UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

Temas selectos de termodinámica y física estadística 1

Ciencia de redes

Grupo 8385

Tarea 1

Daniel Rojo Mata

DanielRM@ciencias.unam.mx

I. Solución Problema 1.

Las *co-citas* de dos nodos i y j, en una red dirigida, se definen como el número de vértices que 'entran' tanto al nodo i como al nodo j. Es decir, si i y j son artículos, entonces son citados por algún otro artículo, llámese k.

Con base a esto, se define $A_{ik}A_{jk} = 1$ si es que i y j *son citados* por el nodo k, en otro caso $A_{ik}A_{jk} = 0$. Entonces, la matriz de *co-citas* se define como:

$$C_{ij} = \sum_{k=1}^{n} A_{ik}A_{kj}^{T}$$

$$\mathbb{C} = \mathbb{A}\mathbb{A}^{\mathbb{T}}$$

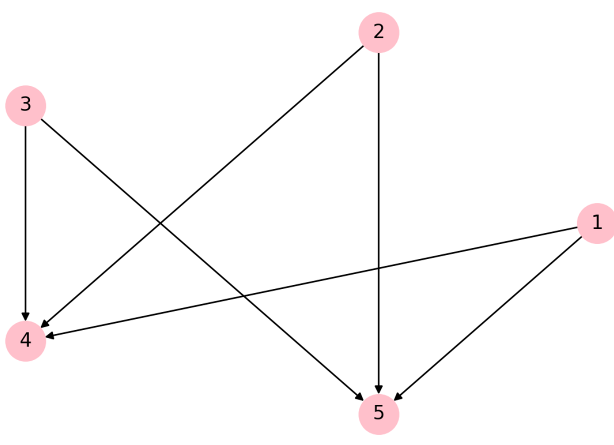


Figura 1: Representación de una red de *co-citas*. Los nodos 4 y 5 son citados por los nodos 1,2 y 3.

Por otro lado, dos nodos i y j, en una red dirigida, están *acoplados bibliográficamente* si los nodos citan el mismo artículo.

Ahora, se define $A_{ki}A_{kj} = 1$ si es que i y j *citan* al nodo k, en otro caso $A_{ki}A_{kj} = 0$.

Entonces, la matriz de *acoplamiento bibliográfico* se define como:

$$B_{ij} = \sum_{k=1}^{n} A_{ik}^{T}A_{jk}$$

$$\mathbb{B} = \mathbb{A}^{\mathbb{T}}\mathbb{A}$$

A diferencia de la red de *co-citas*, los enlaces de los nodos i y j 'salen' y se enlazan con otros nodos.

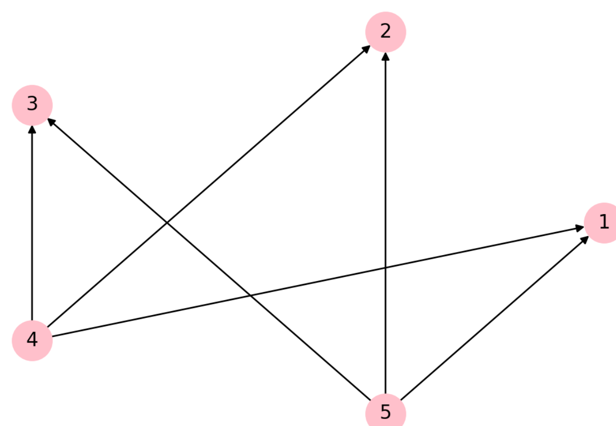


Figura 2: Representación de una red de *acoplamiento bibliográfico*. Los nodos 4 y 5 citan a los nodos 1,2 y 3.

