

Passejos aleatoris en grafs

Gerard Castro, Kim López, Gal·la Mora, Arnau Mas

5 de desembre de 2018

Esquema

- 1 Introducció
- 2 Distribucions estacionàries
- 3 PageRank
- 4 El teorema de Pólya

Introducció

Definició

Definició

Un *passeig aleatori* a un graf G és una seqüència de vèrtexs $v_1, v_2, \dots, v_n, \dots$, tal que $v_k v_{k+1}$ es tria uniformement d'entre les arestes incidents a v_k .

Aspectes probabilístics

$V(t)$ és la variable aleatòria que representa la posició del passeig a l'instant t

Aspectes probabilístics

$V(t)$ és la variable aleatòria que representa la posició del passeig a l'instant t
Les probabilitats de transició

$$P(V(t) = u \mid V(t-1) = v)$$

determinen el passeig:

Aspectes probabilístics

$V(t)$ és la variable aleatòria que representa la posició del passeig a l'instant t
Les probabilitats de transició

$$P(V(t) = u \mid V(t-1) = v)$$

determinen el passeig:

$$P(V(t) = u) = \sum_{v \in V(G)} P(V(t) = u \mid V(t-1) = v) P(V(t-1) = v).$$

Probabilitats de transició

Com que

$$P(V(t) = u \mid V(t-1) = v) = \frac{a(u, v)}{\text{gr}(v)},$$

Probabilitats de transició

Com que

$$P(V(t) = u \mid V(t-1) = v) = \frac{a(u, v)}{\text{gr}(v)},$$

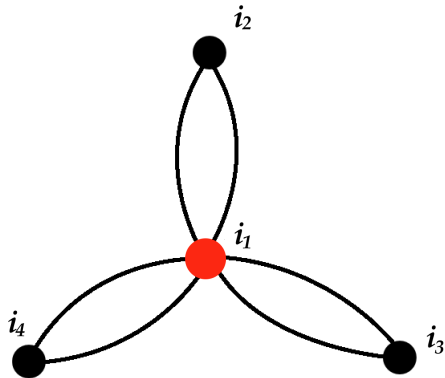
tenim

$$P(V(t) = u) = \sum_{v \in V(G)} \frac{a(u, v)}{\text{gr}(v)} P(V(t-1) = v).$$

Un exemple

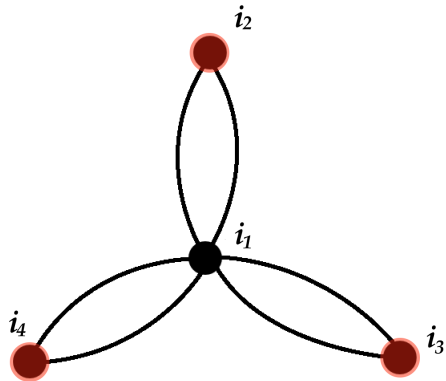
Un exemple

$$P(V(0) = i_1) = 1$$



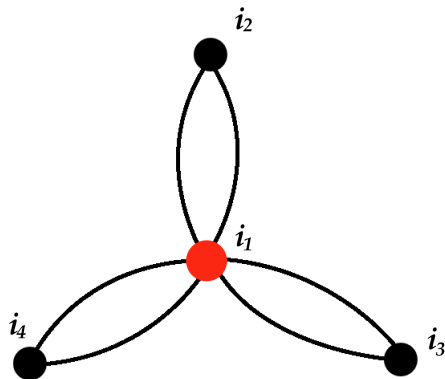
Un exemple

$$\begin{aligned} P(V(1) = i_2) &= P(V(1) = i_3) \\ &= P(V(1) = i_4) = \frac{1}{3} \end{aligned}$$



Un exemple

$$P(V(3) = i_1) = 1$$



Matriu de transició

Definim $\mathbf{p}_t \in \mathbb{R}^{V(G)}$ com $\mathbf{p}_t(u) = P(V(t) = u)$.

Matriu de transició

Definim $\mathbf{p}_t \in \mathbb{R}^{V(G)}$ com $\mathbf{p}_t(u) = P(V(t) = u)$. Aleshores, en forma matricial

$$\mathbf{p}_t = AD^{-1}\mathbf{p}_{t-1}.$$

A és la matriu d'adjacència i D la matriu dels graus. $P = AD^{-1}$ és la *matriu de transició*.

