Fractional diffusion and random walks on graphs

Igor Simunec
Relatore: Michele Benzi

30 aprile 2019

Outline

Outline

Notations

We consider an **undirected graph** $\mathcal{G} = (V, E)$ with n vertices.

The **adjacency matrix** associated to $\mathcal{G}$ is the $n \times n$ matrix A such that

$$A_{ij} = \begin{cases} 1 & \text{if } (i,j) \in E, \\ 0 & \text{otherwise.} \end{cases}$$

The graph is undirected, so A is symmetric.
Define the following:

- $\mathbb{1} = [1 \ \dots \ 1]^T \in \mathbb{R}^n$;
- $d = A\mathbb{1}$ the vector of **degrees**, and $D = \mathrm{Diag}(d) \in \mathbb{R}^{n \times n}$;
- The **stochastic matrix** $P = D^{-1}A$, i.e. such that $P \geq 0$ and $P\mathbb{1} = \mathbb{1}$.

The random walk

The matrix P defines a **random walk** on the graph: if X_k denotes the position of the walker at time k, we have

$$P_{ij} = \mathbf{P}(X_{k+1} = j \,|\, X_k = i) = \mathbf{P}(X_1 = j \,|\, X_0 = i).$$

This is a homogeneous discrete time **Markov chain**.
The probability distribution of X_k is given by

$$x_k^T = x_0^T P^k, \qquad \forall k \in \mathbb{N}.$$

We denote by π a **stationary probability distribution**, i.e. such that

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

For an undirected graph, $A = A^T$ and it is easy to see that $\pi = d/(\mathbb{1}^T d)$ is a stationary distribution.

In order to state the Perron-Frobenius theorem, we need the following definitions:

Definition (Irreducible matrix)

A matrix A is **irreducible** if the associated graph is strongly connected: for every pair of nodes (i, j) there is a path going from i to j and viceversa.

Definition (Primitive matrix)

A non-negative matrix A is **primitive** if there exists $m > 0$ such that $A^m > 0$ (component-wise).

Remark

We have $(A^k)_{ij} > 0 \iff$ there is a path of length k from i to j.
So A primitive means that there exists $m > 0$ such that for any pair of nodes (i, j), there exists a path of length m that goes from i to j.

Perron - Frobenius theorem

Theorem (Perron - Frobenius [4])

Let A be an irreducible $n \times n$ non-negative matrix. Then the following statements hold:

(i) The spectral radius $\rho(A)$ is an eigenvalue of A with algebraic multiplicity one.

(ii) The eigenvector x associated to $\rho(A)$ can be chosen so that $x > 0$ (component-wise), and $\rho(A)$ is the only eigenvalue with this property.

(iii) If A is primitive, all other eigenvalues λ of A satisfy $|\lambda| < \rho(A)$.

(iv) If A is primitive and $\rho(A) = 1$, it holds $\lim_{k\to\infty} A^k = \frac{1}{y^T x} xy^T$, where x and y are respectively the right and left eigenvector associated to $\rho(A)$.

Consequences

- The stationary distribution satisfies $\pi^T = \pi^T P$, i.e. it is a left eigenvector associated to 1.
- By the Perron-Frobenius theorem, a strongly connected graph always has a unique stationary probability distribution.
- For an undirected graph, this is given explicitly by $\pi = d/(\mathbb{1}^T d)$.
- If the adjacency matrix A graph is primitive, the stationary distribution can be computed as

$$\pi^T = \lim_{k \to \infty} x_0^T P^k$$

  for any initial probability vector x_0.

From now on we will assume that the graph is undirected and A is primitive.

Outline

Definition (First hitting probability)

Given a pair of nodes (i, j) and an integer $k > 0$, define the **first hitting probability**

$$F_{ij}(k) = \mathsf{P}(X_k = j, X_{k-1} \neq j, \ldots, X_1 \neq j \mid X_0 = i).$$

This is the probability of going from i to j in exactly k steps. We define $F_{ij}(0) = \delta_{ij}$.