Entonces, resumiendo:

- *Co-citas*; enlaces que *entran* a los nodos i y j. Véase la figura 1.
- *Acoplados bibliográficamente*; enlaces que *salen* de los nodos i y j. Véase la figura 2

$\square$

II. Solución problema 2.

La matriz de adyacencias de la red mostrada en (a) es la siguiente:

$$A = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \end{pmatrix} \quad (1)$$

La matriz de *co-citas* se define como: $C = AA^T$. La matriz transpuesta de 1 es:

$$A^T = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Entonces, la matriz de *co-citas* es de la forma:

$$C = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 2 & 1 & 0 & 1 \\ 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 2 \end{pmatrix}$$

Ahora, considérese el conjunto $a = \{1, 2, 3, 4\}$ y $b = \{1, 2, 3, 4, 5\}$, con esta elección, la matriz de incidencias de la red en (b) es de la siguiente forma:

$$B = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Nótese que justo es una matriz de 4×5, tal como se tiene en la definición.
Por otra parte, la matriz transpuesta es de la siguiente forma:

$$B^T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Ahora, la matriz de proyección sobre el conjunto independiente a (conjunto blanco) se define como: $P_a = BB^T$.
Por tanto, dicha matriz es la siguiente:

$$P_a = \begin{pmatrix} 2 & 1 & 0 & 1 \\ 1 & 2 & 0 & 2 \\ 0 & 0 & 1 & 1 \\ 1 & 2 & 1 & 4 \end{pmatrix} \quad (2)$$

La red de la matriz 2 es la siguiente.

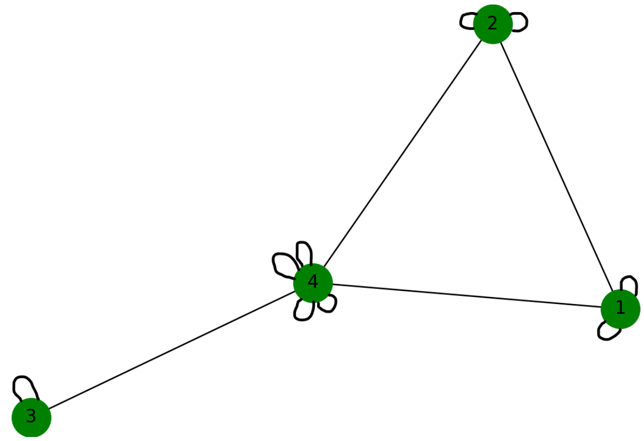


Figura 3: Red de multienlaces de la matriz P_a

Lo próximo es la matriz de proyección sobre el conjunto independiente b (conjunto negro), dicha matriz se define como: $P_b = B^T B$.
Por tanto, la matriz es la siguiente:

$$P_b = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 2 & 2 & 1 & 1 \\ 1 & 2 & 3 & 1 & 1 \\ 0 & 1 & 1 & 2 & 1 \\ 0 & 1 & 1 & 1 & 1 \end{pmatrix} \quad (3)$$

Así, la red asociada a la matriz 3 es la siguiente.

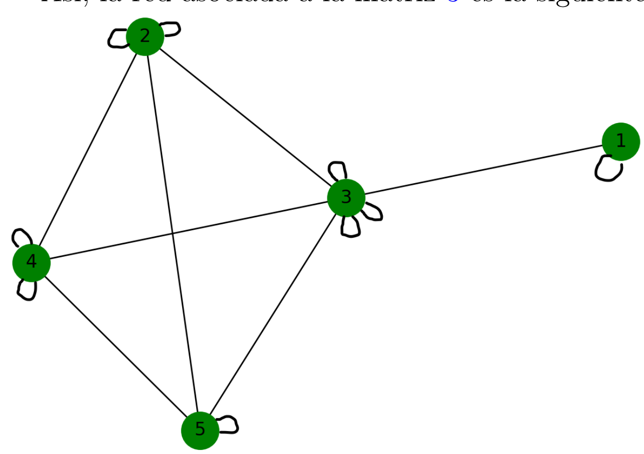


Figura 4: Red de multienlaces de la matriz P_b

[1]Los valores presentes en la tabla fueron obtenidos con ayuda de Python, véase los anexos, figura 12.

Por último se muestra una tabla[1] con algunas medidas de centralidad de la red mostrada en la figura 4.

