Diffusion and Random Walks on Graphs

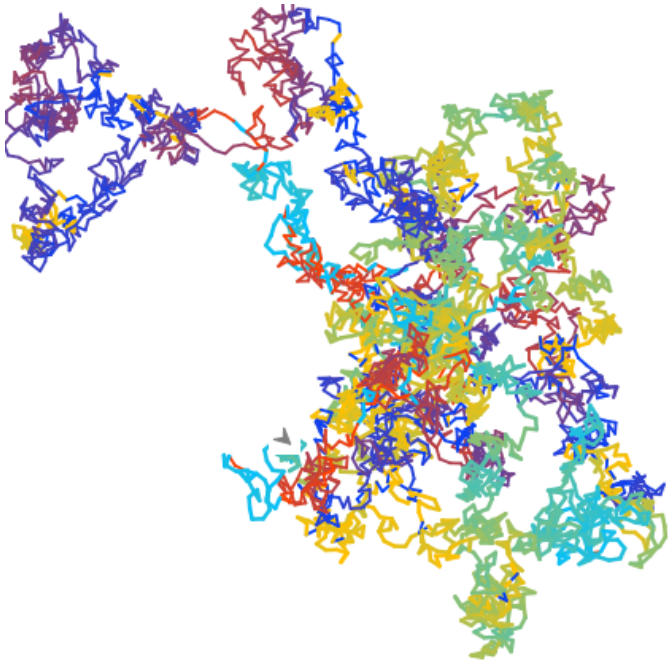
Leonid E. Zhukov

School of Applied Mathematics and Information Science
National Research University Higher School of Economics

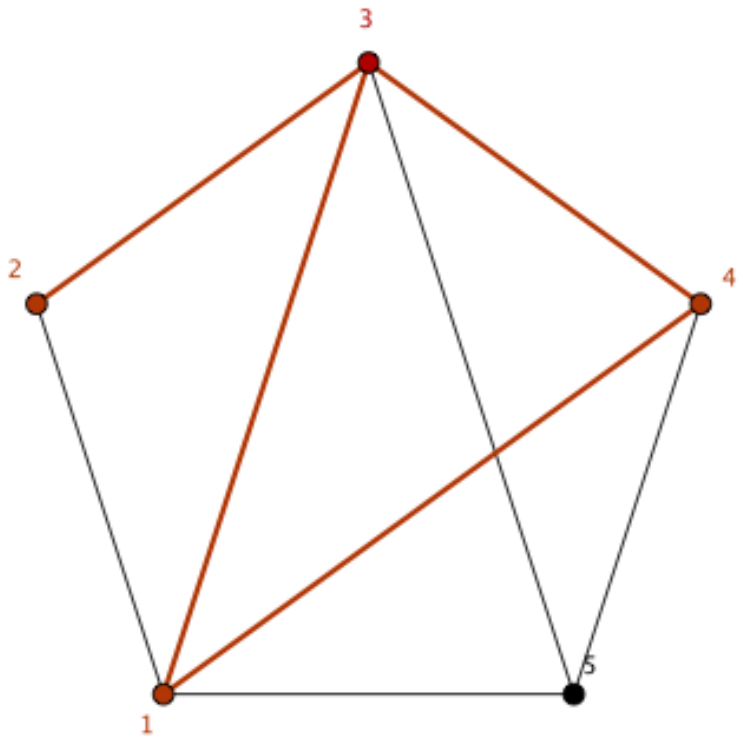
31.03.2014

Random walks on graph

- random walk: start at some vertex v, on every step move to random neihgbour of the current vertex
- if unweighted graph - uniformaly at random
- consider undirected connected graphs
- can go along the same edge more then once
- can visit nodes more then once
- modeling: taking individual location of random walk



