

Dynamics on Networks

29. November 2021

Newman treats processes on networks in chapters 16-19, the dynamical systems we will look at are chapter 19. Random walks are in 6.14, with an easy review of the core results. Besides this the main reference for this part of the course is [2] which is also available at <https://arxiv.org/abs/1612.03281> and gives an excellent if terse overview of random walks on networks.

Important: In the stochastic system literature transition matrices are defined differently from these notes. We have T_{ij} denoting the transition from j to i , the random walk literature usually has it the other way around.

1 Discrete time random walk

A walk can be written as a sequence of nodes $s(t) = [1, 2, 3, 2, 4, \dots]$ such that $(s(t), s(t+1))$ is in the edge set for all t . A stochastic process on the network is then a probability density on the space of walks $p(s)$, it assigns a probability to a particular walk occurring. We are interested in time invariant Markovian random walks, that is, random walks for which the probability of the next time only depends on where we are, not where we've been or how late it is. This implies that we can propagate probability densities in time: $p(s(t+1) = i) = \sum_j p(s(t+1) = i | p(s(t)) = j) p(s(t)) = j$. The conditional probability to be at i given that we were at j is called the transition matrix $T_{i \leftarrow j} = p(s(t+1) = i | p(s(t)) = j)$.

Take a network with adjacency matrix A_{ij} , and degrees $d_i = \sum_j A_{ij}$ and $D_{ij} = \delta_{ij} d_i$. a random walker that at each time step selects a random edge to transverse has transition matrix given by

$$T_{i \leftarrow j} = \frac{A_{ij}}{d_j} = A \cdot D^{-1} \quad (1)$$

The stationary distribution of this random walker is defined by:

$$p_i^* = \sum_j T_{i \leftarrow j} p_j^* = \sum_j \frac{A_{ij} p_j^*}{d_j} \quad (2)$$

Ansatz:

$$p_i^* = c d_i \quad (3)$$

$$c d_i = \sum_j \frac{A_{ij} c d_j}{d_j} \quad (4)$$

$$c d_i = c \sum_j A_{ij} \quad (5)$$

To understand the dynamics we need the spectrum and basis of $A \cdot D^{-1}$. First an observation:

$$(1, 1, 1, 1, \dots)^\top \cdot A = (d_1, d_2, d_3, d_4, \dots)^\top \quad (6)$$

$$(1, 1, 1, 1, \dots)^\top \cdot A \cdot D^{-1} = (1, 1, 1, 1, \dots)^\top \quad (7)$$

So we have an eigenvalue 1 with right eigenvector p^* and left eigenvector $(1, 1, 1, 1, \dots)^\top$.

If the graph is strongly connected the Perron Frobenius theorem gives us a lot of information about A , in particular we know that this is a simple eigenvalue and all others have equal or smaller modulus.

Supplemental reading: Perron–Frobenius theorem for irreducible non-negative matrices. Wikipedia article is a very good summary

Note that:

$$\tilde{A} = D^{-1/2} \cdot A \cdot D^{-1/2} \quad (8)$$

$$T = D^{1/2} \cdot \tilde{A} \cdot D^{-1/2} \quad (9)$$

Thus T and \tilde{A} are similar, and have the same spectrum (and easily related eigenvectors), the eigenvectors we already found are transformed into:

$$\lambda^{(1)} = 1 \quad (10)$$

$$v^1 = (d_1^{0.5}, d_2^{0.5}, d_3^{0.5}, d_4^{0.5}, \dots)^\top \quad (11)$$

The matrix \tilde{A} is symmetric and thus has real eigenvalues which we order $\lambda^{(i)} > \lambda^{(i+1)}$.