- D_C: Centralidad de grado (Degree Centrality).
- C_C: Centralidad de cercanía (Closeness Centrality).
- B_C: Centralidad de intermediación (Betweeness Centrality).

Medidas de centralidad

Nodo	D_C	C_C	B_C
1	0.75	0.571	0
2	1.125	0.8	0
3	1.5	1.0	0.5
4	1.25	0.8	0.0
5	1.25	0.8	0.0

III. Solución problema 3.

Sea A la matriz de adyacencia de una red no dirigida de N nodos.

Se define lo siguiente:

$$\begin{cases} A_{ik}A_{jk} = 1, & \text{si k es vecino de i y j} \\ A_{ik}A_{jk} = 0, & \text{si k no es vecino de i y de j.} \end{cases}$$

Sumamos sobre todos los posibles valores, desde $k = 1$ hasta N.

$$\sum_{k=1}^{N} A_{ik}A_{jk} = \sum_{k=1}^{N} A_{ik}A_{kj}^{T} = \mathbb{A}^2$$

Con base a esto, se define:

$$n = A^2$$

En donde $n_{ij} = (A^2)_{ij}$.

Es decir, la entrada n_{ij} representa el número de caminos de longitud 2 entre el nodo i y el nodo j, lo cual, a su vez, representa los nodos en común entre dichos vértices, pues si $A_{ik}A_{jk} = 1$ es porque $\exists$ k tal que conecta a éstos con otro nodo, lo que implica que los nodos son vecinos en común.

$\square$

IV. Solución problema 4.

Sea A matriz de adyacencia de una red *k-regular*. Con los resultados de la sección *anexos*, la matriz de adyacencias de dicha red es:

$$A = \begin{pmatrix} 0 & 1 & 1 & \cdots & 1 \\ 1 & 0 & 1 & \cdots & 1 \\ 1 & 1 & 0 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 0 \end{pmatrix} \quad (4)$$

En este caso:

- $A_{ii} = 0$
- $\forall i, j \in \{1, ..., k\}$; $A_{ij} = A_{ji} = 1$

Por otro lado, se sabe que el grado del nodo i, para la matriz de dimensión k es de la forma:

$$K_i = \sum_{m=1}^{k} A_{im}$$

Pero, al ser una red *k-regular* el grado de cualquier nodo es k, por ello:

$$K_i = k, \ \forall i \in \{1, ..., k\}$$

Entonces, el *vector de grado*, para este caso es de la forma:

$$\vec{K} = \begin{pmatrix} k \\ k \\ k \\ \vdots \\ k \end{pmatrix} = k \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = k\vec{1} \quad (5)$$

Como se menciona en los anexos: El grado de un nodo i es el número de vecinos que este tiene; se 'proyecta' la matriz A sobre el vector $\vec{1}$, es decir se hace lo siguiente:

$$A\vec{1} = \vec{K}$$

Pero por 5 se tiene:

$$A\vec{1} = \vec{K} = k\vec{1}$$

$$\therefore \ A\vec{1} = k\vec{1}$$

Que es justo la condición de que k sea *valor propio* y que $\vec{1}$ sea *vector* propio de la matriz A.

$\square$

Para probar el siguiente inciso se hace uso del siguiente teorema.

Teorema (Frobenius, 1908-1912) [1]: *Sea* $\mathbf{A}$ *una matriz cuadrada con entradas no negativas. Si* $\mathbf{A}$ *es irreducible, entonces:*

1. *Existe un valor propio (simple)* $\lambda > 0$ *tal que* $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ *donde el vector propio* $\mathbf{v} > 0$. *Además* $\lambda \geq |\mu|$ *para cualquier otro valor propio* μ *de* $\mathbf{A}$.

2. *Cualquier vector propio* ≥ 0 *es un múltiplo de* $\mathbf{v}$.

La condición de que la matriz $\mathbf{A}$ sea *irreducible*, en el caso de *teoría de grafos*, se resume en que la gráfica (red) sea conexa, que por definición; $\forall i, j$ nodos, existe un $ij-camino$.