Random walks on graph



$$3 \to 4 \to 1 \to 3 \to 2 \to \cdots$$

Random walks on graph

- Let $p_i(t)$ - probability, that a walk is at node i at moment t, $\sum_i p_i(t) = 1$
- $p_j(t+1)$ depends on $p_i(t)$, where $j \in N(i)$
- $\mathbf{p}(0) = p_0$ - initial location, typically on one vertex
- Random walk

$$p_j(t+1) = \sum_i \frac{p_i(t)}{k_i} A_{ij}$$

- Transition (walk) matrix

$$P_{ij} = \frac{A_{ij}}{k_i} = D_{ii}^{-1} A_{ij}, \text{ where } D_{ij} = k_i \delta_{ij}$$

- Matrix form

$$\mathbf{p}(t+1) = (\mathbf{D}^{-1}\mathbf{A})^T \mathbf{p}(t) = \mathbf{P}^T \mathbf{p}(t) = (\mathbf{P}^T)^{t+1} \mathbf{p}(0)$$

$$\mathbf{P}^T = \mathbf{D}^{-1}\mathbf{A}$$

Random walks on graph

- Random walk on connected non-bipartite graphs converges to limiting distribution

$$\lim_{t\to\infty} \mathbf{p}(t) = \lim_{t\to\infty} (\mathbf{P}^T)^t \mathbf{p}(0) = \boldsymbol{\pi}$$

- Limiting distribution = stationary distribution

$$\lim_{t\to\infty} \mathbf{p}(t+1) = \lim_{t\to\infty} \mathbf{P}^T \mathbf{p}(t)$$

$$\boldsymbol{\pi} = (\mathbf{D}^{-1}\mathbf{A})^T \boldsymbol{\pi}$$

$$\pi_j = \sum_i \frac{\pi_i}{k_i} A_{ij}$$

- Stationary (stable) distribution

$$\pi_i = \frac{k_i}{\sum_j k_j} = \frac{k_i}{2E}$$

- Proof:

$$\sum_i \frac{\pi_i}{k_i} A_{ij} = \frac{1}{2E} \sum_i \frac{k_i}{k_i} A_{ij} = \frac{k_j}{2E} = \pi_j$$

Random walks on graph

Eigenvalue problem:

$$\boldsymbol{\pi}\mathbf{P} = \lambda\boldsymbol{\pi}$$

Theorem (Perron-Frobenious)

Real square matrix with non-negative entries that

- *stochastic (rows sum up to one)*
- *irreducible (strongly connected graph)*
- *aperiodic (gcd of the length of the closed directed paths =1)*

has unique largest eigenvalue $\lambda_{max} = 1$ *with positive left eigenvector and power iterations coneverges to it.*

Random walks on graph

- Lazy random walk

$$p_j(t+1) = \frac{1}{2}p_j(t) + \frac{1}{2}\sum_i \frac{p_i(t)}{k_i}A_{ij}$$

- Matrix form

$$\mathbf{p}(t+1) = \frac{1}{2}(\mathbf{I} + \mathbf{D}^{-1}\mathbf{A})^T\mathbf{p}(t)$$

- Stationary distribution (always converges)

$$\pi = (\mathbf{D}^{-1}\mathbf{A})^T\pi$$

Random walks on graph

Theorem

Let λ_2 denote second largest eigenvalue of transition matrix $\mathbf{P} = \mathbf{D}^{-1}\mathbf{A}$, $\mathbf{p(t)}$ probability distribution vector and $\boldsymbol{\pi}$ stationary distribution. If walk starts from the vertex i, $p_i(0) = 1$, then after t steps for every vertex:

$$|p_j(t) - \pi_j| \leq \sqrt{\frac{k_j}{k_i}}\lambda_2^t$$

- For stochastic matrix $\lambda_1 = 1$, $\lambda_2 < 1$
- for $\mathbf{P}' = \frac{1}{2}(\mathbf{I} + \mathbf{D}^{-1}\mathbf{A})$, $\lambda_2' = \frac{1}{2}(1 + \lambda_2)$

Physics of Diffusion

Diffusion is a spontaneus penetration of molecules of one type among molecules of the other type, from regions of higher concentration to the regions of lower concentration. Diffusion happens due to random motion of molecules

