma...



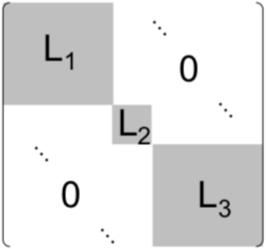
Agrupamiento espectral

Interpretación de camino aleatorio en un grafo

Resultado

La probabilidad de transitar entre nodos de distintos clústeres, si están separados, será menor que la de transitar entre los nodos de un mismo clúster

Si los diferentes clústeres están conectados, la matriz es única



Aprendizaje no supervisado

VC04: Agrupamiento espectral – Conocimientos básicos

Félix José Fuentes Hurtado
felixjose.fuentes@campusviu.es

Universidad Internacional de Valencia