Y como la red, en nuestro problema, es *k-regular* se sigue que es conexa, además se aprecia en 4 que A es no negativa (sus entradas son positivas), de esta manera, por el Teorema de Frobenius, existe un único vector propio con entradas no negativas, que además es el asociado con el valor propio k.

En efecto:

Sea $\vec{u}$ vector propio de A con valor propio λ tal que $\vec{u} \neq \vec{1}$ y $\lambda \neq k$ y las entradas de $\vec{u}$ son todas positivas.

Como los vectores $\vec{1}$ y $\vec{u}$ son ortogonales se satisface:

$$\vec{1}^T \cdot \vec{u}^T = 0$$

$$\vec{1}^T \cdot \vec{u}^T = \sum_{j=1}^{k} u_j = 0$$

De la expresión anterior:

$$u_1 + \sum_{j=2}^{k} u_j = 0 \;\Rightarrow\; u_1 = -\sum_{j=2}^{k} u_j$$

Pero como todas las entradas de $\vec{u}$ son positivas, entonces, por la igualdad anterior:

$$u_1 < 0$$

Lo cual contradice la suposición de que las entradas de $\vec{u}$ son todas positivas, así, se concluye que el vector $\vec{1}$ es el único vector positivo.

Tal y como el Teorema asegura.

$\square$

V. Solución Problema 5.

Se sabe que la distancia entre dos nodos es la longitud mínima entre estos mientras que el diámetro es la mayor de las distancias para cualquier par de nodos.

Se busca una red en la cual el diámetro sea c veces más grande que la distancia promedio.

La idea del problema es hacer una red con una cantidad considerablemente grande de nodos para que, en promedio, estos tengan una distancia igual a 1.

Considérese el siguiente ejemplo visual.

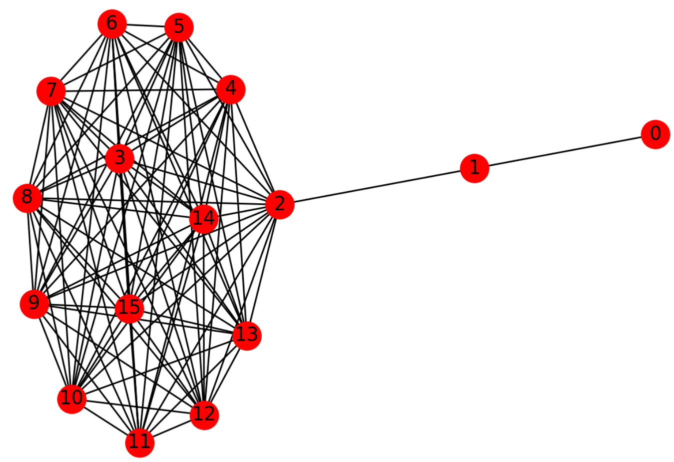


Figura 5: Red con 15 nodos, 2 de estos se encuentran alejados del resto.

En la figura 5 se muestra una red con 15 nodos [2], en el 'cúmulo' de nodos que se aprecia a la izquierda la distancia promedio de los nodos es aproximadamente 1, esto porque todos los nodos están conectados, además de que son bastantes. Ahora, entre los nodos 0,1,2 existe una distancia de 1.

Nótese que para todos los nodos en el cúmulo, la distancia promedio, al nodo 2 es igual 1, esto por estar conectados directamente con éste. Así:

$$\langle d_{i2} \rangle \approx 1, \;\forall i \in \{3, ..., 15\}$$

Luego, el diámetro, en este ejemplo, cumple con tener el valor $diam = 3$, debido a que la distancia máxima es hasta el nodo 0. es decir:

$$diam = max\{d_{ij}\} = 3, \;\forall i \in \{1, ..., 15\}$$

Por tanto se concluye que:

[2] Se realizó una red con 15 nodos para poder ejemplificar la idea principal del problema.

$$\langle d_{ij} \rangle < 3$$

Para el caso general se tiene que extender la cadena de nodos mostrados en la parte derecha, es decir, si queremos obtener $\langle d_{ij} \rangle < c$, se tendrían que posicionar $c-1$ nodos alejados del 'cúmulo', así como advierte la siguiente figura.