- Let $\Phi(r,t)$ -concentration
- Fik's Law

$$J = -C\frac{\partial \Phi}{\partial r} = -C\nabla\Phi$$

- Continuity equation (conserved quantity)

$$\frac{\partial \Phi}{\partial t} + \nabla J = 0$$

- Diffusion Equation (heat equation)

$$\frac{\partial \Phi(r,t)}{\partial t} = C\Delta\Phi(r,t)$$

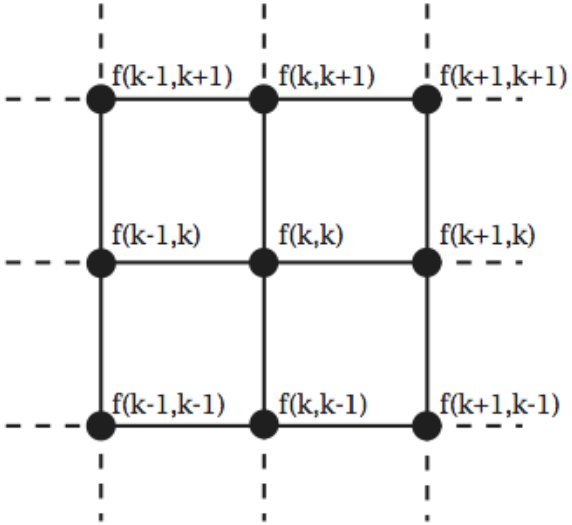
Diffusion

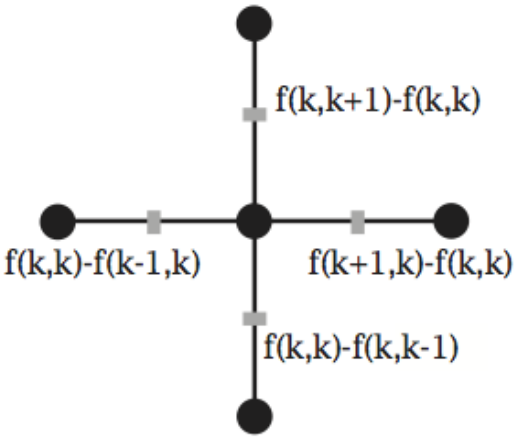
- Laplacian 2D

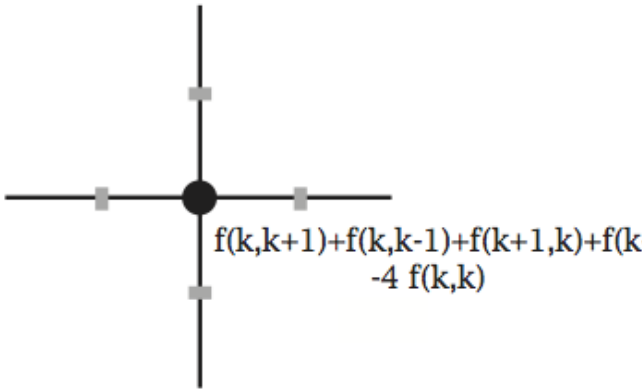
$$\Delta f = \nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

- Discretised Laplacian in 2D

$$\Delta f(x,y) = \frac{f(x+h,y) + f(x-h,y) + f(x,y+h) + f(x,y-h) - 4f(x,y)}{h^2}$$







