

# Una aplicación de digrafos: GooglePage Rank

Idea: desplegar los resultados de búsquedas según su importancia.

Algoritmo de Google:

- ▶ Definir la importancia de las páginas web.
- ▶ Calcular la importancia de cada página.

Considerar el grafo de internet:

- ▶ Vértices: páginas web;
- ▶ Aristas:  $a \rightarrow b$  si hay un hyperlink de  $a$  apuntando hacia  $b$  y  $a \neq b$ .

## Definición

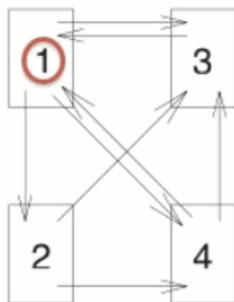
Sea  $x_k$  la importancia (no normalizada) del vértice (página)  $k$  y  $L_k$  el conjunto de vértices que inciden en  $k$ . Entonces

$$x_k = \sum_{j \in L_k} \frac{1}{\delta^-(j)} x_j$$

donde  $\delta^-(j)$  es el número de aristas que salen del vértice  $j$ .

## Ejemplo

Supongamos que el grafo de internet es:



entonces

$$x_1 = \frac{1}{1}x_3 + \frac{1}{2}x_4, \quad x_2 = \frac{1}{3}x_1$$

En total:

$$\begin{aligned} x_1 &= \frac{1}{1}x_3 + \frac{1}{2}x_4 \\ x_2 &= \frac{1}{3}x_1 \\ x_3 &= \frac{1}{3}x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_4 \\ x_4 &= \frac{1}{3}x_1 + \frac{1}{2}x_2 \end{aligned}$$

que podemos abreviar como

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 & 1/2 \\ 1/3 & 0 & 0 & 0 \\ 1/3 & 1/2 & 0 & 1/2 \\ 1/3 & 1/2 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

Se tiene un sistema de ecuaciones del tipo

$$\lambda X = AX$$

en tal caso, el vector  $X$  se llama *eigenvector* del *eigenvalor*  $\lambda$ .

Despejando

$$(A - \lambda I)X = 0$$

donde  $I$  es matriz identidad. En nuestro ejemplo:

$$\begin{pmatrix} -1 & 0 & 1 & 1/2 \\ 1/3 & -1 & 0 & 0 \\ 1/3 & 1/2 & -1 & 1/2 \\ 1/3 & 1/2 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

y por Gauss-Jordan:

$$x_1 = 2r, x_2 = \frac{2}{3}r, x_3 = \frac{3}{2}r, x_4 = r$$

con  $r \in \mathbb{R}$  variable libre.

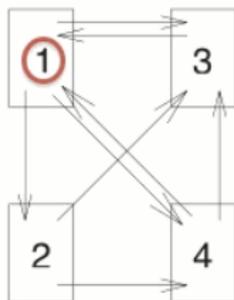
## Definición

La importancia normalizada  $x'_k$  del vertice  $k$  es

$$x'_k = \frac{x_k}{\sum_j x_j}$$

Por ejemplo, las importancias normalizadas, en nuestro ejemplo, son:

- ▶  $x'_1 = \frac{x_1}{x_1+x_2+x_3+x_4} = \frac{2r}{2r+(2/3)r+(3/2)r+r} = 12/31 \approx .3870967741935484$
- ▶  $x'_2 = \frac{x_2}{x_1+x_2+x_3+x_4} = 4/31 \approx .1290322580645161$
- ▶  $x'_3 = \frac{x_3}{x_1+x_2+x_3+x_4} = 9/31 \approx .2903225806451613$
- ▶  $x'_4 = \frac{x_4}{x_1+x_2+x_3+x_4} = 6/31 \approx .1935483870967742$



Entonces, los vértices ordenados por importancia según Google son:

- ▶ 1
- ▶ 3
- ▶ 4
- ▶ 2

En particular, el vértice 1 es el más importante.

*Aclaración:* Google no usa el método de Gauss-Jordan, sino el *método de la potencia* que se basa en el teorema de Perron-Frobenius.