The slowest process to die off is that given by $\lambda^{(2)}$, the Cheeger constant / conductance tells us something here:

Consider a graph partitions S, \bar{S} with $Vol(S) < Vol(\bar{S})$ where

$$Vol(S) = \sum_{i \in S} d_i \quad (12)$$

is the sum of d_i in S , call $b(S)$ the number of links from S to \bar{S} , then

$$h = \min_S \frac{b(S)}{|S|} \quad (13)$$

then we have:

$$\frac{h^2}{2} < 1 - |\lambda^{(2)}| < 2h \quad (14)$$

Proof is omitted here (but I might give a partial one next time...). This material is discussed in [2] 3.2.3, which gives many nice references.

Supplemental reading: We can design dynamical processes to understand properties of the graph. The most famous one is PageRank, section 5.2.1 in [2].

2 Aside: Decomposition for non-symmetric matrices

Transition matrices are not typically symmetric. $T_{ij} \neq T_{ji}$. However, they typically still can be diagonalized. We do need to distinguish left and right eigenvectors though.

$$T \cdot v_R^{(i)} = \lambda^{(i)} v_R^{(i)} \quad (15)$$

$$v_L^{(i)\top} \cdot T = \lambda^{(i)} v_L^{(i)\top} \quad (16)$$

For non-symmetric (not self-adjoint) transition matrices, eigenvectors are no longer orthogonal, so $v_R^{(i)} \cdot v_R^{(j)} \neq \delta_{ij}$. However, instead we can choose a normalization such that we have:

$$v_R^{(i)} \cdot v_L^{(j)} = \delta_{ij} \quad (17)$$

We can decompose the transition matrix in the following way.

$$T = \sum_i \lambda^{(i)} v_R^{(i)} v_L^{(i)\top} \quad (18)$$

$$T \cdot v_R^{(j)} = \sum_i \lambda^{(i)} v_R^{(i)} (v_L^{(i)} \cdot v_R^{(j)}) \quad (19)$$

$$= \sum_i \lambda^{(i)} v_R^{(i)} \delta_{ij} = \lambda^{(j)} v_R^{(j)} \quad (20)$$

$$(21)$$

or with indices:

$$T_{kl} = \sum_i \lambda^{(i)} v_{Rk}^{(i)} v_{Li}^{(i)} \quad (22)$$

$$\sum_l T_{kl} v_{Rl}^{(j)} = \sum_{il} \lambda^{(i)} v_{Rk}^{(i)} v_{Li}^{(i)} v_{Rl}^{(j)} \quad (23)$$

$$= \sum_i \lambda^{(i)} v_{Rk}^{(i)} \delta_{ij} = \lambda^{(j)} v_{Rk}^{(j)} \quad (24)$$

$$(25)$$

Even though the eigenvectors are no longer orthogonal they still typically form a basis, so we can express initial conditions in this basis. The right eigenvectors have the function of telling us what the eigenmodes look like, the left eigenvectors tell us how much initial conditions end up in which eigenmode:

$$p(K) = T^K p(0) \quad (26)$$

$$p(K) = \sum_i \lambda^{(i)K} v_R^{(i)} (v_L^{(i)} \cdot p(0)) \quad (27)$$

$$(28)$$

The overlap of left eigenvector and initial distribution $(v_L^{(i)} \cdot p(0))$ is the weight in the mode decomposition.

3 Linearizing general dynamics on a network

We need to say what we mean by a dynamical system on a network, one possibility is this:

$$\dot{x}_i = f(x_i) + \sum_j A_{ij} g(x_i - x_j) \quad (29)$$

Assume we have a fix point x^* , then we can look at the linear behaviour of the system:

$$x_i(t) = x_i^* + \epsilon z_i(t) \quad (30)$$

$$x_{ij}^* = x_i^* - x_j^* \quad (31)$$

$$\dot{x}_i^* = 0 = f(x_i^*) + \sum_j A_{ij} g(x_{ij}^*) \quad (32)$$

$$\epsilon \dot{z}_i = f(x_i^*) + \epsilon f'(x_i^*) z_i + \sum_j A_{ij} g(x_{ij}^*) + \epsilon \sum_j A_{ij} g'(x_{ij}^*) (z_i - z_j) + O(\epsilon^2) \quad (33)$$

