Solution 19-3

This is an example solution from the forthcoming book Networks and Complexity.
Find more exercises at https://github.com/NC-Book/NCB

Ex 19.3: Epidemic Centrality [3]

Let's find the most important node for a spreading process.

a) Consider an SIS epidemic on a network described by an adjacency matrix $\mathbf{A}$. Explain why $x_i(t)$, the probability that node i is infected at time t can be modelled by

$$\dot{x}_i = -rx_i + p(1 - x_i) \sum_j A_{ij} x_j \tag{1}$$

Solution

In the SIS model infected nodes recover at rate r, and become infected at the rate p per infected neighbor. The probability that node i is currently infected is x_i. Multiplying this probability that the node is till infected with the recovery rate gives us a loss rate for the probability to be in the infected state, which gives us the $-rx_i$ term.

The probability that node i is currently susceptible is $1 - x_i$. We can estimate the number of infected neighbors of i by the sum $\sum A_{ij} x_j$ for each of these neighbors infection occurs at rate p. Multiplying these factors leads to $p(1 - x_i) \sum A_{ij} x_j$, which is the second term.

b) So far the equation is non-linear. To use spectral methods we need to linearize it. Consider the limit in which only a small proportion of nodes is infected at any time and use it to find a linear approximation to the differential equations. (There are two ways to do this, if you see only one check the solution for the other.)

Solution

The physicists way to do this is to use the assumption of rare infection in the form

$$x_i \ll 1 \tag{2}$$

to justify

$$1 - x_i \approx 1. \tag{3}$$

This linearizes the differential equation to

$$\dot{x}_i = -rx_i + p \sum_j A_{ij} x_j. \tag{4}$$

which is a linear equation.

The more formal mathematical way is to use a Taylor approximation of the equation system around zero. To do this we introduce the vector $\boldsymbol{x} = (x_1, x_2, \ldots)^{\mathrm{T}}$ already at this point.

We know that the linear-order Taylor approximation is

$$\dot{\boldsymbol{x}} = f(\boldsymbol{x}) \approx f(\boldsymbol{0}) + \mathbf{J}\boldsymbol{x} \tag{5}$$

In our system the ith element of f is

$$f_i(\boldsymbol{x}) = -rx_i + p(1 - x_i)\sum_j A_{ij}x_j. \tag{6}$$

So at $x_i = x_j = 0$ we have $f_i(\boldsymbol{x}) = 0$, and because this true for all all elements

$$f(\boldsymbol{0}) = \boldsymbol{0} \tag{7}$$

Hence we are left with

$$\boldsymbol{x} = \mathbf{J}\boldsymbol{x} \tag{8}$$

where $\mathbf{J}$ is the Jacobian in $\boldsymbol{x} = \boldsymbol{0}$. Technically this is a linear system, so we are done already. But is this the same result that we got using the first way? Indeed it is, as we will see in the next sub-question.

c) Show that the system can now be written in the matrix form $\dot{\boldsymbol{x}} = \mathbf{J}\boldsymbol{x}$. Find a way to express the matrix $\mathbf{J}$ that appears as a function of $\mathbf{A}$.

Solution
If we used the physical route in the previous subquestion we have

$$\dot{x}_i = -rx_i + p\sum_j A_{ij}x_j. \tag{9}$$

So we now have to transform this to matrix form. We know that $\sum_j A_{ij}x_j$ is the formula for the ith element of $\mathbf{A}\boldsymbol{x}$ so we can write

$$\dot{x} = -r\boldsymbol{x} + p\mathbf{A}\boldsymbol{x} \tag{10}$$

We are already close to the desired result. All that is left to do now is to write the transformations on the right hand side as one matrix. So we want to isolate the factor $\boldsymbol{x}$ from both terms but $-r + p\mathbf{A}$ would not be a meaningful mathematical expression as we don't know how to add $-r$ to a matrix. Instead we can just multiply the identity matrix,

$$\begin{aligned} \dot{x} &= -r\boldsymbol{x} + p\mathbf{A}\boldsymbol{x} & (11) \\ &= -r\mathbf{I}\boldsymbol{x} + p\mathbf{A}\boldsymbol{x} & (12) \\ &= (-r\mathbf{I} + p\mathbf{A})\boldsymbol{x} & (13) \end{aligned}$$

We can now write the system as

$$\dot{\boldsymbol{x}} = \mathbf{J}\boldsymbol{x} \tag{14}$$

where

$$\mathbf{J} = p\mathbf{A} - r\mathbf{I} \tag{15}$$

If we went the mathematical route we already know

$$\dot{\boldsymbol{x}} = \mathbf{J}\boldsymbol{x} \tag{16}$$

2

but we still have to find out what $\mathbf{J}$ is. For this purpose we use the definition

$$J_{ij} = \left.\frac{\partial \dot{x}_i}{\partial x_j}\right|_{\boldsymbol{x}=\boldsymbol{0}}. \tag{17}$$

