Remember where you came from: Hitting times for second-order random walks

Dario Fasino
Univ. Udine

Napoli, Feb. 15, 2022

Outline

At a glance

Second-order random walks are increasingly popular.
How can we define (and compute) their hitting/return times?
And what properties do they have?

D. F., A. Tonetto, F. Tudisco
Hitting times for second-order random walks.
arXiv:2105.14438 (2021).

Intro

Random walk on a graph $\mathcal{G} = (V, E)$

- Start from a node chosen $i \in V$ with some probability
  - Pick one of the outgoing edges
  - Move to the destination of the edge
  - Repeat.

Applications

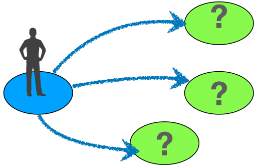
Random walks are widely used in network science to

- model user navigation and diffusive processes
- quantify node centrality and accessibility
- reveal network communities and core-periphery structures.

Random walk on a graph $\mathcal{G} = (V, E)$

- Start from a node chosen $i \in V$ with some probability
  - Pick one of the outgoing edges
  - Move to the destination of the edge
  - Repeat.

First-order random walk



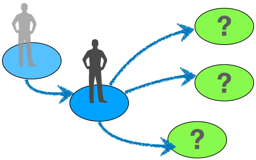
- X_t: node visited at time $t = 0, 1, \ldots$
- x_t: probability vector of X_t
- P: (row stochastic) transition matrix
  $P_{ij} = \mathbb{P}(X_{t+1} = j | X_t = i)$

$$x_{t+1}^{\mathrm{T}} = x_t^{\mathrm{T}} P.$$

Random walk on a graph $\mathcal{G} = (V, E)$

- Start from a node chosen $i \in V$ with some probability
  - Pick one of the outgoing edges
  - Move to the destination of the edge
  - Repeat.

Second-order random walk



- Y_t: joint probability matrix
  $(Y_t)_{ij} = \mathbb{P}(X_t = j, X_{t-1} = i)$
- $\mathbf{P}$: transition tensor
  $\mathbf{P}_{ijk} = \mathbb{P}(X_{t+1} = k | X_t = j, X_{t-1} = i)$

$$\begin{cases} (Y_{t+1})_{jk} = \sum_i \mathbf{P}_{ijk}(Y_t)_{ij} \\ x_{t+1} = \mathbb{1}^{\mathrm{T}}\, Y_{t+1}. \end{cases}$$

Notations

Let $S \subset V$, with $\mathcal{G} = (V, E)$ strongly connected.

- $\tau_{i \to S}$: hitting time to S starting from i.

$$\tau_{i \to S} = \begin{cases} 0 & i \in S \\ 1 + \sum_{k=1}^{n} P_{ik} \tau_{k \to S} & i \notin S. \end{cases}$$

- $\tau_{i \to S}^{+}$: return time to S starting from i.

$$\tau_{i \to S}^{+} = \begin{cases} 1 + \sum_{k=1}^{n} P_{ik} \tau_{k \to S} & i \in S \\ 0 & i \notin S. \end{cases}$$

- $\pi = (\pi_1, \ldots, \pi_n)^{\mathrm{T}}$: the stationary probability vector,

$$\pi > 0, \qquad \pi^{\mathrm{T}} = \pi^{\mathrm{T}} P, \qquad \mathbb{1}^{\mathrm{T}} \pi = 1.$$

For a fixed $S \subset V$, let t^S and b^S be the vectors with the hitting times and return times to S, respectively:

$$t_i^S = \begin{cases} 0 & i \in S \\ \tau_{i \to S} & i \notin S \end{cases} \qquad b_i^S = \begin{cases} \tau^+_{i \to S} & i \in S \\ 0 & i \notin S. \end{cases}$$

The vector t^S solves the singular equation $(I - P)t^S = \mathbb{1} - b^S$.