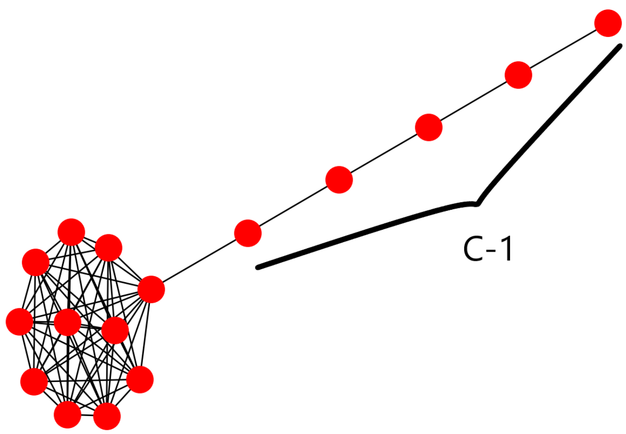


Figura 6: Red con N nodos, de los cuales, una cantidad $c-1$ está alejada del 'cúmulo'.

Es importante notar que $\langle d_{ij} \rangle < cdiam$ dependerá del número de nodos que se tenga en la red y el diámetro requerido.
Entre mayor sea el valor requerido para el diámetro mayor será el número de nodos de la red, en particular en el 'cúmulo'.

$\square$

VI. Solución Problema 6.

Considérese una red no dirigida de tamaño N donde se sabe que todos los nodos tienen grado 1, por comodidad, llámese a esta G.
Para que esto ocurra, debe haber una cantidad *par* de nodos para así poder formar parejas entre estos, de otra forma, los nodos no tendrían grado 1.
En todo lo próximo $x_i x_{i+1}$ es el enlace entre el nodo x_i y x_{i+1}. Si se tienen los nodos:

$$V_1(G) = \{x_1, x_2, x_3\}$$

Entonces un posible conjunto de enlaces es el siguiente:

$$E_1(G) = \{x_1x_2, x_1x_3, x_2x_3\}$$

La red es la que se muestra en la figura 7

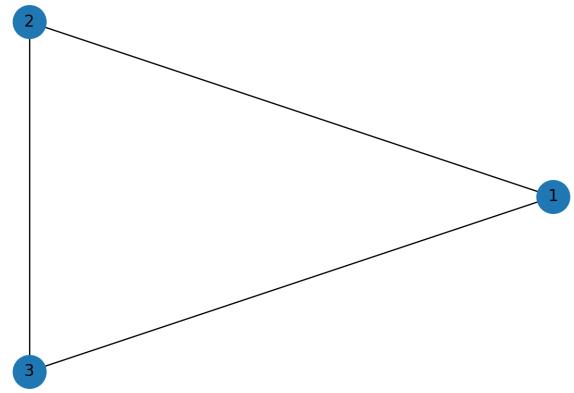


Figura 7: Representación visual de la red (V_1, E_1). Todos los nodos tienen grado 2.

Nótese que, en este ejemplo, todos los nodos tienen grado 2, lo cual no es lo requerido.
Sin embargo, con estos vértices se puede definir otro tipo de red.

$$V_2(G) = \{x_1, x_2, x_3\}$$

Ahora, considere los enlaces:

$$E_2(G) = \{x_1x_2\}$$

En este contexto, solo existe un enlace entre el nodo x_1 y x_2 teniendo ambos nodos grado 1, sin embargo, x_3 es un nodo aislado, por lo que no se cumple la condición de que todos los nodos tengan grado 1.
Tal y como advierte la siguiente figura.