$$\dot{z}_i = f'(x_i^*) z_i + \sum_j A_{ij} g'(x_{ij}^*) (z_i - z_j) \quad (34)$$

$$(35)$$

Consider the weighted adjacency matrix $A_{ij}^w = A_{ij} g'(x_{ij}^*)$ and the corresponding degree matrix D^w , and introduce $F_{ij} = \delta_{ij} f'(x_i^*)$. We then have

$$\dot{z}_i = f'(x_i^*) z_i + \sum_j A_{ij}^w (z_i - z_j) \quad (36)$$

$$\dot{z}_i = f'(x_i^*) z_i + z_i \sum_j A_{ij}^w - \sum_j A_{ij}^w z_j \quad (37)$$

$$\dot{z}_i = f'(x_i^*) z_i + z_i d_i^w - \sum_j A_{ij}^w z_j \quad (38)$$

$$\dot{z} = F \cdot z + [D^w - A^w] \cdot z \quad (39)$$

4 The Network Laplacian and Diffusion

For a network of adjacency matrix A the matrix

$$L = D - A \quad (40)$$

is the network Laplacian. As we will see below it controls diffusion processes on networks. In the case of a linearized network dynamics, the strength of diffusion across edges is determined by the linear response of the non-linear dynamics. We will return to this when we will analyze the stability conditions of fixed points in more detail. First we will study the network Laplacian more thoroughly.

Diffusion is closely related to flows on the network. To efficiently speak of flows we need to provide a fiducial orientation on the edges. Then we can have a positive flow value representing a flow in one direction, and a negative flow value denoting a flow in the opposite direction.

Consider a network with node set \mathcal{N} and edge set \mathcal{E} , pick two functions that determine for each edge which end is its source and target: $s, t : \mathcal{E} \rightarrow \mathcal{N}$. We have $s(k) \neq t(k)$ and denoting an edge connecting the nodes i and j by the set $\{i, j\}$ we have $k = \{s(k), t(k)\} \in \mathcal{E}$. This information can be encoded in the signed incidence matrix $B_{ik} \in \{0, 1, -1\}^{\mathcal{N} \times \mathcal{E}}$:

$$B_{ik} = \begin{cases} +1 & \text{if } i = s(k) \\ -1 & \text{if } i = t(k) \\ 0 & \text{otherwise} \end{cases} \quad (41)$$

With this we can obtain the signed flows:

$$f_k = [B^\top \cdot z]_k = \sum_i B_{ik} z_i = s(k) - t(k) \quad (42)$$

on the other hand, given a set of flows on the *edges*

$$[B \cdot f]_i = \sum_k B_{ik} f_k = \sum_{k:s(k)=i} f_k - \sum_{k:t(k)=i} f_k \quad (43)$$

is the total flow going out of node i .

So given some potential z , which induces a flow on the edges of $z_i - z_j$, $B^\top \cdot z$ is the state of flows induced by the potential on the edges, and $B \cdot B^\top \cdot z$ is the total flow out of the nodes induced by the potential. If the potential z is what is flowing, we have $\dot{z} = -B \cdot B^\top \cdot z$. This is a diffusion process.

Let us study this:

$$[B \cdot B^\top]_{ij} = \sum_k B_{ik} B_{jk} \quad (44)$$

$$\text{for } i \neq j = \sum_k B_{ik} B_{jk} = \begin{cases} -1 & \text{if } k = \{i, j\} \\ 0 & \text{otherwise} \end{cases} \quad (45)$$

$$\text{for } i = j = \sum_k B_{ik} B_{ik} = d_i \quad (46)$$

so we have:

$$B \cdot B^\top = L \quad (47)$$

Further reading: These ideas are related to simplicial homology, https://en.wikipedia.org/wiki/Simplicial_homology. In this context B_{ik} is the boundary operator. A cycle flow is in the kernel of B_{ik} , so the linear algebra of B counts cycles on graphs. Simplicial homology is concerned with higher dimensional analogies of this observation.

