Coursework II: Stochastic Processes and Networks

Max Taylor-Davies
Modelling in Biology
Imperial College London

March 8, 2020

1 Independent Pairs

1.1

1.1.1

We have a network of M nodes, where each node represents an animal. Since we are not considering self-interactions, each node may possibly connect to M-1 other nodes. Each of these possible connections exists with probability q. Thus the expected degree of each node is given by $q(M-1)$. We can choose to model the network as a simple (undirected) Erdos-Renyi network; the degree distribution is then binomial, given by

$$p(k) = P(\text{degree} = k) = \binom{M-1}{k} q^k (1-q)^{M-1-k}.$$

1.1.2

Fig. 1a shows a graphical representation of the evolution of $A_{ij}(n)$ between states 0 and 1. The transition matrix arising from this representation is given by

$$\mathbf{T} = \begin{bmatrix} 1-q & 1-q \\ q & q \end{bmatrix}$$

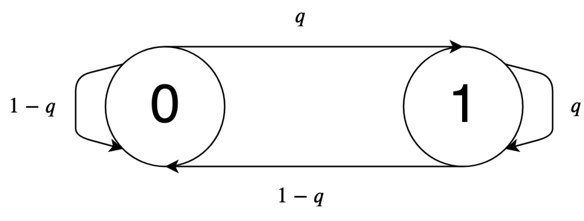
1.2

1.2.1

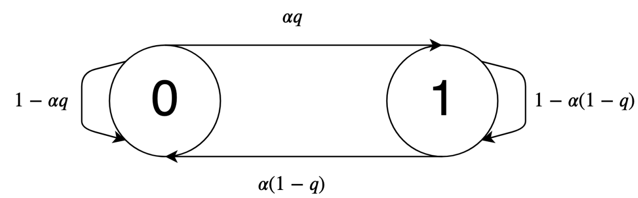
For animals that interact on one day to be more likely to interact on the next, we require that $T_{11} > T_{10}$. Using the transition probabilities from the model in Fig. 1b:

$$1 - \alpha(1-q) > \alpha q$$
$$1 - \alpha + \alpha q > \alpha q$$
$$1 - \alpha > 0$$
$$\rightarrow \alpha < 1$$

.



(a)



(b)

Figure 1: Graphical representation of evolution of $A_{ij}(n)$ between states 0 and 1 for different models of interaction.

Additionally, we require that all transition probabilities are nonnegative, and so our fully defined range for α is

$$0 \leq \alpha < 1$$

1.2.2

The stationary distribution π must satisfy $\mathbf{T}\pi = \pi$:

$$\begin{bmatrix} 1-q & 1-q \\ q & q \end{bmatrix} \begin{bmatrix} \pi_0 \\ \pi_1 \end{bmatrix} = \begin{bmatrix} \pi_0 \\ \pi_1 \end{bmatrix}$$

$$(1-\alpha q)\pi_0 + \alpha(1-q)\pi_1 = \pi_0$$

$$-\alpha q \pi_0 + \alpha(1-q)\pi_1 = 0$$

$$\pi_1 = \frac{q}{1-q}\pi_0$$

Combining this with the requirement that $\pi_0 + \pi_1 = 1$, we find

$$\pi = \begin{bmatrix} 1-q \\ q \end{bmatrix}$$

1.2.3

Let the vector $\mathbf{x}(n)$ represent the probability distribution of $A_{ij}(n)$. We can find the distribution at timestep $n+8$ by repeated application of the transition matrix:

$$\mathbf{x}(n+8) = T^8\mathbf{x}(n).$$

For $q = 0.3,\ \alpha = 0.4$, this gives us

$$\mathbf{x}(n+8) = \begin{bmatrix} 0.6882 \\ 0.3118 \end{bmatrix}$$

and therefore

$$P(A_{ij}(n+8) = 1 | A_{ij}(n) = 1) = 0.3118.$$

1.2.4

The autocorrelation function for a gap of 8 days is given by

$$R_A(8) = \frac{\langle A_{ij}(n+8)A_{ij}(n)\rangle - \langle A_{ij}(n)\rangle^2}{\mathrm{Var}\{A_{ij}(n)\}}.$$

Using our computed stationary distribution for $\langle A_{ij}(n)\rangle$ and $\mathrm{Var}\{A_{ij}(n)\}$, we can write this as

$$\begin{aligned} R_A(8) &= \frac{P[A_{ij}(n+8) = 1|A_{ij}(n) = 1]P[A_{ij}(n) = 1] - 0.3^2}{0.3 - 0.3^2} \\ &= 0.0169 \end{aligned}$$

1.2.5

Assuming that our model is accurate, the ecologists' previously recorded interactions will be of very little predictive value following an 8 day holiday, since we have just seen that the correlation between two states of the system separated by 8 days is close to zero.

