Modeling Complex Systems (CS/CSYS 302), Fall 2020: Mean-field Models on Networks

Compartmental models

Compartmental models are a class of models where we divide the parts of our systems in different "compartments". The main assumption is that we assume that all parts in the same compartment are indistinguishable (same state, same behaviour). These models are a canonical approach to epidemiology, but also popular in pharmacokinetic (where compartments are organs or tissues, and parts are particles or concentrations of substances following blood flow, diffusion or ventilation processes. The mechanisms underlying these models can always be represented by a schematic of boxes (compartments) and arrows (flows), and modelled by differential equations.

The Susceptible-Infectious-Susceptible (or SIS for short) model of disease spread is a classic example of a compartmental model since there is a discrete number of states for individuals (S or I) and every individual in the same state follows the same rules: Infectious individuals recover at rate α and susceptible individual get infected by an infectious neighbour at rate β. If the system is *fully connected*, meaning everyone is everyone else's neighbour, we can follow the number of infectious individuals with the following differential equation:

$$\frac{dI}{dt} = \beta SI - \alpha I = \beta(N - I)I - \alpha I \ . \tag{1}$$

Where the first term is the logistic growth of the contagion — meaning we have an exponential growth with rate β up to a constant carrying capacity N — and the second term is the number of recovery per infinitesimal timestep dt. It is also common to use a dimensionless renormalized time $\tilde{t} = \alpha t$ and rate $\lambda = \beta/\alpha$ such that $d\tilde{t} = \alpha dt$ and

$$\frac{dI}{d\tilde{t}} = \frac{1}{\alpha}\frac{dI}{dt} = \lambda(N - I)I - I \ . \tag{2}$$

Homogeneous mean-field on networks

The equations used above are correct if the system is fully connected, but on a network only a subset of all pairs of susceptible and infectious individuals are actually neighbours. Really, we want something like

$$\frac{dI}{d\tilde{t}} = \lambda[SI] - I \ , \tag{3}$$

where $[SI]$ is the number of of edges that connect susceptible individuals to infectious individuals. There are no exact ways to mathematically keep track of $[SI]$, instead we rely on mean-field approximations.

Mean-field approximations simply mean that we will approximate interaction terms beyond a certain scale. For example, if we know the network is an Erdős-Rényi random graph where every edge exists with probability p, we can write

$$\frac{dI}{d\tilde{t}} = \lambda[SI] - I = \lambda pI(N - I) - I \ . \tag{4}$$

To highlight how the mean-field works, let us now focus on the *fraction* $\tilde{I}$ of infectious individuals rather than the *number* I of infectious individuals. The equation becomes

$$\begin{aligned}
\frac{d\tilde{I}}{d\tilde{t}} = \frac{1}{N}\frac{dI}{d\tilde{t}} &= \lambda p\tilde{I}(N - I) - \tilde{I} \\
&= \lambda p\tilde{I}(N - I) \cdot \frac{N}{N} - \tilde{I} \\
&= \lambda pN\tilde{I}(1 - \tilde{I}) - \tilde{I} \\
&= \lambda\langle k\rangle\tilde{I}(1 - \tilde{I}) - \tilde{I} \ .
\end{aligned} \tag{5}$$

In this last equation, we can now see that we are essentially taking every susceptible individual (a fraction $1 - \tilde{I}$ of the population) and connecting them with the average degree $\langle k\rangle$ to average or random neighbours that are infectious with probability $\tilde{I}$. So we are not following a network per se, but a population of nodes connected with strength $\langle k\rangle$ to a mean-field, or an approximation of the real system.

Pairwise models

The previous section assumes that the average degree $\langle k\rangle$ is a good description of the network. We know that that is not true in heterogeneous systems, which is why we want to consider networks in the first place: Some individuals are way more connected and central in the netwoirk than the average. One way to incorporate that is by also looking at the excess degree of nodes.