Definition (Mean first hitting time)

Given a pair of nodes (i, j), the **mean first hitting time** is defined as

$$T_{ij} = \sum_{k=1}^{\infty} k F_{ij}(k).$$

Note that with this definition we have $T_{ii} = 0$.
The **mean first return times** are instead given by

$$\tau_i = 1 + \sum_{\alpha=1}^{n} P_{i\alpha} T_{\alpha i} = 1 + (PT)_{ii}.$$

With a short computation one can easily prove that

$$T_{ij} = 1 + \sum_{\alpha=1}^{n} P_{i\alpha} T_{\alpha j} - \delta_{ij}\tau_j = 1 + (PT)_{ij} - \delta_{ij}\tau_j.$$

This can be written in the equivalent matrix form

$$(I - P)T = \mathbb{1}\mathbb{1}^T - \mathrm{Diag}(\tau_1, \ldots, \tau_n).$$

Consider the equation we obtained:

$$(I - P)T = \mathbb{1}\mathbb{1}^T - \operatorname{Diag}(\tau_1, \dots, \tau_n).$$

- By multiplying it on the left by π^T, we get

$$0 = \mathbb{1}^T - \pi^T \operatorname{Diag}(\tau_1, \dots, \tau_n) \quad \Rightarrow \quad \tau_i = 1/\pi_i.$$

- By multiplying it on the right by π, we get

$$(I - P)T\pi = 0 \quad \Rightarrow \quad T\pi \in \ker(I - P).$$

From the Perron-Frobenius theorem, $\ker(I - P) = \operatorname{span}(\mathbb{1})$, so there exists a constant K such that $T\pi = K\mathbb{1}$, or

$$\sum_{j=1}^{n} T_{ij}\pi_j = K, \qquad \forall i = 1, \dots, n.$$

The constant K is known as **Kemeny's constant** for the graph $\mathcal{G}$.

We can get an explicit expression for T in terms of the **graph Laplacian matrix** $L = D - A$.

Define the (symmetric) normalized graph Laplacian

$$\mathcal{L} = D^{-1/2} L D^{-1/2} = I - D^{1/2} P D^{-1/2}.$$

We have

$$T = (\mathbb{1}^T d)\left(\mathbb{1} v^T - D^{-1/2}\mathcal{L}^+ D^{-1/2}\right), \qquad v_i = \mathcal{L}^+_{ii}/d_i,$$

where $\mathcal{L}^+$ denotes the Moore-Penrose generalized inverse of $\mathcal{L}$.

A **centrality measure** is a function that associates a value to each node: this provides a ranking of the nodes based on their importance, in some sense.

The above expression for T can be used to define a **random walk centrality** for the graph $\mathcal{G}$.

From the previous expression for T, we can see that the skew-symmetric part of T has rank 2:

$$T - T^T = (\mathbb{1}^T d)(\mathbb{1} v^T - v \mathbb{1}^T) = \mathbb{1} k^T - k \mathbb{1}^T, \qquad k_i = \mathcal{L}^+_{ii} / \pi_i.$$

Component-wise, this reads $T_{ij} - T_{ji} = k_j - k_i$, so

$$T_{ij} > T_{ji} \iff \frac{1}{k_i} > \frac{1}{k_j}.$$

Definition

Given any vector k such that $T - T^T = \mathbb{1} k^T - k \mathbb{1}^T$, the vector with entries $1/k_i$ is a **random walk centrality** for the graph $\mathcal{G}$.

- k can be interpreted as a ranking of the nodes based on their **accessibility**: i is "more accessible" than j if and only if $1/k_i > 1/k_j$.

Consider again the equation

$$T = (\mathbb{1}^T d)\left(\mathbb{1} v^T - D^{-1/2}\mathcal{L}^+ D^{-1/2}\right), \qquad v_i = \mathcal{L}^+_{ii}/d_i,$$

By multiplying it on the left by $\pi^T = d^T/(\mathbb{1}^T d)$, we get

$$\pi^T T = (\mathbb{1}^T d) v^T - d^T D^{-1/2}\mathcal{L}^+ D^{-1/2} = (\mathbb{1}^T d) v^T$$

since $d^T D^{-1/2}\mathcal{L}^+ = 0$.
Hence

$$(\pi^T T)_j = (\mathbb{1}^T d) v_j \quad \Rightarrow \quad \sum_{i=1}^n \pi_i T_{ij} = \mathcal{L}^+_{jj}/\pi_j.$$

This is similar to the expression for Kemeny's constant: recall that

$$\sum_{j=1}^n T_{ij}\pi_j = K.$$

If we replace k with $\ell = k + \alpha\mathbb{1}$ for some $\alpha \in \mathbb{R}$, we still have

$$T - T^T = \mathbb{1}\ell^T - \ell\mathbb{1}^T.$$

We have seen that the entries of ℓ and Kemeny's constant K satisfy

$$\ell_j = \sum_{i=1}^{n} \pi_i T_{ij} + \alpha,$$
$$K = \sum_{i=1}^{n} \pi_i T_{ji} \qquad \forall j.$$

If we choose $\alpha = K$, we get $\ell_j = \sum_{i=1}^{n} \pi_i (T_{ij} + T_{ji})$.