Matriu de transició

Definim $\mathbf{p}_t \in \mathbb{R}^{V(G)}$ com $\mathbf{p}_t(u) = P(V(t) = u)$. Aleshores, en forma matricial

$$\mathbf{p}_t = AD^{-1}\mathbf{p}_{t-1}.$$

A és la matriu d'adjacència i D la matriu dels graus. $P = AD^{-1}$ és la *matriu de transició*.

Per tant

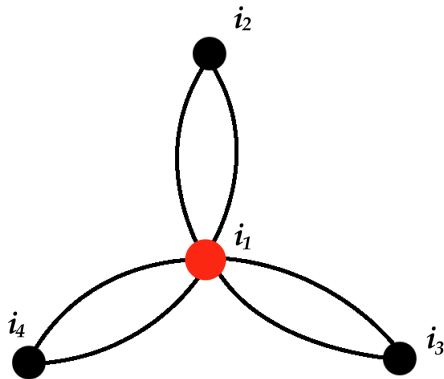
$$\mathbf{p}_t = P^t \mathbf{p}_0.$$

Un exemple

Un exemple

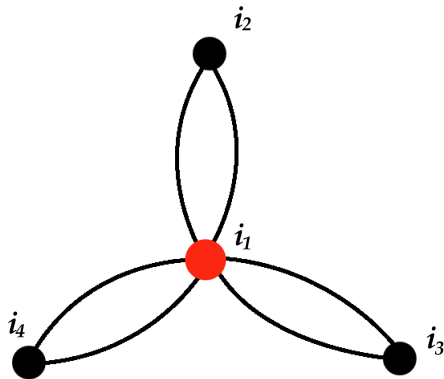
$$A = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \quad D = \begin{pmatrix} 6 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

$$P = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1/3 & 0 & 0 & 0 \\ 1/3 & 0 & 0 & 0 \\ 1/3 & 0 & 0 & 0 \end{pmatrix}$$



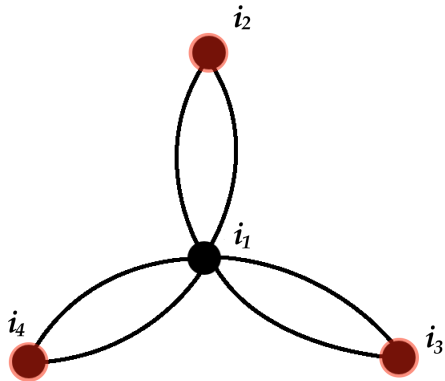
Un exemple

$$\mathbf{p}_0 = (1, 0, 0, 0)$$



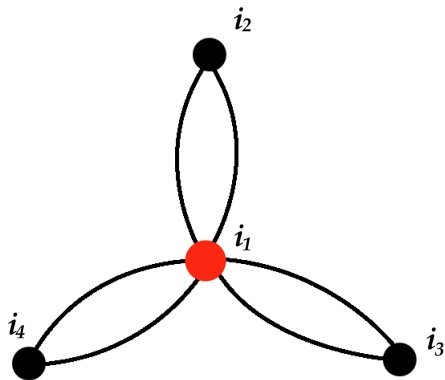
Un exemple

$$\mathbf{p}_1 = \left(0, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) = P\mathbf{p}_0$$



Un exemple

$$\mathbf{p}_2 = (1, 0, 0, 0) = P\mathbf{p}_1 = P^2\mathbf{p}_0$$



Distribucions estacionàries

Distribució estacionària

Com és el passeig per a temps grans?

Distribució estacionària

Com és el passeig per a temps grans?

Definició

Una distribució és *estacionària* si $P(V(t) = u) = P(V(t - 1) = u)$ per tot $u \in V(G)$.

Distribució estacionària

Com és el passeig per a temps grans?

Definició

Una distribució és *estacionària* si $P(V(t) = u) = P(V(t-1) = u)$ per tot $u \in V(G)$.

En termes de la matriu de transició,

$$P\pi = \pi.$$

Distribució estacionària

Com és el passeig per a temps grans?

Definició

Una distribució és *estacionària* si $P(V(t) = u) = P(V(t - 1) = u)$ per tot $u \in V(G)$.

