

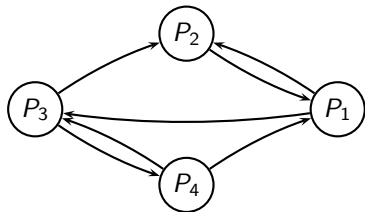
Un Esempio di Modellistica Matematica: Page Ranking di Google

Marco Verani

MOX, Dipartimento di Matematica, Politecnico di Milano
`marco.verani@polimi.it`

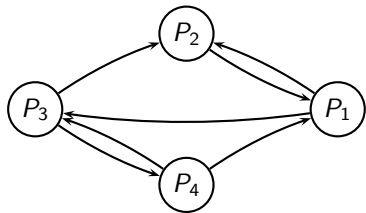
Milano, 16 dicembre 2011

Un Modello Matematico della rete internet (I)



- Siano P_1, \dots, P_4 pagine web
- Rappresentazione matematica dei link presenti fra essi \rightarrow **grafo orientato**

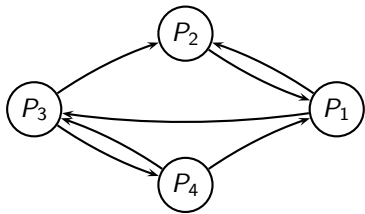
Un Modello Matematico della rete internet (I)



- Siano P_1, \dots, P_4 pagine web
- Rappresentazione matematica dei link presenti fra essi \rightarrow grafo orientato

Page ranking: stabilire una classifica d'importanza tra le pagine web P_1, \dots, P_4

Un Modello Matematico della rete internet (I)

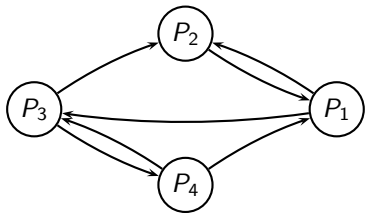


- Siano P_1, \dots, P_4 pagine web
- Rappresentazione matematica dei link presenti fra essi \rightarrow grafo orientato

Page ranking: stabilire una classifica d'importanza tra le pagine web P_1, \dots, P_4

Criteri d'importanza:

Un Modello Matematico della rete internet (I)



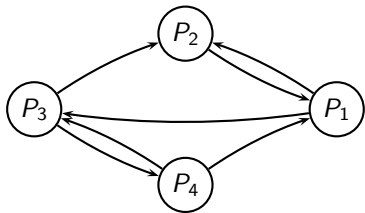
- Siano P_1, \dots, P_4 pagine web
- Rappresentazione matematica dei link presenti fra essi → grafo orientato

Page ranking: stabilire una classifica d'importanza tra le pagine web P_1, \dots, P_4

Criteri d'importanza:

- se una pagina web A ha un collegamento (link) verso una pagina B , questo è interpretato come un voto di A in favore di B (cioè alza B in graduatoria);

Un Modello Matematico della rete internet (I)



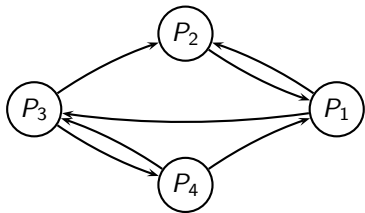
- Siano P_1, \dots, P_4 pagine web
- Rappresentazione matematica dei link presenti fra essi \rightarrow **grafo orientato**

Page ranking: stabilire una classifica d'importanza tra le pagine web P_1, \dots, P_4

Criteri d'importanza:

- se una pagina web A ha un collegamento (link) verso una pagina B , questo è interpretato come un voto di A in favore di B (cioè alza B in graduatoria);
- i votanti non sono tutti uguali: il voto di chi è alto in classifica (cioè ha ricevuto molti link) vale maggiormente di quello di chi è in basso;

Un Modello Matematico della rete internet (I)



- Siano P_1, \dots, P_4 pagine web
- Rappresentazione matematica dei link presenti fra essi \rightarrow **grafo orientato**

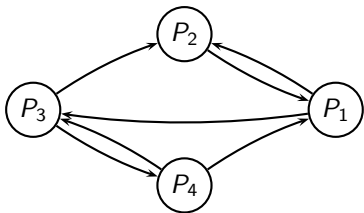
Page ranking: stabilire una classifica d'importanza tra le pagine web P_1, \dots, P_4

Criteri d'importanza:

- se una pagina web A ha un collegamento (link) verso una pagina B , questo è interpretato come un voto di A in favore di B (cioè alza B in graduatoria);
- i votanti non sono tutti uguali: il voto di chi è alto in classifica (cioè ha ricevuto molti link) vale maggiormente di quello di chi è in basso;
- un votante che esprime molti voti (cioè una pagina ha molti link verso altre pagine) perde importanza.