2 Social groups

2.1

2.1.1

For the first column of the transition matrix, we take our current state (at day n) to be $(0,0,0)$, corresponding to $A_{12} = A_{13} = A_{23} = 0$. Since there are no interactions on day n, all interaction probabilities on day $n+1$ are given by βq - our first column therefore looks like this:

$$T_1 = \begin{bmatrix} (1-\beta q)^3 \\ (1-\beta q)^2\beta q \\ (1-\beta q)^2\beta q \\ (1-\beta q)^2\beta q \\ (1-\beta q)(\beta q)^2 \\ (1-\beta q)(\beta q)^2 \\ (1-\beta q)(\beta q)^2 \\ (\beta q)^3 \end{bmatrix}.$$

For the fifth column, our state at day n is $(0,1,1)$, corresponding to $A_{12} = 0,\ A_{13} = A_{23} = 1$. Since animal 1 and animal 2 both interacted with animal 3 on day n, the probability of interaction on day $n+1$ between animals 1 and 2 is equal to q, while both other interaction probabilities remain at βq. Our fifth column therefore has the form:

$$T_5 = \begin{bmatrix} (1-q)(1-\beta q)^2 \\ (1-q)(1-\beta q)\beta q \\ (1-q)(1-\beta q)\beta q \\ q(1-\beta q)^2 \\ (1-q)(\beta q)^2 \\ \beta q^2(1-\beta q) \\ \beta q^2(1-\beta q) \\ \beta^2 q^3 \end{bmatrix}.$$

For the eighth column, out state at day n is $(1,1,1)$. So now, on day $n+1$, every pair of animals has previously interacted with a common third animal, and so all interaction probabilities are equal to q. Our eighth column is thus:

$$T_8 = \begin{bmatrix} (1-q)^3 \\ (1-q)^2 q \\ (1-q)^2 q \\ (1-q)^2 q \\ (1-q)q^2 \\ (1-q)q^2 \\ (1-q)q^2 \\ q^3 \end{bmatrix}.$$

2.1.2

After loading the transition matrix into matlab, we can determine the stationary distribution by using `eig` to find the eigenvector with a corresponding eigenvalue of 1, and then normalise the eigenvector by dividing by its sum. This gives us a stationary distribution of

$$\pi = \begin{bmatrix} 0.5560 \\ 0.1167 \\ 0.1167 \\ 0.1167 \\ 0.0254 \\ 0.0254 \\ 0.0254 \\ 0.0179 \end{bmatrix}$$

2.1.3

The probability that the interaction A_{23} exists if both $A_{12} = A_{13} = 1$ in the stationary distribution is given by

$$\begin{aligned} P[A_{23} = 1 | A_{12} = A_{13} = 1] &= \frac{P[\text{state} = (1,1,1)]}{P[\text{state} = (1,1,1)] + P[\text{state} = (1,1,0)]} \\ &= \frac{\pi_8}{\pi_8 + \pi_7} = 0.4134 \end{aligned}$$

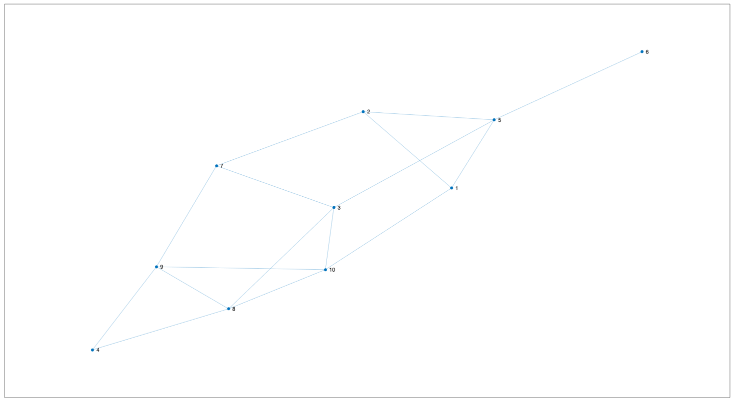


Figure 2: Animal social network

Similarly,

$$P[A_{23} = 1 | A_{12} = A_{13} = 0] = \frac{P[\text{state} = (0,0,1)]}{P[\text{state} = (0,0,1)] + P[\text{state} = (0,0,0)]}$$
$$= \frac{\pi_2}{\pi_2 + \pi_1} = 0.1734$$

2.2

2.2.1

Loading the adjacency matrix in matlab, a graph can be created with `G = graph(A)`. The average degree of the graph is then calculated to be 3.2.

2.2.2

If we have some adjacency matrix $\mathbf{A}$, then $(\mathbf{A}^n)_{ij}$ gives us the number of distinct paths of length n from node i to node j. So $(\mathbf{A}^n)_{ii}$ will give us the number of distinct cycles of length n, starting from node i, that exist in the graph. Therefore, computing trace($\mathbf{A}^3$) yields the total number of distinct cycles of length 3 contained in the graph. This value is not the same, however, as the number of triangles. Each triangle will actually contribute 6 distinct 3-cycles to the total - because a cycle can start at any of the triangle's three points, and can go in either one of two directions around the triangle. This leads us to the following formula for the number of triangles in the graph:

$$N_{\text{tri}} = \frac{\text{trace}(\mathbf{A}^3)}{6}.$$

Evaluating this for the adjacency matrix provided, we get $N_{\text{tri}} = 4$. This result can be verified by examining the plot of the network, shown in Fig. 2.

2.2.3

In an undirected Erdos-Renyi network with M nodes and average degree $\bar{k}$, any one specific edge exists with probability given by

$$P(E_1) = q = \frac{\bar{k}}{M-1}$$

.

Therefore the probability of any three specific edges existing is given by

$$P(E_1, E_2, E_3) = q^3 = (\frac{\bar{k}}{M-1})^3$$

(since the existence of any one particular edge in the network is independent from the existence of any other edge). In a network of M nodes, we can form $\binom{M}{3}$ distinct sets of three nodes. Our expected number of triangles is then given by

$$\langle N_{\text{tri}} \rangle = \binom{M}{3} (\frac{\bar{k}}{M-1})^3$$

which, for $\bar{k} = 3.2$, $M = 10$, comes out as 5.3939. If the triangular motif truly were overrepresented in our social network, we would expect the number of triangles in the network to be significantly higher than the expected number of triangles in the corresponding null Erdos-Renyi network. In fact, the number is slightly lower, suggesting that the motif is not in fact overrepresented in our social network.