Diffusion on network

- Some substance that occupy vertices, on each time step diffusises out $\phi_i(t)$ - quantity per node

$$\begin{aligned}\phi_i(t+1) &= \phi_i(t) + \sum_j A_{ij}(\phi_j(t) - \phi_i(t))C\delta t \\ \frac{d\phi_i(t)}{dt} &= C\sum_j A_{ij}(\phi_j(t) - \phi_i(t))\end{aligned}$$

$$\frac{d\phi_i}{dt} = C(\sum_j A_{ij}\phi_j - \sum_j A_{ij}\phi_i) = C(\sum_j A_{ij}\phi_j - k_i\phi_i) = C\sum_j (A_{ij} - \delta_{ij}k_j)\phi_j$$

$$\frac{d\phi_i}{dt} = -C\sum_j L_{ij}\phi_j$$

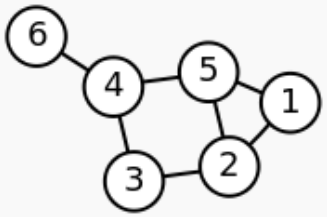
- Graph Laplacian

$$L_{ij} = k_j\delta_{ij} - A_{ij} = D_{ij} - A_{ij}, \quad D_{ij} = k_j\delta_{ij}$$

Graph Laplacian

- Discrete Laplace operator

$$\mathbf{L} = \mathbf{D} - \mathbf{A}$$

Labeled graph	Degree matrix	Adjacency matrix	Laplacian matrix
	$\begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 2 & -1 & 0 & 0 & -1 & 0 \\ -1 & 3 & -1 & 0 & -1 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & -1 & 3 & -1 & -1 \\ -1 & -1 & 0 & -1 & 3 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 \end{pmatrix}$

- Smoothing operator

$$(L\phi)_i = \sum_j (D_{ij} - A_{ij})\phi_j = \sum_j (k_i \delta_{ij}\phi_j - A_{ij}\phi_j) = k_i(\phi_i - \frac{1}{k_i}\sum_j A_{ij}\phi_j)$$

- Laplace equation $\nabla\phi = 0$, $\quad (L\phi)_i = 0$, solution - harmonic function

$$\phi_i = \frac{1}{k_i}\sum_j A_{ij}\phi_j$$

Diffusion on Graph

- Diffusion equation

$$\frac{d\phi}{dt} + C\mathbf{L}\phi = 0$$

- Eigenvector basis

$$\mathbf{L}\mathbf{v}_k = \lambda \mathbf{v}_k$$

$$\phi(t) = \sum_k a_k(t)\mathbf{v}_k, \quad a_k(t) = \phi(t)^T \mathbf{v}_k$$

- ODE

$$\sum_k \left(\frac{da_k(t)}{dt} + C\lambda_k a_k(t) \right) \mathbf{v}_k = 0$$

$$\frac{da_k(t)}{dt} + C\lambda_k a_k(t) = 0$$

$$a_k(t) = a_k(0)e^{-C\lambda_k t}$$

- Solution

$$\phi(t) = \sum_k a_k(0)\mathbf{v}_k e^{-C\lambda_k t}$$

Laplace matrix

- **L** - symmetric positive semidefinite

$$\phi^T L\phi = \sum_{ij} L_{ij}\phi_i\phi_j = \sum_{ij}(k_i\delta_{ij} - A_{ij})\phi_i\phi_j = \frac{1}{2}\sum_{ij} A_{ij}(\phi_i - \phi_j)^2$$

- Spectral properties

$$\mathbf{L}\mathbf{v}_i = \lambda\mathbf{v}_i$$

- real non-negative eigenvalues $\lambda_i \geq 0$ and othogonal eigenvectors $\mathbf{v}_i$
- smallest eigenvalue always $\lambda_1 = 0$ for $\mathbf{v}_1 = \mathbf{e} = [1, 1, 1...1]^T$

$$\mathbf{L}\mathbf{e} = (\mathbf{D} - \mathbf{A})\mathbf{e} = 0$$

- Number of zero eigenvalues = number of connected components
- In connected graph $\lambda_2 \neq 0$ - algebraic connectivity of a graph (spectral gap), $\mathbf{v}_2$ - Fiedler vector