Un Modello Matematico della rete internet (II)

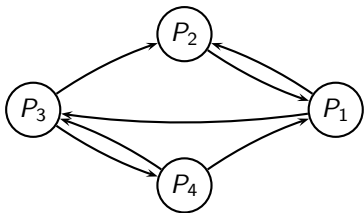
Grafo orientato \Leftrightarrow matrice (tabella) A



$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}$$

Un Modello Matematico della rete internet (II)

Grafo orientato \Leftrightarrow matrice (tabella) A

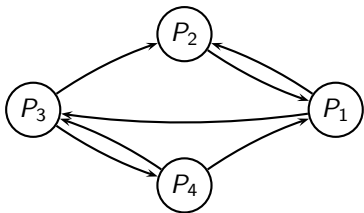


$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}$$

- a_{ij} si riferisce al link da P_j a P_i

Un Modello Matematico della rete internet (II)

Grafo orientato \Leftrightarrow matrice (tabella) A

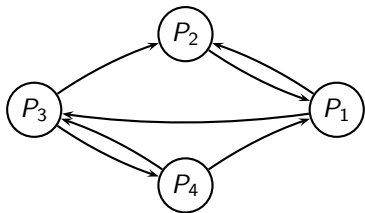


$$A = \begin{pmatrix} 0 & a_{12} & a_{13} & a_{14} \\ a_{21} & 0 & a_{23} & a_{24} \\ a_{31} & a_{32} & 0 & a_{34} \\ a_{41} & a_{42} & a_{43} & 0 \end{pmatrix}$$

- a_{ij} si riferisce al link da P_j a P_i
- $a_{11} = 0$: non c'è link da P_1 a P_1 ($a_{22} = \dots = a_{44} = 0$)

Un Modello Matematico della rete internet (II)

Grafo orientato \Leftrightarrow matrice (tabella) A

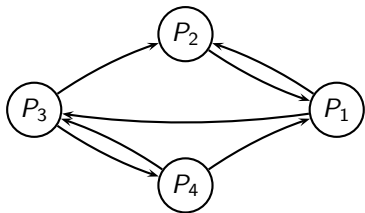


$$A = \begin{pmatrix} 0 & a_{12} & a_{13} & a_{14} \\ a_{21} & 0 & a_{23} & a_{24} \\ a_{31} & a_{32} & 0 & a_{34} \\ a_{41} & a_{42} & a_{43} & 0 \end{pmatrix}$$

- a_{ij} si riferisce al link da P_j a P_i
- $a_{11} = 0$: non c'e' link da P_1 a P_1 ($a_{22} = \dots = a_{44} = 0$)
- $a_{32} = 0$: non c'e' link da P_2 a P_3 ($a_{41} = a_{42} = a_{32} = a_{13} = a_{24} = 0$)

Un Modello Matematico della rete internet (II)

Grafo orientato \Leftrightarrow matrice (tabella) A

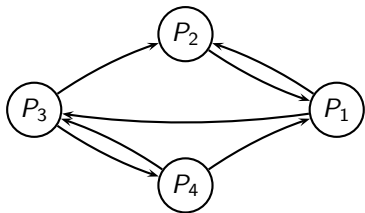


$$A = \begin{pmatrix} 0 & a_{12} & 0 & a_{14} \\ a_{21} & 0 & a_{23} & 0 \\ a_{31} & 0 & 0 & a_{34} \\ 0 & 0 & a_{43} & 0 \end{pmatrix}$$