4.1 A continuum limit consideration

Let's first consider a discretized line, with spacing ϵ , and indexed by n . On it we have a discretized field z_n . The discrete derivative is

$$[\partial z]_n = \frac{z_{n+1} - z_n}{\epsilon} \quad (48)$$

the second derivative, which in one dimension is the Laplacian is

$$[\partial^2 z]_{n-1} = \frac{z_{n+1} - 2z_n + z_{n-1}}{\epsilon^2} \quad (49)$$

In the continuum limit the shift by $n - 1$ on the left does not matter, so an equally permissible discretization is

$$[\partial^2 z]_n = \frac{z_{n+1} - 2z_n + z_{n-1}}{\epsilon^2} \quad (50)$$

In terms of the adjacency matrix of the line, this can be written as

$$[\partial^2 z]_n = \frac{1}{\epsilon^2} \sum_m A_{nm} z_m - d_n z_n = -\frac{1}{\epsilon^2} [L \cdot z]_n \quad (51)$$

So we obtain the negative network Laplacian.

In a higher dimensional regular grid, the Laplacian is the sum over the second derivatives in all possible directions. Take n to be a vector of whole numbers. Take $n +_k a$ to be the vector n with a added onto the k th entry, then we again have:

$$[\Delta z]_n = \frac{1}{\epsilon^2} \sum_k z_{n+k-1} - 2z_n + z_{n+k-1} \quad (52)$$

$$= \frac{1}{\epsilon^2} \sum_m A_{nm} z_m - d_n z_n = -\frac{1}{\epsilon^2} [L \cdot z]_n \quad (53)$$

as the positive parts of the sum run over all neighbors of n , the negative parts sum up to twice the dimension of the grid, which is the degree.

4.2 What diffusion says about networks

Diffusion can tell us things about the network topology, and vice versa.

Consider the question: What is the dimension of a network?

This is a question that, for example, people who try to build space times from scratch ask [1], but the concept of spectral dimension goes far back (I haven't really found a good account of its origin so far yet but I'll keep looking). The core observation is that the property of diffusion depends on the dimension of the space we are diffusing in. Particularly the probability to return to the starting location. Consider the heat kernel solution to diffusion from a delta peak at the origin:

$$K_{\text{heat}}(t, x) = (4\pi t)^{-d/2} e^{-\frac{x^2}{4t}} \quad (54)$$

The density at the origin at time t goes down as $t^{-d/2}$, if we interpret this density as the density of a continuous time random walker, this represents the probability density to revisit the origin after time t . If we can measure this dependence we can fit a (possibly fractal) notion of dimension here.

This translates immediately into the network context. We can observe how $z_i(t)$ changes with time under the diffusion evolution $\dot{z} = -L \cdot z$, and if we observe a fixed polynomial behavior we can interpret this as the dimension of the network at a certain scale.

There are other ways to characterize the spectral dimension directly in terms of the spectrum of L . Just as with the discrete time random walker we will now continue to investigate the spectrum of L to understand better how dynamics given in terms of L and network properties relate.

4.3 Spectrum of L

First observe that L is a semi-positive matrix:

$$z^\top \cdot L \cdot z = z^\top \cdot B \cdot B^\top \cdot z \quad (55)$$

$$= |B^\top \cdot z|^2 \geq 0 \quad (56)$$

$$(57)$$

This means that diffusion is stable. Further it is clear that if all nodes are the same, there is nothing flowing on the network, and thus the total flow at every node is zero.

$$z^* = (1, 1, 1, \dots)^\top \quad (58)$$

$$L \cdot z^* = 0 \quad (59)$$

$$[L \cdot z^*]_i = \sum_j A_{ij}(1 - 1) = 0 \quad (60)$$

Thus we have an equilibrium state of diffusion in which every node has the same potential.

