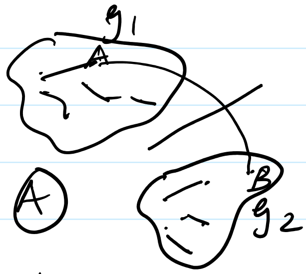
Modularity Maximiza^n.

29 May 2023 09:54

Modularity : A measure to check how well is the graph of nodes and its connections.

Each vertex/node i has a type c_i

There are n_c type of labels.



Actual no. of edges.

$$\sum_{\text{edges }(i,j)} \delta(c_i, c_j) = \frac{1}{2} \sum_{i,j} A_{ij}\, \delta(c_i, c_j) \longrightarrow (A)$$

Every pair of vertices (i, j) is counted twice (j, i) ($\because$ We assume undirected graph)

$\delta(c_i, c_j) = 1$ if $c_i = c_j$
$= 0$, otherwise

Expected no. of edges.

$$\frac{1}{2} \sum_{i,j} \frac{k_i \cdot k_j}{2m}\, \delta(c_i, c_j) \longrightarrow (B)$$

Modularity Actual − Expected.

$\underline{A - B}$

Expected no. of edges bet^n 2 nodes v_i & v_j

$$\frac{1}{2} \sum_{i,j} A_{ij}\, \delta(c_i, c_j) - \frac{1}{2} \sum_{i,j} \frac{k_i k_j}{2m}\, \delta(c_i, c_j)$$

$$= \frac{1}{2} \sum_{i,j} \left(A_{ij} - \frac{k_i k_j}{2m}\right) \delta(c_i, c_j)$$

Total edges m.

Modularity matrix

$$Q = \frac{1}{2m} \sum \boxed{\left(A_{ij} - \frac{k_i k_j}{2m}\right)}\, \delta(c_i, c_j)$$

Modularity

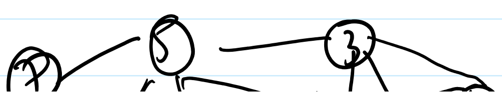
$Q > 0 \rightarrow$ Assortative Mixing

$Q < 0 \rightarrow$ Dissassortative "

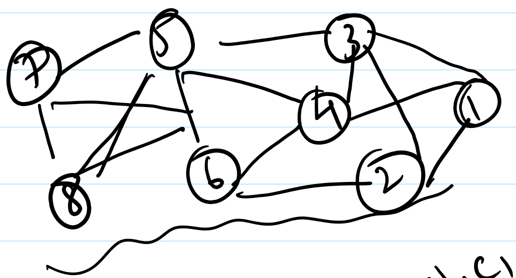
$Q = 0 \rightarrow$ Random

Always $Q \leq 1$

Expected no. of edges bet^n V_1 & V_4

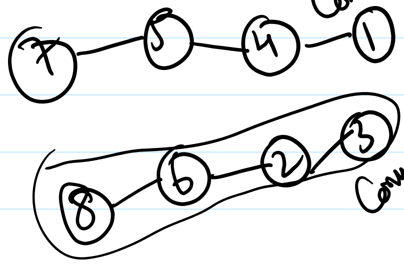


$\ldots = 0.4$



Expected no. of edges bet^n v_1 & v_4

$$= \frac{k_1 \cdot k_4}{2m} = \frac{3*4}{2*15} = 0.4$$



$$\sum_{i,j \in C_2} \left(A_{ij} - \frac{k_i \cdot k_j}{2m} \right)$$

$\Rightarrow$

$A_{ij}=1$	$A_{ij}=0$
$\to$ 2−3	3−6
$\to$ 2−6	3−8
$\to$ 6−8	2−8

C_1 [1, 4, 5, 7].

$A_{ij}=1$:
- 1−4 $\Rightarrow 1 - \frac{3*4}{30} = 0.6$
- 4−5 $\Rightarrow 1 - \frac{4*5}{30} = 0.33$
- 5−7 $\Rightarrow 1 - \frac{5*3}{30} = 0.5$

$A_{ij}=0$:
- 1−5 $\Rightarrow 0 - \frac{3*5}{30} = -0.50$
- 1−7 $\Rightarrow 0 - \frac{3*3}{30} = -0.30$
- 4−7 $\Rightarrow 0 - \frac{4*3}{30} = -0.40$

Total modularity score of Community C1 $= 0.60 + 0.33 + 0.5 - 0.50 - 0.3 - 0.4 \approx 0.23$

||ly for C2.