- a_{ij} si riferisce al link da P_j a P_i
- $a_{11} = 0$: non c'e' link da P_1 a P_1 ($a_{22} = \dots = a_{44} = 0$)
- $a_{32} = 0$: non c'e' link da P_2 a P_3 ($a_{41} = a_{42} = a_{32} = a_{13} = a_{24} = 0$)

Un Modello Matematico della rete internet (II)

Grafo orientato \Leftrightarrow matrice (tabella) A

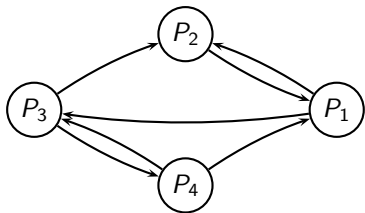


$$A = \begin{pmatrix} 0 & a_{12} & 0 & a_{14} \\ a_{21} & 0 & a_{23} & 0 \\ a_{31} & 0 & 0 & a_{34} \\ 0 & 0 & a_{43} & 0 \end{pmatrix}$$

- a_{ij} si riferisce al link da P_j a P_i
- $a_{11} = 0$: non c'è link da P_1 a P_1 ($a_{22} = \dots = a_{44} = 0$)
- $a_{32} = 0$: non c'è link da P_2 a P_3 ($a_{41} = a_{42} = a_{32} = a_{13} = a_{24} = 0$)
- $a_{21} = \frac{1}{2}$: $1 \rightarrow$ c'è link da P_1 a P_2 ; $2 \rightarrow$ numero link uscenti da P_1 ;

Un Modello Matematico della rete internet (II)

Grafo orientato \Leftrightarrow matrice (tabella) A

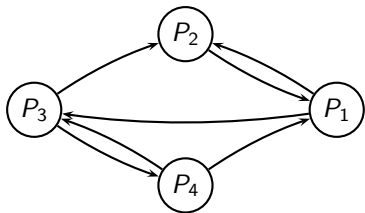


$$A = \begin{pmatrix} 0 & a_{12} & 0 & a_{14} \\ a_{21} & 0 & a_{23} & 0 \\ a_{31} & 0 & 0 & a_{34} \\ 0 & 0 & a_{43} & 0 \end{pmatrix}$$

- a_{ij} si riferisce al link da P_j a P_i
- $a_{11} = 0$: non c'è link da P_1 a P_1 ($a_{22} = \dots = a_{44} = 0$)
- $a_{32} = 0$: non c'è link da P_2 a P_3 ($a_{41} = a_{42} = a_{32} = a_{13} = a_{24} = 0$)
- $a_{21} = \frac{1}{2}$: 1 \rightarrow c'è link da P_1 a P_2 ; 2 \rightarrow numero link uscenti da P_1 ;
- $a_{31} = \frac{1}{2}$

Un Modello Matematico della rete internet (II)

Grafo orientato \Leftrightarrow matrice (tabella) A

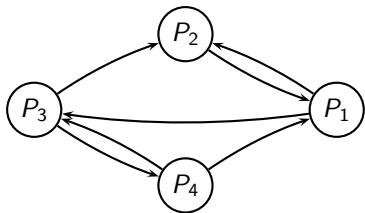


$$A = \begin{pmatrix} 0 & a_{12} & 0 & a_{14} \\ 1/2 & 0 & a_{23} & 0 \\ 1/2 & 0 & 0 & a_{34} \\ 0 & 0 & a_{43} & 0 \end{pmatrix}$$

- a_{ij} si riferisce al link da P_j a P_i
- $a_{11} = 0$: non c'è link da P_1 a P_1 ($a_{22} = \dots = a_{44} = 0$)
- $a_{32} = 0$: non c'è link da P_2 a P_3 ($a_{41} = a_{42} = a_{32} = a_{13} = a_{24} = 0$)
- $a_{21} = \frac{1}{2}$: 1 \rightarrow c'è link da P_1 a P_2 ; 2 \rightarrow numero link uscenti da P_1 ;
- $a_{31} = \frac{1}{2}$

Un Modello Matematico della rete internet (II)