Further, if the graph is connected, then this is the only equilibrium state. This follows by considering the flow vector for an equilibrium state:

$$0 = z^{*\top} \cdot L \cdot z^* \quad (61)$$

$$= z^{*\top} \cdot B \cdot B^\top \cdot z^* \quad (62)$$

$$= |B^\top \cdot z|^2 \quad (63)$$

$$= |f^*|^2 = \sum_{k \in \mathcal{E}} |f_k^*|^2 \quad (64)$$

However, this sum of positive terms can only be zero if all individual terms are zero. This mean $f_k = 0 = z_{s(k)} - z_{t(k)}$, thus within the component every z has to be the same.

The number of zeros of the Laplacian spectrum counts the number of components of the graph. We can choose a basis of the 0 eigenspace such that each eigenvector is constant on one component and zero on all others. All other eigenvalues are strictly positive.

In fact the spectrum of L is intimately related with the graph structure. Spectral graph theory is an entire field of mathematics in its own. It is (maybe) surprisingly rich and still actively developing. An introductory article can be found here [3], where chapters 8-12 cover what we are looking at here.

Presenting the results of spectral graph theory would be a course onto itself, so we will focus on those most relevant to dynamics. Just as some very simple flavor though: $Tr(A^r)$ counts the number of closed walks of length r , this can be written as $\sum_i a(i)^r$.

Lets next consider the maximum eigenvalue. This is the fastest process on the graph. We could expect that this has something to do with highly localized diffusion across many edges. We can get a very simple result here, take i to be the node for which $v_i^{(n)}$ is maximal, then:

$$v_i^{(n)} \leq v_j^{(n)} \quad \forall j \neq i \quad (65)$$

$$\lambda^{(n)} v_i^{(n)} = \sum_j L_{ij} v_j^{(n)} \quad (66)$$

$$= d_i v_i^{(n)} - \sum_j A_{ij} v_j^{(n)} \quad (67)$$

$$\leq d_i v_i^{(n)} - \sum_j A_{ij} v_i^{(n)} \quad (68)$$

$$= 2d_i v_i^{(n)} \quad (69)$$

$$\Rightarrow \lambda^{(n)} \leq 2d_i \quad (70)$$

$$\Rightarrow \lambda^{(n)} \leq 2d_{\max} \quad (71)$$

$$(72)$$

Thus the spectrum of L is in $[0, 2d_{\max}]$.

Exercises: Pure diffusion is stable by the spectral results above, however, if there is more dynamics going on on the network, interesting stuff starts to happen at the intersection of diffusion and local dynamics, and it's possible to obtain non-trivial stability results. This will be the subject of the seminar talk January 3rd. Newman contains a number of illuminating exercises on this topic at the end of chapter 18. Exercises 18.1-18.3 take the type of arguments we have here and apply them to stability analysis. 18.4 considers the connection to regular lattices, 18.5 synchronization (in 18.5 the unstated assumption is made that $g(x) = -g(-x)$).

4.4 The slowest process and graph partitioning

As for stochastic processes we will now look at the slowest process, for a connected network this is $\lambda^{(2)}$. This can be characterized by a variational principle:

$$\lambda^{(2)} = \min_{x: x \cdot 1^T = 0} \frac{\langle x, Lx \rangle}{\langle x, x \rangle} \quad (73)$$

This will be extremely useful in what's to come. We will now go through the argument of Newman, chapter 11.5.

Say we want to partition the graph into two clusters, of size n_1 and n_2 , such that the number of links between these clusters is minimized. Newman shows in chapter 11.5. how this can be approximated by the second eigenvector, that is, the vector that realizes the minimization in (73).

Literatur

- [1] Jan Ambjørn, Jerzy Jurkiewicz und Renate Loll. „The spectral dimension of the universe is scale dependent“. In: *Physical review letters* 95.17 (2005), S. 171301.

- [2] Naoki Masuda, Mason A Porter und Renaud Lambiotte. „Random walks and diffusion on networks“. In: *Physics reports* 716 (2017), S. 1–58.
- [3] Bogdan Nica. „A brief introduction to spectral graph theory“. In: *arXiv preprint arXiv:1609.08072* (2016).