This choice produces the natural random walk centrality $1/\ell_j$, which has a direct random walk interpretation.

Outline

The graph Laplacian

The Laplacian matrix $L = D - A$ represents the **discrete Laplace operator** with Neumann boundary conditions on the graph $\mathcal{G}$.
It is used to model **diffusion** on the graph:

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = -Lx(t), \qquad x(0) = x_0.$$

Its normalization $\tilde{\mathcal{L}} = D^{-1}L$ is also used to describe a **continuous time random walk** on the graph:

$$\frac{\mathrm{d}}{\mathrm{d}t}P(t) = -P(t)\tilde{\mathcal{L}}, \qquad P(0) = I.$$

The solution to this differential equation is the matrix exponential

$$P(t) = \exp(-\tilde{\mathcal{L}}t).$$

The fractional Laplacian

In order to model long-range dynamics on the graph, we define fractional powers of the graph Laplacian matrix: L^γ, for $\gamma \in (0,1)$.
In the case we consider, the graph is undirected and the definition can be given in terms of the eigendecomposition of L.
There exists an orthogonal matrix Q such that

$$L = Q \Lambda Q^T, \qquad \Lambda = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n),$$

and we can define

$$L^\gamma = Q \Lambda^\gamma Q^T, \qquad \Lambda^\gamma = \mathrm{Diag}(\lambda_1^\gamma, \ldots, \lambda_n^\gamma).$$

In a more general setting, the definition can be given using Hermite polynomial interpolation or the Jordan canonical form of L.

Fractional dynamics: motivation

- In some applications, the random walker can perform "long-range jumps" and move directly to a node not connected by an edge to the previous one, with probability that is lower the more distant the new node is.
- The fractional Laplacian L^γ is usually a full matrix, with entries that decay when going "far" from the sparsity pattern of L.

Thus fractional dynamics are useful to capture this long-range behaviour. Using the normalized fractional Laplacian $\mathcal{L}^{(\gamma)} = \mathrm{Diag}(L^\gamma)^{-1} L^\gamma$, we define

$$W = I - \mathcal{L}^{(\gamma)}.$$

Then W is a stochastic matrix, and it can be interpreted as the transition matrix of a fractional random walk on the graph $\mathcal{G}$.

Fractional dynamics: summary

- Fractional diffusion:

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = -L^{\gamma}x(t), \qquad x(0) = x_0.$$

- Discrete time fractional random walk:

$$\begin{cases} x_{k+1}^T = x_k^T W \\ x_0^T \mathbb{1} = 1, \quad x_0 \geq 0. \end{cases}$$

- Continuous time fractional random walk:

$$\frac{\mathrm{d}}{\mathrm{d}t}P(t) = -P(t)\mathcal{L}^{(\gamma)}, \qquad P(0) = I.$$

Decay in the fractional Laplacian

To show a theoretical result on the decay properties of the fractional Laplacian, we will use the following approximation theorem:

Theorem (Jackson [5])

Let $f : [a, b] \to \mathbb{R}$ be a function with modulus of continuity ω. Denote by $\mathcal{P}_n$ the set of polynomials of degree $\leq n$. Then it holds

$$E_n(f) := \inf_{p_n \in \mathcal{P}_n} \|f - p_n\|_\infty \leq c\omega(1/n),$$

where $c = \frac{1}{2}(1 + \pi^2/2)(b - a)$ is a constant that only depends on the interval $[a, b]$.