Grafo orientato \Leftrightarrow matrice (tabella) A

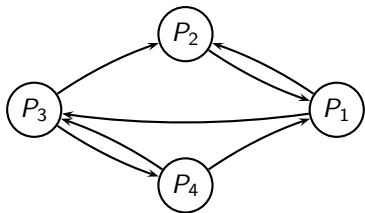


$$A = \begin{pmatrix} 0 & 1 & 0 & a_{14} \\ 1/2 & 0 & a_{23} & 0 \\ 1/2 & 0 & 0 & a_{34} \\ 0 & 0 & a_{43} & 0 \end{pmatrix}$$

- a_{ij} si riferisce al link da P_j a P_i
- $a_{11} = 0$: non c'è link da P_1 a P_1 ($a_{22} = \dots = a_{44} = 0$)
- $a_{32} = 0$: non c'è link da P_2 a P_3 ($a_{41} = a_{42} = a_{32} = a_{13} = a_{24} = 0$)
- $a_{21} = \frac{1}{2}$: 1 \rightarrow c'è link da P_1 a P_2 ; 2 \rightarrow numero link uscenti da P_1 ;
- $a_{31} = \frac{1}{2}$

Un Modello Matematico della rete internet (II)

Grafo orientato \Leftrightarrow matrice (tabella) A



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

- a_{ij} si riferisce al link da P_j a P_i
- $a_{11} = 0$: non c'è link da P_1 a P_1 ($a_{22} = \dots = a_{44} = 0$)
- $a_{32} = 0$: non c'è link da P_2 a P_3 ($a_{41} = a_{42} = a_{32} = a_{13} = a_{24} = 0$)
- $a_{21} = \frac{1}{2}$: 1 \rightarrow c'è link da P_1 a P_2 ; 2 \rightarrow numero link uscenti da P_1 ;
- $a_{31} = \frac{1}{2}$