En termes de la matriu de transició,

$$P\pi = \pi.$$

Per tant

$$\pi(u) = \sum_{v \in V(G)} \frac{a(u, v)}{\text{gr}(v)}$$

Existència i unicitat de distribucions estacionàries

Teorema

Tot graf connex admet una única distribució estacionària.

Existència i unicitat de distribucions estacionàries

Teorema

Tot graf connex admet una única distribució estacionària.

Demostració.

π estacionària. Veurem que $\pi(u) \propto \text{gr } u$.

Existència i unicitat de distribucions estacionàries

Teorema

Tot graf connex admet una única distribució estacionària.

Demostració.

π estacionària. Veurem que $\pi(u) \propto \text{gr } u$. Sigui u^* tal que $\frac{\pi(u^*)}{\text{gr}(u^*)}$ és màxim:

Existència i unicitat de distribucions estacionàries

Teorema

Tot graf connex admet una única distribució estacionària.

Demostració.

π estacionària. Veurem que $\pi(u) \propto \text{gr } u$. Sigui u^* tal que $\frac{\pi(u^*)}{\text{gr}(u^*)}$ és màxim:

$$\pi(u^*) = \sum_{v \in V(G)} \frac{a(u^*, v)}{\text{gr}(v)} \pi(v)$$

Existència i unicitat de distribucions estacionàries

Teorema

Tot graf connex admet una única distribució estacionària.

Demostració.

π estacionària. Veurem que $\pi(u) \propto \text{gr } u$. Sigui u^* tal que $\frac{\pi(u^*)}{\text{gr}(u^*)}$ és màxim:

$$\pi(u^*) = \sum_{v \in V(G)} \frac{a(u^*, v)}{\text{gr}(v)} \pi(v) \leq \frac{\pi(u^*)}{\text{gr}(u^*)} \sum_{v \in V(G)} a(u^*, v)$$

Existència i unicitat de distribucions estacionàries

Teorema

Tot graf connex admet una única distribució estacionària.

Demostració.

π estacionària. Veurem que $\pi(u) \propto \text{gr } u$. Sigui u^* tal que $\frac{\pi(u^*)}{\text{gr}(u^*)}$ és màxim:

$$\pi(u^*) = \sum_{v \in V(G)} \frac{a(u^*, v)}{\text{gr}(v)} \pi(v) \leq \frac{\pi(u^*)}{\text{gr}(u^*)} \sum_{v \in V(G)} a(u^*, v) = \pi(u^*)$$

Aleshores

$$\sum_{v \in V(G)} a(u^*, v) \frac{\pi(v)}{\text{gr}(v)} = \sum_{v \in V(G)} a(u^*, v) \frac{\pi(u^*)}{\text{gr}(u^*)}.$$



Existència i unicitat de distribucions estacionàries

Demostració.

Per tant $\frac{\pi(v)}{\text{gr}(v)} = \frac{\pi(u^*)}{\text{gr}(u^*)}$ per tot v adjacent a u^* .

Existència i unicitat de distribucions estacionàries

Demostració.

Per tant $\frac{\pi(v)}{\text{gr}(v)} = \frac{\pi(u^*)}{\text{gr}(u^*)}$ per tot v adjacent a u^* . Podem estendre l'argument a tots els vèrtexs de G fent servir que és connex.

Existència i unicitat de distribucions estacionàries

Demostració.

Per tant $\frac{\pi(v)}{\text{gr}(v)} = \frac{\pi(u^*)}{\text{gr}(u^*)}$ per tot v adjacent a u^* . Podem estendre l'argument a tots els vèrtexs de G fent servir que és connex.

Com que $\pi(u) \propto \text{gr}(u)$ aleshores $\pi(u) = C \text{gr}(u)$.

Existència i unicitat de distribucions estacionàries

Demostració.

Per tant $\frac{\pi(v)}{\text{gr}(v)} = \frac{\pi(u^*)}{\text{gr}(u^*)}$ per tot v adjacent a u^* . Podem estendre l'argument a tots els vèrtexs de G fent servir que és connex.

Com que $\pi(u) \propto \text{gr}(u)$ aleshores $\pi(u) = C \text{gr}(u)$.