Decay in the fractional Laplacian

Using Jackson's theorem we can prove the following:

Proposition

Let L be the Laplacian matrix of an undirected graph $\mathcal{G}$ and let $\gamma \in (0,1)$. Denote by $d(i,j)$ the length of the shortest path connecting nodes i and j in $\mathcal{G}$. Then the following holds:

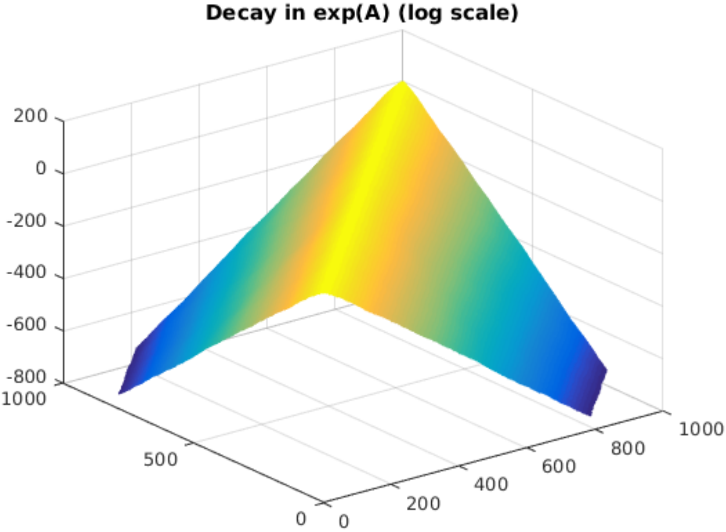
$$|(L^\gamma)_{ij}| \leq C \frac{1}{|d(i,j)-1|^\gamma}, \qquad C = (1+\pi^2/2)\frac{\rho(L)}{2}.$$

Corollary

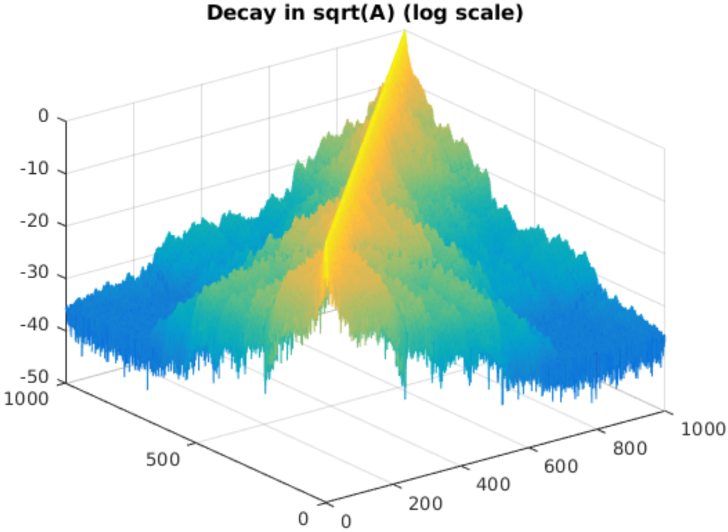
The off-diagonal entries of $W = I - \mathcal{L}^{(\gamma)}$ satisfy:

$$|W_{ij}| \leq (1+\pi^2/2)\frac{\rho(L)^{2-\gamma}}{2\min_i d_i} \cdot \frac{1}{|d(i,j)-1|^\gamma}.$$

- Decay in e^A for A positive semidefinite, bandwidth $k = 5$, and a simple eigenvalue at 0:



- Decay in $A^{1/2}$ for A positive semidefinite, bandwidth $k = 5$, and a simple eigenvalue at 0:



Speed of exploration

- The fractional random walk with transition matrix $W = I - \mathcal{L}^{(\gamma)}$ explores the graph faster than the standard random walk, both for continuous and discrete time.

The differential equation for the continuous time fractional random walk is

$$\frac{\mathrm{d}}{\mathrm{d}t} P(t) = -P(t)\mathcal{L}^{(\gamma)}, \qquad P(0) = I \in \mathbb{R}^{n \times n}.$$

To quantify the "speed of exploration", we define the **average fractional return probability** (for continuous time)

$$p_0^{(\gamma)}(t) = \frac{1}{n} \sum_{i=1}^{n} P(t)_{ii} = \frac{1}{n} \operatorname{tr}\left(\exp(-\mathcal{L}^{(\gamma)} t)\right) = \frac{1}{n} \sum_{i=1}^{n} \exp(-\lambda_i^{(\gamma)} t).$$

The limit for $t \to \infty$ of this probability is $p_0^{(\gamma)}(\infty) = \dfrac{1}{n}$.

