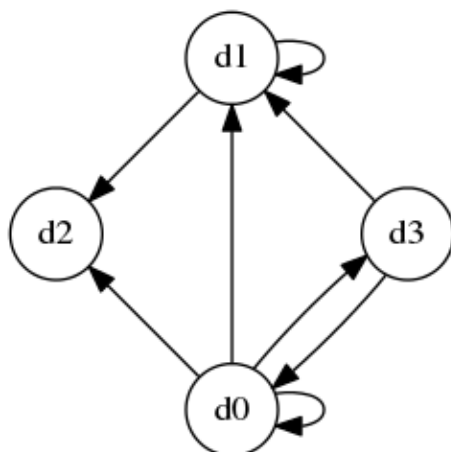


Dadas las siguientes páginas web y los enlaces entre ellas representadas como un grafo, se pide calcular el pagerank de cada página. Se debe calcular:

- i) la matriz de enlaces,
- ii) la matriz de probabilidades de transición,
- iii) la matriz de probabilidades de transición con teletransporte (utiliza un  $\alpha=0,15$  para el teletransporte),
- iv) todas las iteraciones para calcular el pagerank. Realiza como máximo cinco iteraciones.



- i) Matriz de enlaces:

```

d0 d1 d2 d3
d0 [1  1  1  1]
d1 [0  1  1  0]
d2 [0  0  0  0] <- callejón sin salida
d3 [1  1  0  0]
  
```

- ii) Matriz de probabilidades de transición inicial. Recordemos que se normaliza por filas:

```

      d0    d1    d2    d3
d0 [ 0.250 0.250 0.250 0.250 ]
d1 [ 0.000 0.500 0.500 0.000 ]
d2 [ 0.000 0.000 0.000 0.000 ]
d3 [ 0.500 0.500 0.000 0.000 ]
  
```

iii) Matriz de probabilidades de transición con teletransporte ( $\alpha=0.15$ ):

Recordatorio del teletransporte:

- Si estamos en un callejón sin salida (d2 en nuestro ejemplo):  
Saltar equiprobablemente a cualquier página web con probabilidad  $\frac{1}{N}$
- Si no estamos en un callejón sin salida se suman los siguientes valores:
  - Con probabilidad  $\alpha$ : Saltar a cualquier página web. Salto a cualquier página con probabilidad  $\frac{\alpha}{N}$  (en nuestro ejemplo  $0.15/4 = 0.0375$ , aprox. 0.038). Esta parte es igual para todas las páginas que nos son callejones sin salida
  - Con probabilidad  $1-\alpha$ : Saltar utilizando alguno de los enlaces de la página de forma equiprobable. Esta parte depende del número de enlaces de salida

		d0	d1	d2	d3
d0	[	0.250	0.250	0.250	0.250 ]
d1	[	0.038	0.463	0.463	0.038 ]
d2	[	0.250	0.250	0.250	0.250 ]
d3	[	0.463	0.463	0.038	0.038 ]

iv) Cálculo del Pagerank, recordemos que la probabilidad de estar en el estado  $j$  en el instante  $t+1$  será:  $\sum_i x_i P_{ij}$ , es decir, cada nuevo componente  $j$  del vector  $x$  en el tiempo  $t+1$  se calcula como el producto escalar del vector  $x$  del tiempo  $t$  por la columna  $j$  de la matriz de probabilidades.

		d0	d1	d2	d3
$x_{t=0} =$	[	1	0	0	0 ]
$x_{t=1} =$	[	0.250	0.250	0.250	0.250 ]
$x_{t=2} =$	[	0.250	0.356	0.250	0.144 ]
$x_{t=3} =$	[	0.205	0.356	0.295	0.144 ]
$x_{t=4} =$	[	0.205	0.356	0.295	0.144 ]

Como ya ha convergido el resultado, no hacemos más iteraciones, el enunciado indicaba que se podían realizar máximo 5 iteraciones.

Pagerank, es el valor del vector  $x$  cuando se ha alcanzado la convergencia:

$$\pi = [ \quad 0.205 \quad 0.356 \quad 0.295 \quad 0.144 \quad ]$$