Hitting time matrix

The hitting time matrix $T = (\tau_{i \to j})$ solves the equation

$$(I - P)T = \mathbb{1}\mathbb{1}^{\mathrm{T}} - \mathrm{Diag}(\rho_1, \ldots, \rho_n)$$

where $\rho_i = \tau^+_{i \to i}$ is the return time to node i.

For a fixed $S \subset V$, let t^S and b^S be the vectors with the hitting times and return times to S, respectively:

$$t_i^S = \begin{cases} 0 & i \in S \\ \tau_{i \to S} & i \notin S \end{cases} \qquad b_i^S = \begin{cases} \tau_{i \to S}^{+} & i \in S \\ 0 & i \notin S. \end{cases}$$

The vector t^S solves the singular equation $(I - P)t^S = \mathbb{1} - b^S$.

Kac's lemma

Let $\rho_S = \sum_{i \in S} \pi_i \tau_{i \to S}^{+}$ be the average return time to S. Then,

$$\rho_S = \left(\sum_{i \in S} \pi_i\right)^{-1}.$$

In particular, $\rho_i = 1/\pi_i$.

From first- to second-order rw.s

Classical random walks suffer from a few drawbacks:
localization, limited expressiveness, slow mixing rates...

Extend 1-order random walk idea by accounting for an earlier step in navigation:

$$\mathbb{P}(X_{n+1} = k | X_n = j, X_{n-1} = i) = p_{ijk}$$

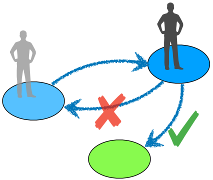
From first- to second-order rw.s

Classical random walks suffer from a few drawbacks:
localization, limited expressiveness, slow mixing rates...

Extend 1-order random walk idea by accounting for an earlier step in navigation:

$$\mathbb{P}(X_{n+1} = k | X_n = j, X_{n-1} = i) = p_{ijk}$$

non-backtracking rw

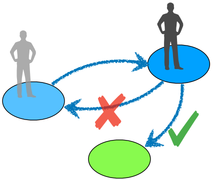


Classical random walks suffer from a few drawbacks: localization, limited expressiveness, slow mixing rates...

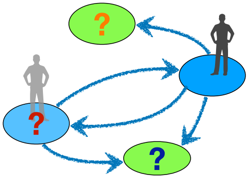
Extend 1-order random walk idea by accounting for an earlier step in navigation:

$$\mathbb{P}(X_{n+1} = k | X_n = j, X_{n-1} = i) = p_{ijk}$$

non-backtracking rw



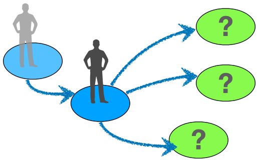
`node2vec` algorithm



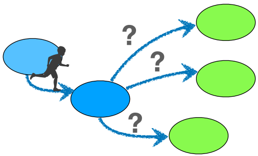
The "lifting" idea: Introduce a walker moving from edge to edge,

$$p_{ijk} = \mathbb{P}(W_{n+1} = (j,k) | W_n = (i,j))$$

This converts the second-order rw on V . . .



. . . to a first-order rw on E

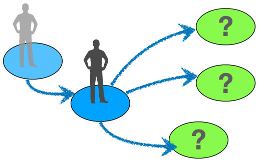


From second- to first-order random walks

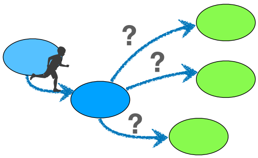
The “lifting” idea: Introduce a walker moving from edge to edge,

$$p_{ijk} = \mathbb{P}(W_{n+1} = (j,k) | W_n = (i,j))$$

This converts the second-order rw on V . . .



. . . to a first-order rw on E

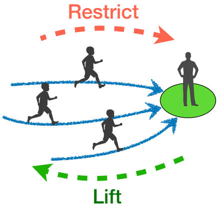


Second-order rw.s on $\mathcal{G} = (V, E)$ correspond to first-order rw.s on $\widehat{\mathcal{G}} = (E, \widehat{E})$, the (directed) line graph of $\mathcal{G}$.