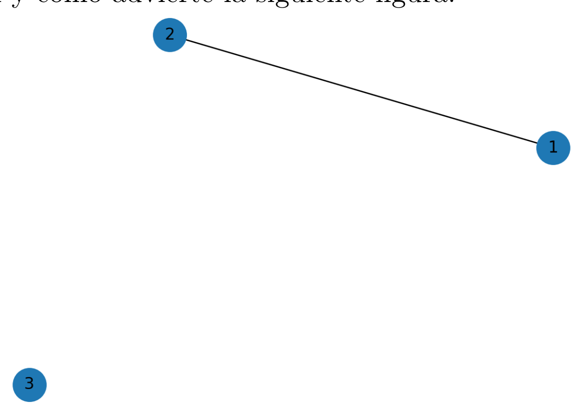


Figura 8: Representación visual de la red (V_2, E_2). Puesto que hay un vértice aislado, no todos los nodos tienen grado 1.

Esta es la motivación de que el número de nodos debe ser par.
Considérese:

$$V_3(G) = \{x_1, x_2, x_3, x_4\}$$

Ahora, considere los enlaces:

$$E_3(G) = \{x_1x_2, x_3x_4\}$$

De esta manera, no se tienen vértices aislados e inclusive todos los nodos tienen grado 1.

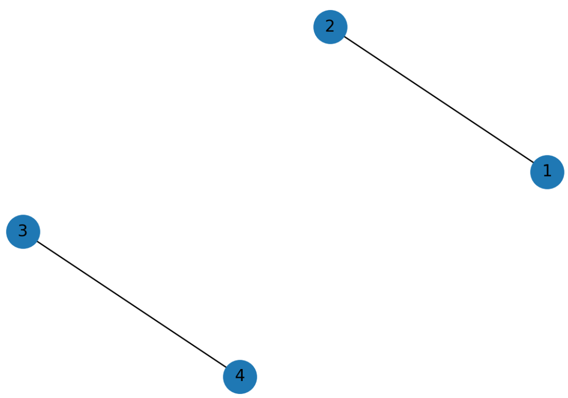


Figura 9: Representación visual de la red (V_3, E_3). Todos los nodos tienen grado 1.

Para generalizar lo anterior se definen los nodos y los enlaces de la siguiente forma:

Para los nodos:

$$V(G) = \{x_1, x_2, ..., x_N\}$$

En donde se entiende que hay una cantidad par de nodos.

Ahora, para los enlaces:

$$E_p(G) = \{x_{2p-1}x_{2p} \mid x_{2p-1}, x_p \in V(G) \wedge p \in \mathbf{N} - \{0\}\ \}$$

En donde $x_{2p-1}x_{2p}$ es el enlace entre el nodo x_{2p-1} y x_{2p}.

Así, por ejemplo, los conjuntos:

$$E_1(G) = \{x_1x_2 \mid x_1, x_2 \in V(G)\}$$

$$E_2(G) = \{x_3x_4 \mid x_3, x_4 \in V(G)\}$$

etc.

Entonces, todos los enlaces de la red son la unión de los siguientes conjuntos:

$$E(G) = E_1(G) \cup E_2(G) \cup ... \cup E_{N-1}(G)$$

$$E(G) = \bigcup_{i=1}^{N-1} E_i(G)$$

Entonces, por construcción, la red que se pide es la siguiente:

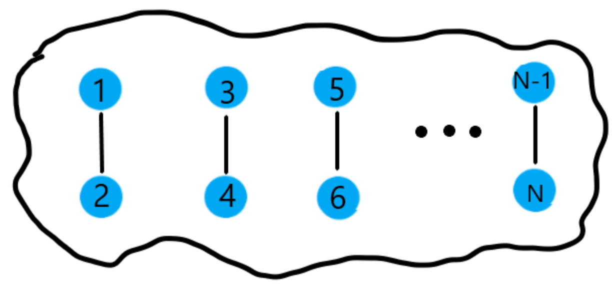


Figura 10: Representación visual de la red (V, E). La red consta de $\frac{N}{2}$ componentes, con N un número par.