Formalizziamo i criteri d'importanza

x_i denota l'importanza (ranking) della pagina P_i con $0 \leq x_i \leq 1$

Formalizziamo i criteri d'importanza

x_i denota l'importanza (ranking) della pagina P_i con $0 \leq x_i \leq 1$

Formalizzazione matematica dei criteri d'importanza scelti



$x_i =$ somma "pesata" dei ranking delle pagine che puntano P_i

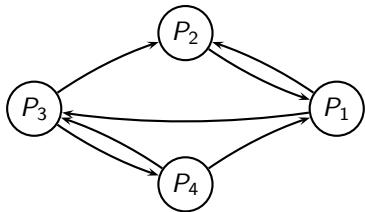
Formalizziamo i criteri d'importanza

x_i denota l'importanza (ranking) della pagina P_i con $0 \leq x_i \leq 1$

Formalizzazione matematica dei criteri d'importanza scelti



$x_i =$ somma "pesata" dei ranking delle pagine che puntano P_i



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

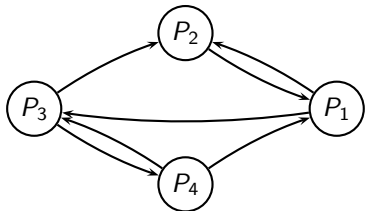
Formalizziamo i criteri d'importanza

x_i denota l'importanza (ranking) della pagina P_i con $0 \leq x_i \leq 1$

Formalizzazione matematica dei criteri d'importanza scelti



$x_i =$ somma "pesata" dei ranking delle pagine che puntano P_i



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

$$x_1 = 0 \cdot x_0 + 1 \cdot x_2 + 0 \cdot x_3 + 1/2 \cdot x_4$$

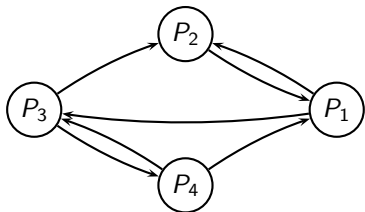
Formalizziamo i criteri d'importanza

x_i denota l'importanza (ranking) della pagina P_i con $0 \leq x_i \leq 1$

Formalizzazione matematica dei criteri d'importanza scelti



$x_i =$ somma "pesata" dei ranking delle pagine che puntano P_i



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

$$x_1 = 0 \cdot x_0 + 1 \cdot x_2 + 0 \cdot x_3 + 1/2 \cdot x_4$$

$$x_2 = 1/2 \cdot x_0 + 0 \cdot x_2 + 1/2 \cdot x_3 + 0 \cdot x_4$$

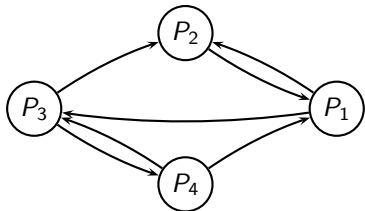
Formalizziamo i criteri d'importanza

x_i denota l'importanza (ranking) della pagina P_i con $0 \leq x_i \leq 1$

Formalizzazione matematica dei criteri d'importanza scelti



$x_i =$ somma "pesata" dei ranking delle pagine che puntano P_i



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

$$x_1 = 0 \cdot x_0 + 1 \cdot x_2 + 0 \cdot x_3 + 1/2 \cdot x_4$$

$$x_2 = 1/2 \cdot x_0 + 0 \cdot x_2 + 1/2 \cdot x_3 + 0 \cdot x_4$$

$$x_3 = 1/2 \cdot x_0 + 0 \cdot x_2 + 0 \cdot x_3 + 1/2 \cdot x_4$$

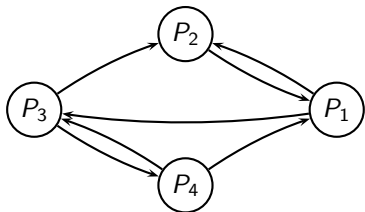
Formalizziamo i criteri d'importanza

x_i denota l'importanza (ranking) della pagina P_i con $0 \leq x_i \leq 1$

Formalizzazione matematica dei criteri d'importanza scelti



$x_i =$ somma "pesata" dei ranking delle pagine che puntano P_i



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

$$x_1 = 0 \cdot x_0 + 1 \cdot x_2 + 0 \cdot x_3 + 1/2 \cdot x_4$$

$$x_2 = 1/2 \cdot x_0 + 0 \cdot x_2 + 1/2 \cdot x_3 + 0 \cdot x_4$$

$$x_3 = 1/2 \cdot x_0 + 0 \cdot x_2 + 0 \cdot x_3 + 1/2 \cdot x_4$$

$$x_4 = 0 \cdot x_0 + 0 \cdot x_2 + 1/2 \cdot x_3 + 0 \cdot x_4$$

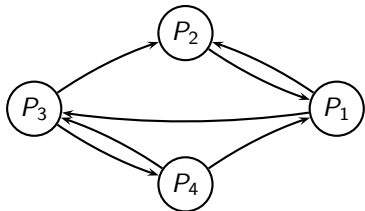
Formalizziamo i criteri d'importanza

x_i denota l'importanza (ranking) della pagina P_i con $0 \leq x_i \leq 1$

Formalizzazione matematica dei criteri d'importanza scelti



$x_i =$ somma "pesata" dei ranking delle pagine che puntano P_i



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

$$x_1 = 0 \cdot x_0 + 1 \cdot x_2 + 0 \cdot x_3 + 1/2 \cdot x_4$$

$$x_2 = 1/2 \cdot x_0 + 0 \cdot x_2 + 1/2 \cdot x_3 + 0 \cdot x_4$$

$$x_3 = 1/2 \cdot x_0 + 0 \cdot x_2 + 0 \cdot x_3 + 1/2 \cdot x_4$$

$$x_4 = 0 \cdot x_0 + 0 \cdot x_2 + 1/2 \cdot x_3 + 0 \cdot x_4$$

$$\mathbf{x} = \mathbf{Ax}$$

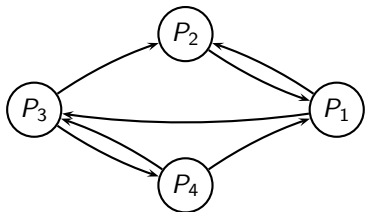
Formalizziamo i criteri d'importanza

x_i denota l'importanza (ranking) della pagina P_i con $0 \leq x_i \leq 1$

Formalizzazione matematica dei criteri d'importanza scelti



$x_i =$ somma "pesata" dei ranking delle pagine che puntano P_i



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

$$\mathbf{x} = [0.3529 \ 0.2941 \ 0.2353 \ 0.1176]$$

P_1 è la pagina più "importante" !