To substitute the present system we change the index of the summation to k to avoid having to deal with two different j in the same equation, which means our system is now

$$\dot{x}_i = -rx_i + p(1-x_i)\sum A_{ik}x_k \tag{18}$$

Substituting into the definition of the Jacobian yields

$$J_{ij} = -r\delta_{ij} + pA_{ij} \tag{19}$$

There is quite a lot to explain here: First, the δ that appears in the first term on the right is the Kronecker delta

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \tag{20}$$

This delta appears because

$$\frac{\partial x_i}{\partial x_j} = \delta_{ij}. \tag{21}$$

Second, the term

$$p(1-x_i)\sum A_{ik}x_k$$

is a product so we should differentiate it using a product rule. However we can make our life a bit easier by multiplying it out

$$p(1-x_i)\sum A_{ik}x_k = \left(p\sum A_{ik}x_k\right) - \left(p\sum A_{ik}x_i x_k\right) \tag{22}$$

We can now see that every term in the second sum on the right contains a product of variables $x_i x_k$ when we differentiate this with respect to x_j this will leave us with terms that all contain x_i. Since we set $x_i = 0$ in the next step all these terms vanish.

This leaves only the derivative of $p\sum A_{ik}x_k$ to worry about. If we differentiate sum all the derivatives are zero except for the term that contains an x_j, i.e. the only term that survives the derivative is the term where $k = j$. Setting $x_i = 0$ after differentiating yields the term pA_{ij}.

So far we arrived at the equation

$$J_{ij} = -r\delta_{ij} + pA_{ij} \tag{23}$$

Note that δ_{ij} also describes the elements of an identity matrix

$$I_{ij} = \delta_{ij} \tag{24}$$

Hence we can translate the equation for the elements of $\mathbf{J}$ into the matrix equation

$$\mathbf{J} = p\mathbf{A} - r\mathbf{I} \tag{25}$$

so we get the same result that we arrived at via the physical route.

If we wanted to we could even translate this result back into a linearized differential equation for the individual components. We know

$$\dot{\boldsymbol{x}} = \mathbf{J}\boldsymbol{x} \tag{26}$$

Hence,

$$\dot{x}_i = \sum J_{ij}x_j \tag{27}$$

$$= \sum (pA_{ij} - rI_{ij})x_j \tag{28}$$

$$= \left(-r\sum I_{ij}x_j\right) + \left(p\sum A_{ij}x_j\right) \tag{29}$$

$$= \left(-r\sum \delta_{ij}x_j\right) + \left(p\sum A_{ij}x_j\right) \tag{30}$$

$$= -rx_i + p\sum A_{ij}x_j \tag{31}$$

which takes us all the way back to the equation we guessed as our first step on the physical way.

In summary the physical approach is a bit faster and easier, but it requires a bit more insight as we have to spot the right way to lineraize the system. The more formal mathematical approach needs a little bit more work, but it can be always applied to linearize systems even if we don't have the right insight, hence it is often what we use for more complicated systems.

d) The elements of the leading eigenvector of **J** now indicate who is most at risk in the epidemic. Explore how the largest eigenvalue λ and the corresponding eigenvector $\boldsymbol{v}$ is linked to the adjacency matrix **A**.

Solution
Using the spectral shift property the we can write the eigenvalues of **J** as

$$\lambda_n = p\kappa_n - r \tag{32}$$

where κ_n is the nth eigenvalue of the adjacency matrix. Moreover,

$$\boldsymbol{v_n} = \boldsymbol{w_n} \tag{33}$$

where $\boldsymbol{w_n}$ is the nth eigenvalue of the adjacency matrix. So to identify the nodes that are most at risk of being infected we only need to compute the leading eigenvector of the adjacency matrix. In other words epidemic centrality is spectral centrality.

e) Bonus questions: In an epidemic the nodes who are most at risk also poses the greatest risk to others. This may be intuitive but how can we see it mathematically?

Solution
We know that the left eigenvectors of the Jacobian indicate sensitivity to perturbations whereas the right eigenvectors indicate impact on the rest of the network. The right eigenvectors of a matrix are are also the left eigenvectors of the transpose of the matrix. However, in the epidemic example the Jacobian is a symmetric matrix, $\mathbf{J} - \mathbf{J}^{\mathrm{T}}$ and hence the left and right eigenvectors are identical: whoever is likely to get infected is also likely to spread the disease to others.