Nótese que, por construcción, cada nodo tiene grado 1, entonces, la restricción que hay para N es que el número de nodos sea un número par. Luego, la distribución de grado se define como:

$$P_k = \frac{N_k}{N}$$

Con N_k número de nodos con grado k y N el número total de nodos.

Como todos los nodos de la red tienen grado 1, entonces $N_k = N$.

Por ende:

$$P_k = \frac{N}{N} = 1$$

Por último, la red tendría $\dfrac{N}{2}$ componentes.

$\square$

Se considera una red en donde cada nodo tiene grado $k = 2$ y el coeficiente de acumulación *clustering* es igual a 1, para todos los nodos.

Se sabe que el coeficiente de acumulación, C_i, es de la forma:

$$C_i = \frac{2L_i}{k_i(k_i - 1)} \tag{6}$$

En palabras, el coeficiente de clustering es el *número de enlaces entre los vecinos del nodo i* entre el *máximo número de enlaces posibles entre los vecinos de i.*

Como cada nodo tiene $C_i = 1$ y grado 2, en la ecuación 6 se sigue lo próximo:

$$1 = \frac{2L_i}{2(2-1)} \;\Rightarrow\; L_i = 1$$

Es decir, con base en las condiciones establecidas, se debe satisfacer que el número de enlaces entre los

vecinos del nodo i debe ser estrictamente 1.
Y como el grado de cada nodo es 2, solo existe la posibilidad de que la red conste de 3 nodos.
La imagen de la red es la siguiente.

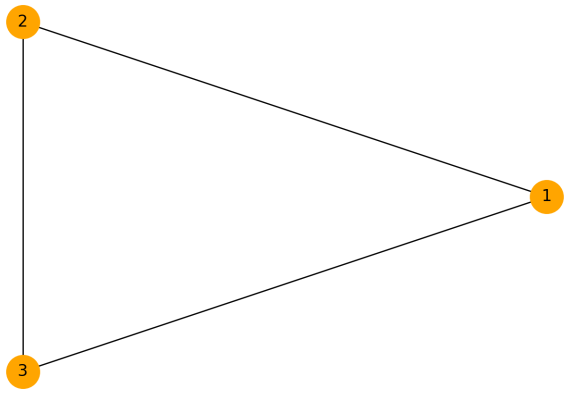


Figura 11: Red con coeficiente de acumulación igual a 1 donde todos los nodos tienen grado 2.

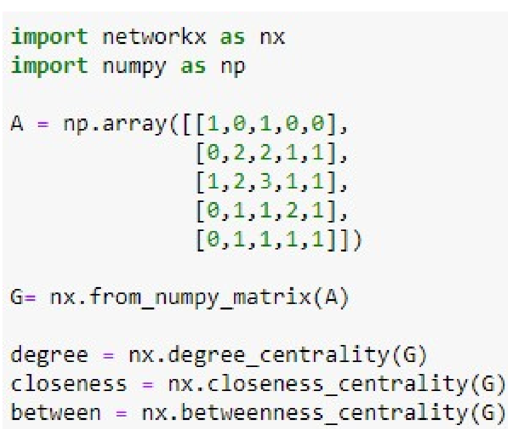
□

Referencias

[1] http://verso.mat.uam.es/~pablo.fernandez/index.html

VII. Anexos

Programa medidas de centralidad, problema 2.



```
import networkx as nx
import numpy as np

A = np.array([[1,0,1,0,0],
              [0,2,2,1,1],
              [1,2,3,1,1],
              [0,1,1,2,1],
              [0,1,1,1,1]])

G= nx.from_numpy_matrix(A)

degree = nx.degree_centrality(G)
closeness = nx.closeness_centrality(G)
between = nx.betweenness_centrality(G)
```

Figura 12: Programa realizado en Python.