The speed of the continuous time exploration is quantified by the **global time**

$$\bar{T}_{\text{cont}} = \int_0^\infty \left(p_0^{(\gamma)}(t) - p_0^{(\gamma)}(\infty) \right) dt = \frac{1}{n} \sum_{i=2}^{n} \frac{1}{\lambda_i^{(\gamma)}},$$

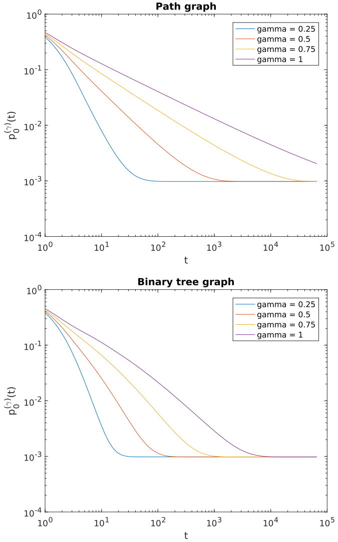
where $0 = \lambda_1^{(\gamma)} < \lambda_2^{(\gamma)} \leq \cdots \leq \lambda_n^{(\gamma)} \leq 2$ are the eigenvalues of $\mathcal{L}^{(\gamma)}$.

We can define an equivalent time for the discrete time random walk, which is related to the fractional fundamental matrix and Kemeny's constant:

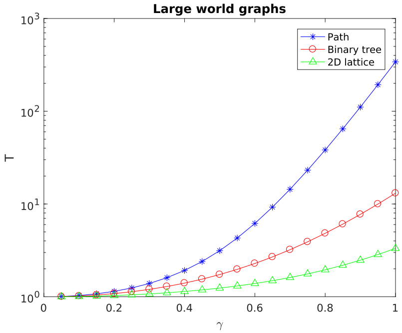
$$\bar{T}_{\text{disc}} = \sum_{k=0}^{\infty} \left(\frac{1}{n} \sum_{i=1}^{n} \left(W^k - \mathbb{1}\pi^T \right)_{ii} \right) = \frac{1}{n} \sum_{i=1}^{n} R_{ii}^{(\gamma)} = \frac{1}{n} K.$$

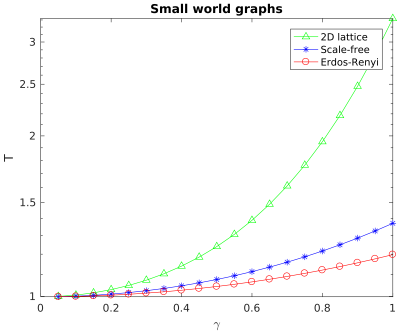
It turns out that $\bar{T}_{\text{cont}}$ and $\bar{T}_{\text{disc}}$ are actually the same.

- Average return probabilities $p_0^{(\gamma)}(t)$ for different graphs and values of γ:



- Average global times $\bar{T}_{\text{cont}} = \bar{T}_{\text{disc}}$ for different graphs and values of γ:





Outline

Conclusions

- We have presented expressions for the matrix of first hitting times T in terms of the normalized Laplacian $\mathcal{L}$.
- We have used those expressions for T to obtain Kemeny's constant and define the random walk centrality.
- We introduced the fractional Laplacian $\mathcal{L}^{(\gamma)}$ in order to model long-range dynamics on the graph.
- We have seen that fractional dynamics explore the graph faster than the standard ones, more significantly for large world graphs.
- The exploration speed is related to the fact that the standard Laplacian is a sparse matrix, while the fractional Laplacian is a full matrix with decay.

References

A. P. Riascos and J. L. Mateos, "Fractional dynamics on networks: Emergence of anomalous diffusion and Lévy flights," *Phys. Rev. E*, vol. 90, p. 032809, Sep 2014.

A. P. Riascos and J. L. Mateos, "Long-range navigation on complex networks using Lévy random walks," *Physical review. E, Statistical, nonlinear, and soft matter physics*, vol. 86, p. 056110, Nov 2012.

D. Fasino, "So, what is the random walk centrality?," Due giorni di Algebra Lineare Numerica, Roma, Feb 2019.

C. Meyer, *Matrix Analysis and Applied Linear Algebra*.
SIAM, 2000.

G. Meinardus, *Approximation of Functions: Theory and Numerical Methods*.
Springer, Berlin, 1967.