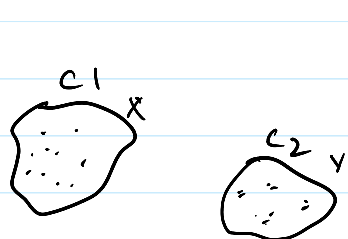
$$\sum_{l=1}^{s} \sum_{i,j \in C_l} \left(A_{ij} - \frac{k_i \cdot k_j}{2m} \right)$$

We require higher values of Q

Aim: To maximize Modularity

$$Q = \frac{1}{2m} \sum_{ij} \boxed{A_{ij} - \frac{k_i k_j}{2m}} \, \delta(c_i, c_j).$$

$\to$ Modularity matrix. B_{ij}



$$= \frac{1}{2m} \sum_{ij} B_{ij} \, \delta(c_i, c_j)$$

$c_i \leftarrow$ label for vertex i

$$Z = \begin{array}{c|ccc} & v_1 & v_2 & v_3 \\ \hline v_1 & & & \\ v_2 & 1 & 0 & 1 \\ v_3 & & & \end{array} =$$

$$\sum_j B_{ij} = \sum_j A_{ij} - \frac{k_i k_j}{2m} = \left(\sum_j A_{ij} \right) - \frac{k_i}{2m} \boxed{\sum_j k_j}$$

$$\sum_j B_{ij} = \sum_j A_{ij} - \frac{k_i \sum_j k_j}{2m} = \left(\sum_j A_{ij}\right) - \frac{k_i}{2m}\left[\sum_j k_j\right]$$

$$k_i - \frac{k_i}{\cancel{2m}} \cdot \cancel{2m} = \underline{0}$$

$$\sum_j B_{ij} = 0$$

Let $s_i = 1$ if vertex i belongs to group 1
$s_i = -1$ if vertex i " " 2

$$\delta(c_i, c_j) = \frac{1}{2}(s_i s_j + 1)$$

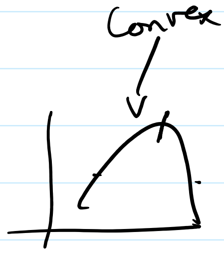
$\delta(c_i, c_j) = 1$ if $c_i = c_j$ $\rightarrow \frac{1}{2}(1 \cdot 1 + 1) = 1$ (group 1)

$\frac{1}{2}(-1 \cdot -1 + 1) = 1$ (group 2)

$$\underline{\underline{Q}} = \frac{1}{2m} \sum_{i,j} B_{ij}\, \delta(c_i, c_j)$$

$$= \frac{1}{2m} \sum_{i,j} B_{ij} \frac{1}{2} (s_i \cdot s_j + 1)$$

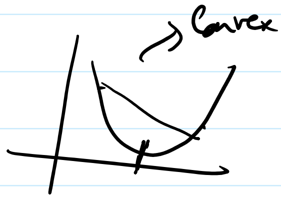
$$= \frac{1}{4m} \sum B_{ij}\, s_i \cdot s_j + \sum B_{ij} = 0.$$



$$= \frac{1}{4m} \sum B_{ij}\, s_i \cdot s_j$$

$$= \boxed{\frac{1}{4m} s^T B s} \longrightarrow \text{Quadratic form}$$

Is this convex



To maximize a f^n, we take derivative

$$\max_{s_i \in \mathbb{R}} \frac{1}{4m} s^T B s$$

$\frac{\partial}{\partial x} x^2 = 2x$

To make it tractable, $s^T s = n$

if there are 6 vertices in a graph
3 vertices w/ $s_i = 1$
3 vertices w/ $s_i = -1$ then

$$(-1^2 + (-1)^2 + (-1)^2 + (1)^2 + (1)^2 + 1^2) = 6$$

$\frac{\partial}{\partial s} s^2 B = 2sB$

$\therefore$ $n - \underline{s^T s} = 0$ ← Constraint

$$\partial(s^T B s) = 0 \Rightarrow \partial\left[s^T B s + \beta(n - s^T s)\right]$$

$\frac{\partial}{\partial S}$

$$\frac{\partial}{\partial s}(s^T B s) = 0 \Rightarrow \frac{\partial}{\partial s}\left[s^T B s + \beta \underbrace{(n - s^T s)}_{} \right]$$

Lagrangian Multiplier method

$$2SB + \frac{\partial}{\partial s}\beta n - \frac{\partial}{\partial s}\beta s^T s$$

$$2SB + 0 - 2\beta S = 0$$

$$2BS - 2\beta S = 0$$

$$\cancel{2}BS = \cancel{2}\beta S$$

$Ax = \lambda x$

Eigen vector

Eigen value

eq^n

$$\underline{\underline{B}} S = \beta S$$

Maximize $\frac{1}{4m} s^T B s = \frac{1}{4m} s^T \beta s$ → eigen vector

→ eigen ~~val~~ value

Modularity Matrix

$S; \cancel{x}\ U$ → vertices → edges

$\underline{\underline{S}}$ will be leading eigen vector $S \equiv [\,\_\ \_\ \_\ \_\ \_\,]$ 5 entries

This vector will have either +ve entries or -ve entries

Make +ve entries as +1

-ve entries as -1

S will have n entries (components) because there are n vertices in the graph

Each component will have +1 or -1

Spectral partitioning

Modularity Maximiz^n algo

1. Calculate leading eigen vector of Modularity Matrix
2. Divide vertices according to the Signs of the components in the leading eigen vector.