Main results

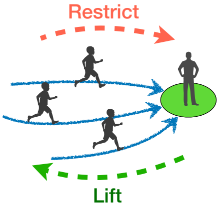
- $L : V \mapsto E$, the "lifting matrix"
- $R : E \mapsto V$, the "restriction matrix"

$$\{W_t\} \underset{L}{\overset{R}{\rightleftarrows}} \{X_t\}$$



- $L : V \mapsto E$, the "lifting matrix"
- $R : E \mapsto V$, the "restriction matrix"

$$\{W_t\} \underset{L}{\overset{R}{\rightleftarrows}} \{X_t\}$$



Theorem

Let $\widehat{P}, \widehat{\pi}$ be the transition matrix and stationary density of $\{W_t\}$. The matrix $P = L\widehat{P}R$ is irreducible, row stochastic, and

$$P_{ij} = \mathbb{P}(X_{t+1} = j | X_t = i), \qquad t \geq 1.$$

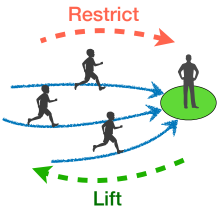
The stationary density of P is $\pi^T = \widehat{\pi}^{\mathrm{T}} R$.

We call P the pullback of $\widehat{P}$.

Main results

- $L : V \mapsto E$, the "lifting matrix"
- $R : E \mapsto V$, the "restriction matrix"

$$\{W_t\} \underset{L}{\overset{R}{\rightleftarrows}} \{X_t\}$$



Theorem

Let $\widetilde{T} = (\widetilde{\tau}_{i \to j})$ be the second-order hitting times matrix for $\{X_t\}$ and $\widehat{T} = (\widehat{\tau}_{e \to f})$ the hitting time matrix for $\{W_t\}$. Then,

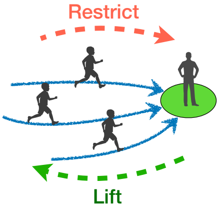
$$\widetilde{T} = L\widehat{T}\,\mathrm{Diag}(a)R - \mathbb{1}b^T,$$

for some (explicitly known) vectors a, b.

Main results

- $L : V \mapsto E$, the "lifting matrix"
- $R : E \mapsto V$, the "restriction matrix"

$$\{W_t\} \underset{L}{\overset{R}{\rightleftarrows}} \{X_t\}$$



Corollary

Let X be any solution of $(I - \widehat{P})X = \mathbb{1}\mathbb{1}^{\mathrm{T}} - \mathrm{Diag}(\widehat{\pi})^{-1}$.
Then $\widetilde{T} = (\widetilde{\tau}_{i \to j})$ is given by

$$\widetilde{T} = LX\,\mathrm{Diag}(a)R - \mathbb{1}b^T,$$

where b is chosen so that $\widetilde{T}_{ii} = 0$.

Corollary

Let $\widetilde{\rho}_S$ be the average 2nd-order average return time to $S \subset V$ of $\{X_t\}$. Then

$$\widetilde{\rho}_S = 1/(\textstyle\sum_{i \in S} \pi_i),$$

where $\pi = (\pi_1, \ldots, \pi_n)^{\mathrm{T}}$ is the stationary density of the pullback matrix P.

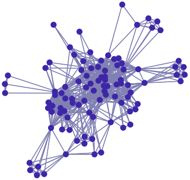
Proof. Let $\mathcal{I} = \cup_{i \in S} \{(j,i) \in E\}$. By Kac's lemma,

$$\begin{aligned} \widetilde{\rho}_S = \widehat{\rho}_{\mathcal{I}} &= \Big(\textstyle\sum_{(j,i) \in \mathcal{I}} \widehat{\pi}_{(j,i)} \Big)^{-1} \\ &= \Big(\textstyle\sum_{i \in S} (\widehat{\pi}^T R)_i \Big)^{-1} = (\textstyle\sum_{i \in S} \pi_i)^{-1}. \end{aligned}$$