Si imposem que $\sum_{u \in V(G)} \pi(u) = 1$ podem determinar C i

$$\pi(u) = \frac{\text{gr}(u)}{\sum_{v \in V(G)} \text{gr}(v)}$$

Existència i unicitat de distribucions estacionàries

Demostració.

Per tant $\frac{\pi(v)}{\text{gr}(v)} = \frac{\pi(u^*)}{\text{gr}(u^*)}$ per tot v adjacent a u^* . Podem estendre l'argument a tots els vèrtexs de G fent servir que és connex.

Com que $\pi(u) \propto \text{gr}(u)$ aleshores $\pi(u) = C \text{gr}(u)$.

Si imposem que $\sum_{u \in V(G)} \pi(u) = 1$ podem determinar C i

$$\pi(u) = \frac{\text{gr}(u)}{\sum_{v \in V(G)} \text{gr}(v)} = \frac{\text{gr}(u)}{2|E(G)|}$$



Convergència a la distribució estacionària

Teorema

Tota distribució de probabilitats en un graf connex no bipartit convergeix a la distribució estacionària.

Convergència a la distribució estacionària

Teorema

Tota distribució de probabilitats en un graf connex no bipartit convergeix a la distribució estacionària.

Demostració.

P és similar a una matriu simètrica:

$$D^{-1/2}PD^{1/2} = D^{-1/2}AD^{-1/2}.$$

Convergència a la distribució estacionària

Teorema

Tota distribució de probabilitats en un graf connex no bipartit convergeix a la distribució estacionària.

Demostració.

P és similar a una matriu simètrica:

$$D^{-1/2}PD^{1/2} = D^{-1/2}AD^{-1/2}.$$

Per tant diagonalitza.

Convergència a la distribució estacionària

Teorema

Tota distribució de probabilitats en un graf connex no bipartit convergeix a la distribució estacionària.

Demostració.

P és similar a una matriu simètrica:

$$D^{-1/2}PD^{1/2} = D^{-1/2}AD^{-1/2}.$$

Per tant diagonalitza. π és l'únic vector propi de valor propi 1. Posem

$$\mathbf{p}_0 = \alpha_1 \pi + \sum_k \alpha_k \mathbf{v}_k$$



Convergència a la distribució estacionària

Demostració.

Es pot demostrar que $\alpha_1 = 1$.

Convergència a la distribució estacionària

Demostració.

Es pot demostrar que $\alpha_1 = 1$. Aleshores

$$P^t \mathbf{p}_0 = \boldsymbol{\pi} + \sum_k \lambda_k^t \alpha_k \mathbf{v}_k.$$

Convergència a la distribució estacionària

Demostració.

Es pot demostrar que $\alpha_1 = 1$. Aleshores

$$P^t \mathbf{p}_0 = \boldsymbol{\pi} + \sum_k \lambda_k^t \alpha_k \mathbf{v}_k.$$

Per qualsevol graf, $|\lambda_k| \leq 1$ i per un graf no bipartit $\lambda_k > -1$.

Convergència a la distribució estacionària

Demostració.

Es pot demostrar que $\alpha_1 = 1$. Aleshores

$$P^t \mathbf{p}_0 = \boldsymbol{\pi} + \sum_k \lambda_k^t \alpha_k \mathbf{v}_k.$$

Per qualsevol graf, $|\lambda_k| \leq 1$ i per un graf no bipartit $\lambda_k > -1$. Per tant

$$P^t \mathbf{p}_0 \xrightarrow{t \rightarrow \infty} \boldsymbol{\pi}.$$

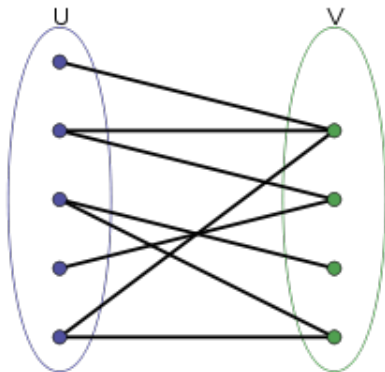


Grafs bipartits

Pels grafs bipartits la convergència no es dóna, tal i com hem vist a l'exemple al principi.

Grafs bipartits

Pels grafs bipartits la convergència no es dóna, tal i com hem vist a l'exemple al principi.



Grafs bipartits

Si permetem repetir vèrtexs aleshores obtenim un *passeig aleatori mandrós*.

