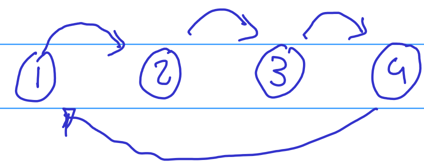
Incontro 1 del 1-10-21

Problema del Page-Rank

Grafi (V, A)
orientati
(diretti)



$H = (h_{ij})$ matrice di adiacenza

$$h_{ij} = \begin{cases} 1 & i \to j \\ 0 & \text{altrimenti} \end{cases}$$

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

Web: nodi $\leftrightarrow$ pagine web

$i \to j$ se la pagina i ha un link alla pagina j

$n \simeq 10^{10}$

10^{20} numeri da memorizzare

H si può memorizzare in modalità sparsa

cioè si memorizzano le coppie (i, j) : $a_{ij} \neq 0$

Ordinare le pagine (nodi del grafo) in base alla loro importanza.

- Ogni nodo i ha la sua importanza w_i
- Un nodo i distribuisce la sua importanza in parti uguali a tutti i nodi a cui punta

Modello matematico:

$$He = d \qquad e = (1, 1, \dots, 1)^T$$

d_i = numero di archi che partono da i

$\boxed{\text{Supponiamo } d_i \neq 0 \quad i = 1, \dots, n}$

$$\boxed{w_j = \sum_{i=1}^{n} h_{ij} \frac{w_i}{d_i}}$$

$h_{ij} = 1$ se $i \to j$

$$w^T = w^T (D^{-1} H) \qquad D = \begin{bmatrix} d_1 & & & 0 \\ & d_2 & & \\ & & \ddots & \\ 0 & & & d_n \end{bmatrix}$$

w è autovettore sinistro di $D^{-1}H$ corrispondente a $\lambda = 1$

Verifica della coerenza del modello

$M := D^{-1} H \qquad d = He, \; D = \mathrm{diag}(d)$

M ha autovalore 1?

$$Me = D^{-1} He = D^{-1} d = e$$

e: autovettore destro corrispond. a $\lambda = 1$

$$\|M\|_\infty = \max_i \sum_j |m_{ij}| = \max_i \sum_j m_{ij} = \max_i \sum_j \frac{h_{ij}}{d_i}$$

$$= \max_i \frac{d_i}{d_i} = 1$$

$\forall \lambda_i$ autovalore di M vale $|\lambda_i| \leq \|M\|_\infty = 1$

1 è autovalore di modulo massimo

Il modello non si applica se $d_i = 0$ per qualche i

Dangling Nodes

[cercare con Google: SNAP Stanford]

Estensione del modello nel caso di dangling nodes

La matrice H viene sostituita da $\hat{H}$

$$\hat{H} = H + u\, e^T \qquad u_i = \begin{cases} 1 & \text{se } d_i = 0 \\ 0 & \text{se } d_i \neq 0 \end{cases}$$

$$H = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix} \qquad u = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad \hat{H} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

Seconda modalità: si aggiunge al grafo un nodo fittizio che punta tutti gli altri nodi ed è puntato da tutti

$$\hat{H} = \left[\begin{array}{ccccc|c} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ \hline 1 & 1 & 1 & 1 & 1 & 0 \end{array}\right]$$

D'ora in poi lavoriamo con $\hat{H}$

$$\hat{M} = \hat{D}^{-1} \hat{H}, \qquad \hat{D} = \operatorname{diag}(\hat{d}), \qquad \hat{d} = \hat{H} e$$

$$\hat{M} e = e \qquad \pi^T \hat{M} = \pi^T$$

$\hat{M}$ ha autov. $\lambda = 1$ e gli altri autov. $|\lambda_i| \leq 1$

Se fosse $|\lambda_i| < 1$ allora

$$S^{-1} \hat{M} S = [\, J \,] = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & \tilde{J} \end{array}\right] \qquad \rho(\tilde{J}) < 1$$

$$S^{-1} \hat{M}^k S = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & \tilde{J}^k \end{array}\right]$$

$$\boxed{\begin{array}{l} w^T \hat{M} = w^T \\ S^{-1} \hat{M} S = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & \tilde{J}^n \end{array}\right] \\ e_1^T S^{-1} \hat{M} S = e_1^T \\ e_1^T S^{-1} \hat{M} = e_1^T S^{-1} \end{array}}$$

$$\lim_{n \to \infty} S^{-1} \hat{M}^n S = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & 0 \end{array}\right] = e_1 e_1^T$$

$$\lim \hat{M}^n = (S e_1) \cdot (e_1^T S^{-1}) = e \cdot w^T = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} [w_1\ w_2 \dots w_n]$$

$$w^T e = 1$$

$$\hat{M} e = e, \quad w^T \hat{M} = w^T$$

$$\begin{cases} x^{(0)} \text{ arbitrario} : x^{(0)T} e = 1 \\ \underline{x^{(k+1)T} = x^{(k)T} \hat{M}} \end{cases} \qquad x^{(k)T} = x^{(0)T} \hat{M}^k$$

$$\longrightarrow x^{(0)T} e\, w^T = w^T$$

(☆) Se $\lambda = 1$ é l'unico autovalore di modulo 1 e tutti gli altri autovalori hanno modulo minore di 1 allora

il METODO DELLE POTENZE fornisce la soluzione computazionale al problema del PageRank.

Purtroppo la condizione (☆) NON SEMPRE è verificata

Siamo costretti a modificare il modello.

Abbiamo bisogno di un teorema:

Teorema di Perron-Frobenius

Sia $A = (a_{ij})$ matrice $n \times n$ $a_{i,j} \geq 0$
allora $\exists$ autovalore di A
uguale a $\rho(A) \geq 0$, esistono in corrispondenza
un autovettore sinistro y e uno destro x :
$x_i \geq 0$, $y_i \geq 0$, $i = 1, \dots, n$

$$A x = \rho(A) x, \qquad y^T A = \rho(A) y^T$$

Se inoltre A è irriducibile allora $\rho(A) > 0$

$\rho(A)$ è autovalore semplice, e $x_i > 0$, $y_i > 0$ $i = 1, \dots, n$

Se $A : a_{ij} > 0$ $\forall i, j = 1, \dots, n$ allora

ogni altro autovalore $\lambda_i \neq \rho(A)$ è: $|\lambda_i| < \rho(A)$

Modificheremo il modello in modo da soddisfare
la terza versione del teorema di Perron-Frobenius