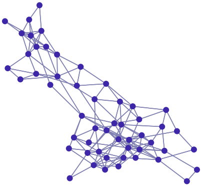
Examples

network	nodes	edges	diam.
`Guppy`	98	725	5
`Dolphins`	53	150	7
`Householder93`	73	180	5

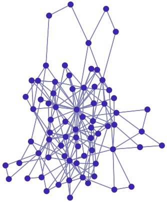
`Guppy`



`Dolphins`



`Householder93`



Examples

network	nodes	edges	diam.
Guppy	98	725	5
Dolphins	53	150	7
Householder93	73	180	5

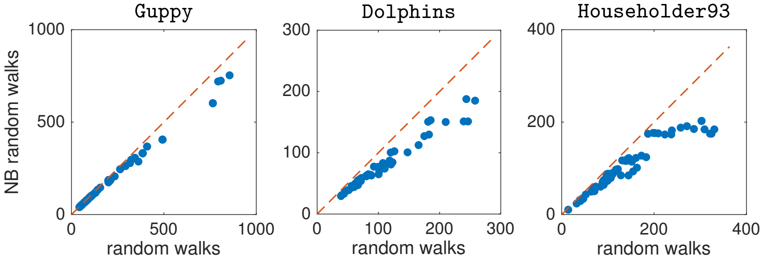


Figure: The mean hitting time $c_i = (\sum_{j=1}^{n} \tau_{i \to j})/n$ computed from 1-order (x-axis) and non-backtracking (y-axis) random walks. Red dotted line: the $y = x$ line.

Examples

Experiments with a second-order rw depending on $\alpha \in [0,1]$ that interpolates between classical ($\alpha = 1$) and non-backtracking ($\alpha = 0$) random walks.

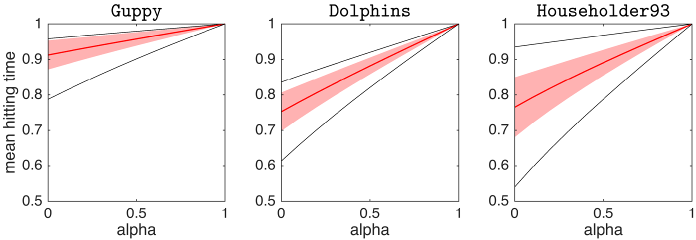


Figure: Mean hitting times $c_i = (\sum_{j=1}^{n} \tau_{i \to j})/n$ normalized to the $\alpha = 1$ case. Solid lines: maximum, average, and minimum values as $\alpha \in [0,1]$.

Kemeny's (almost) constant

Corollary

Let $\mathcal{G} = (V, E)$ be such that, for every pair of edges $(j, i), (k, i) \in E$ there exists a graph automorphism $\varphi : V \mapsto V$ such that $\varphi(j) = k$ and $\varphi(k) = j$. Then

$$\sum_{j=1}^{n} \pi_j \widetilde{\tau}_{i \to j} = \kappa'$$

for some constant $\kappa' < \kappa$, where π is the stationary density of the pullback and κ the Kemeny's constant of the rw on $\mathcal{G}$.

Examples

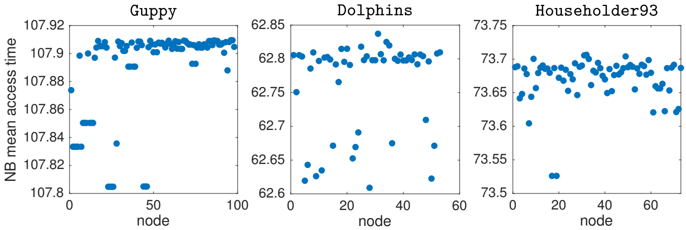


Figure: The non-backtracking mean access time $m_i = \sum_{j=1}^{n} \pi_j \tau_{i \to j}$.

In classical (first-order) rw.s we have $m_i = \kappa$, Kemeny's constant.

network	Guppy	Dolphins	Householder93
Kemeny	119.03	84.524	97.697

Thank you for your attention.