Grafs bipartits

Si permetem repetir vèrtexs aleshores obtenim un *passeig aleatori mandrós*. La matriu de transició és

$$\frac{1}{2}(I + P).$$

PageRank

L'algorisme PageRank

- Google desenvolupa un algoritme per determinar la rellevància de pàgines web.

L'algorisme PageRank

- Google desenvolupa un algoritme per determinar la rellevància de pàgines web.
- Considera el graf dirigit dels híperenllaços entre pàgines i assigna a cada pàgina la probabilitat que li correspon a la distribució estacionària.

L'algorisme PageRank

- Google desenvolupa un algoritme per determinar la rellevància de pàgines web.
- Considera el graf dirigit dels híperenllaços entre pàgines i assigna a cada pàgina la probabilitat que li correspon a la distribució estacionària.
- Per garantir la convergència es fan servir passejos mandrosos.

L'algorisme PageRank

- Google desenvolupa un algoritme per determinar la rellevància de pàgines web.
- Considera el graf dirigit dels híperenllaços entre pàgines i assigna a cada pàgina la probabilitat que li correspon a la distribució estacionària.
- Per garantir la convergència es fan servir passejos mandrosos.
- Com que el passeig és dirigit ja no tenim l'expressió de la distribució estacionària per a grafs no dirigits.

Probabilitat de teletransport

El caminant es teletransporta a un vèrtex aleatori amb probabilitat α .

Probabilitat de teletransport

El caminant es teletransporta a un vèrtex aleatori amb probabilitat α .

$$P_{\text{PR}} = \frac{\alpha}{2|V(G)|} \mathbf{1} + \frac{1-\alpha}{2} (I + P).$$

Probabilitat de teletransport

El caminant es teletransporta a un vèrtex aleatori amb probabilitat α .

$$P_{\text{PR}} = \frac{\alpha}{2|V(G)|}\mathbf{1} + \frac{1-\alpha}{2}(I + P).$$

A PageRank personalitzat ens trasalladem a un vèrtex concret: obtenim un rànquin relatiu a aquell vèrtex.

Probabilitat de teletransport

El caminant es teletransporta a un vèrtex aleatori amb probabilitat α .

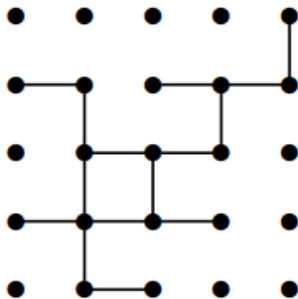
$$P_{\text{PR}} = \frac{\alpha}{2|V(G)|} \mathbf{1} + \frac{1-\alpha}{2}(I + P).$$

A PageRank personalitzat ens trasalladem a un vèrtex concret: obtenim un rànquin relatiu a aquell vèrtex. En general podem teletransportar-nos seguint una distribució qualsevol i obtindrem resultats rellevants a aquesta. En general podem teletransportar-nos seguint una distribució qualsevol i obtindrem resultats rellevants a aquesta.

El teorema de Pólya

Grafs infinits

Podem considerar passejos a grafs infinits. L'exemple més senzill són passejos aleatoris a \mathbb{Z}^d



Passejos recurrents i transitoris

Definició

Anomenarem *probabilitat d'escapament*, p_{esc} , a la probabilitat que el passeig aleatori mai retorni a l'origen.

Passejos recurrents i transitoris

Definició

Anomenarem *probabilitat d'escapament*, p_{esc} , a la probabilitat que el passeig aleatori mai retorni a l'origen.

Definició

Un passeig aleatori és *recurrent* si i només si $p_{\text{esc}} = 0$. Un passeig aleatori és *transitori* si i només si $p_{\text{esc}} > 0$.

Teorema de Pólya

Teorema

Un passeig aleatori simple en una xarxa d -dimensional \mathbb{Z}^d és recurrent si $d = 1$ o $d = 2$, i transitori si $d \geq 3$.

Teorema de Pólya

Teorema

Un passeig aleatori simple en una xarxa d -dimensional \mathbb{Z}^d és recurrent si $d = 1$ o $d = 2$, i transitori si $d \geq 3$.

"A drunk man will find his way home, but a drunk bird may get lost forever."

Shizuo Kakutani