Cálculos para el problema 4

Sea A la matriz de adyacencia de una red no dirigida de N nodos y $\vec{1}$ el vector columna de dimensión N, tal que $\forall i = 1, ..., N$; $\vec{1}_i = 1$.
La matriz A al ser de una red no dirigida satisface las siguientes propiedades:

- $A_{ii} = 0$
- $\forall i, j \in \{1, ..., N\}$; $A_{ij} = A_{ji}$

Con base a lo anterior, la matriz A puede ser expresada de la siguiente forma:

$$A = \begin{pmatrix} 0 & A_{12} & A_{13} & \cdots & A_{1N} \\ A_{21} & 0 & A_{23} & \cdots & A_{2N} \\ A_{31} & A_{32} & 0 & \cdots & A_{2N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{N1} & A_{N2} & A_{N3} & \cdots & 0 \end{pmatrix}$$

Mientras que el vector $\vec{1}$ es:

$$\vec{1} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

Como se sabe, el grado de un nodo i es el número de vecinos que este tiene. Para poder hallarlo se 'proyecta' la matriz A sobre el vector $\vec{1}$, es decir se hace lo siguiente:

$$A\vec{1} = \vec{K}$$

En la expresión anterior $\vec{K}$ es el *vector de grado*, en donde la componente K_i representa el grado del nodo i.
Nótese que la expresión anterior está bien definida, pues las dimensiones de A son $N \times N$ y la dimensión de $\vec{1}$ es $N \times 1$.
De esta modo, la dimensión de $\vec{K}$ es $N \times 1$. Por ende, desarrollando se tiene:

$$A\vec{1} = \begin{pmatrix} 0 & A_{12} & A_{13} & \cdots & A_{1N} \\ A_{21} & 0 & A_{23} & \cdots & A_{2N} \\ A_{31} & A_{32} & 0 & \cdots & A_{2N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{N1} & A_{N2} & A_{N3} & \cdots & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

$$A\vec{1} = \begin{pmatrix} 0 + A_{12} + A_{13} + ... + A_{1N} \\ A_{21} + 0 + A_{23} + ... + A_{2N} \\ A_{31} + A_{32} + 0 + ... + A_{3N} \\ \vdots \\ A_{N1} + A_{N2} + A_{N3} + ... + 0 \end{pmatrix}$$

Así, el vector $\vec{K}$ es:

$$\vec{K} = \begin{pmatrix} \sum_{m=1}^{N} A_{1m} \\ \sum_{m=1}^{N} A_{2m} \\ \sum_{m=1}^{N} A_{3m} \\ \vdots \\ \sum_{m=1}^{N} A_{Nm} \end{pmatrix}$$

En donde el elemento K_i es de la forma:

$$K_i = \sum_{m=1}^{N} A_{im}$$

Algoritmos para el problema 5

Todos los algoritmos fueron realizados y compartidos por Diego Ilhuitemok Barrios Martínez (estudiante del mismo grupo).

```python
import numpy as np
import networkx as nx

def A(n,c):
    n += 1
    A = np.zeros((n, n))

    for i in range(1,n):
      if i < c:
        A[i][i-1] = 1
      if i >= c-1:
        for j in range(c-1, i):
          A[j][i] = 1

    A = A + A.T

    return(A)
```

Figura 13: Con este algoritmo se generan las redes en 6 y 5

```python
#Buscando el camino mas corto
#(distancia) entre los nodos a y b
#spl = shortest path length
def spl(A, n, a, b):

    if a == b: return 0

    elif b < a:
        s = a; f = b
        b = s; a = f

    B = A

    t = 1
    while t < n:
        if B[b][a] != 0: return t
        else:
            t += 1
            B = np.dot(B, A)
```

Figura 14: Con este algoritmo se encuentra el camino más corto entre dos nodos.

```python
n = 300
c = 3

a = A(n, c)

d = 0
e = []

for i in range(n):
    for j in range(n):
        d += spl(a, n, i, j)
        e.append(spl(a, n, i, j))

D = max(e)

pd = d / (n * (n-1))

print("Diámetro", D)

print("Distancia promedio", pd)

print("D = c*pd")
print("c = ", D / pd)
```

Figura 15: Se encuentra el diámetro y se imprimen los resultados.