Diffusion on Graph

- Solution

$$\phi(t) = \sum_k a_k(0)\mathbf{v}_k e^{-C\lambda_k t}$$

- all $\lambda_i > 0$ for $i > 1$, $\lambda_1 = 0$:

$$\lim_{t\to\infty} \phi(t) = a_1(0)\mathbf{v_1}$$

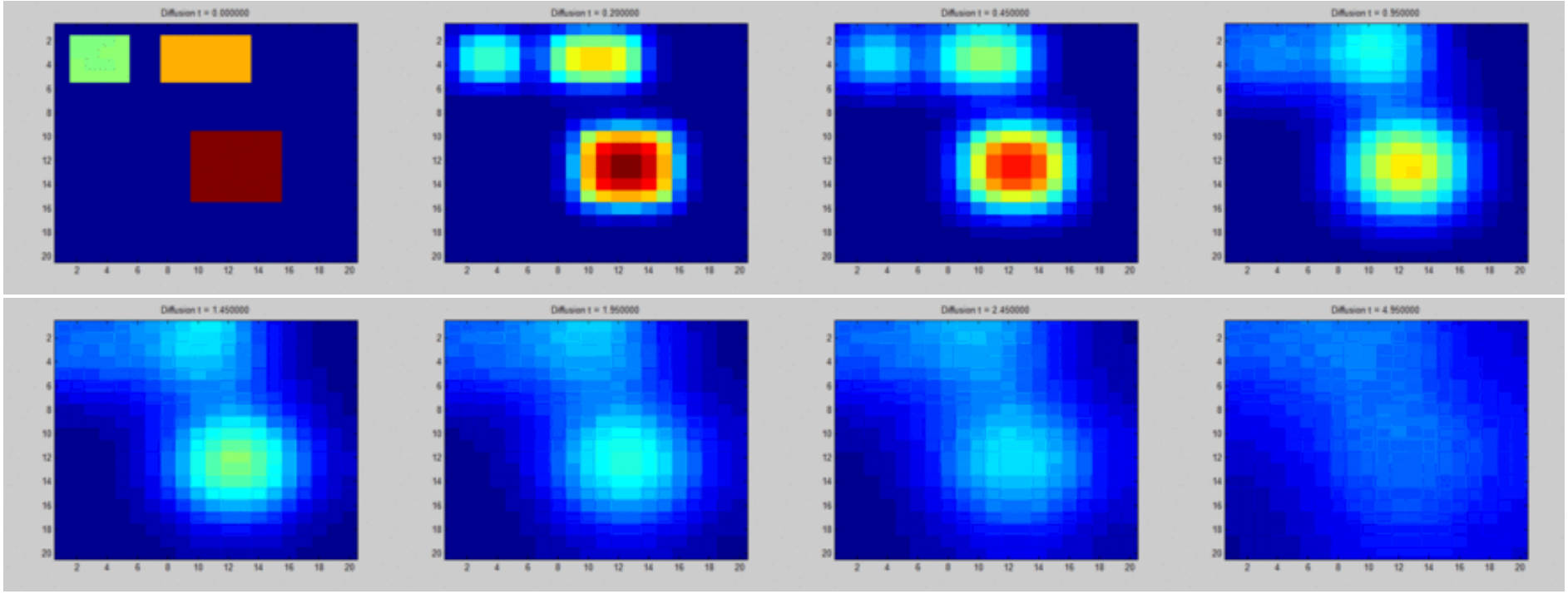
- Normalized solution $\mathbf{v}_1 = \frac{1}{\sqrt{N}}\mathbf{e}$

$$a_1(0) = \phi(0)^T\mathbf{v}_1 = \frac{1}{\sqrt{N}}\sum_j \phi_j(0)$$

- Steady state

$$\lim_{t\to\infty} \phi(t) = (\frac{1}{N}\sum_j \phi_j(0))\mathbf{e} = const$$

Diffusion on Graph



References

- Lovasz, L. (1993). Random walks on graphs: a survey. In Combinatorics, Paul Erdos is eighty (pp. 353 – 397). Budapest: Janos Bolyai Math. Soc.
- Chung, Fan R.K. (1997). Spectral graph theory (2ed.). Providence, RI: American Math. Soc.