Excess degree is a quantity relevant when following an edge and it refers to the number of *other* edges that the node reached has. Think of it as your friend's number of other friends (their number of friends minus one). As we have seen in class, the average excess degree is not simply the average degree since edges are more likely to lead to high degree nodes then low degree nodes. To model this fact, we can directly follow the state of edges by adding three compartments describing the numbers of edges that are in the three following states: $[SS]$, $[SI]$ and $[II]$.

If you draw this new compartmental models you will realize that there are no arrows between our two previous compartments (nodes in S or I) and the three new compartments. Indeed, there are no mechanisms to turn nodes into edges, or edges into nodes. But the state of nodes and edges are still coupled by the transition rates associated with the arrows.

For example, the state of nodes follow a differential equation that depends on $\tilde{I}$ and $[SI]$:

$$\frac{d\tilde{I}}{d\tilde{t}} = \lambda[SI] - \tilde{I} \, , \tag{6}$$

if we assume that $[SI]$ describes the number of edges in state $[SI]$ divided by N (i.e., $[SI]$ edges per node). We know that the number of edges must be equal to $\langle k \rangle N/2$ since there are $\langle k \rangle N$ stubs (or half-links) and it takes two stubs to create an edge. Therefore, under the normalization by N, the sum $[SS] + [SI] + [II]$ must be equal to $\langle k \rangle/2$. While the sum is constant, the density of pairs in one given state can fluctuate. For example, $[SI]$ turns to $[II]$ is the susceptible node is infected or to $[SS]$ if the infectious node recovers.

It is useful to draw the compartments and model all the possible transitions, and it is important to remember that edges are affected by events on other edges. For example, a $[SS]$ edge can turn to $[SI]$ if that node is infected. In fact, for every of the $\beta[SI]$ infections, the infected susceptible node has an average excess degree $\langle k_{\text{ex}} \rangle$ which can be $[SS]$ edges turning to $[SI]$ edges if they lead to susceptible nodes. The probability that a stub of a susceptible node reaches another susceptible node is equal to the ratio of stubs in $[SS]$ to stubs of S nodes: $2[SS]/\langle k \rangle S$. We can thus write:

$$\frac{d[SS]}{d\tilde{t}} = [SI] - \lambda[SI]\frac{\langle k_{\text{ex}} \rangle}{\langle k \rangle}\frac{2[SS]}{S} \, , \tag{7}$$

where the ratio $\langle k_{\text{ex}} \rangle/\langle k \rangle$ allows us to account for heterogeneous degree distributions. Really, the term $[SI]\left(\langle k_{\text{ex}} \rangle/\langle k \rangle\right)\left(2[SS]/S\right)$ is an approximation of the number of triplets $[ISS]$ using only pairwise information. This approximation, similar to a mean-field approximation, is called a **moment-closure**, which is formally any approximation of higher order terms (or moments) using lower order terms. In this particular case, the fact that we rely on pairwise information (i.e. the edges) is why we call this approach a pairwise model.

Similarly, writing out every possible term for $[SI]$ leads to

$$\frac{d[SI]}{d\tilde{t}} = 2[II] - [SI] + \lambda[SI]\left(\frac{\langle k_{\text{ex}} \rangle}{\langle k \rangle}\frac{2[SS]}{S} - \frac{\langle k_{\text{ex}} \rangle}{\langle k \rangle}\frac{[SI]}{S} - 1\right) \, , \tag{8}$$

where the terms in the parenthesis account respectively for $[SS]$ edges turning to $[SI]$, $[SI]$ edges turning to $[II]$ because the S is infected through its excess degree, or $[SI]$ edges turning to $[II]$ because the S is infected through that edge.

Finally, we do not need to write out $d[II]/d\tilde{t}$ since at every point in time, we know that $[II]$ is given by

$$[II] = \frac{\langle k \rangle}{2} - [SS] - [SI] \, . \tag{9}$$

3

Heterogeneous mean-field models

The previous section accounts for heterogeneity of the degree distribution in its moment-closure, but fails to account for dynamical correlations. In factm the average degree of susceptible nodes is expected to be smaller than the average degree of infectious nodes since nodes of higher degree get more exposure to the disease.

The problem is that our compartments are bad; they are too coarse-grained. We put all susceptible nodes in the same compartment, ignoring the fact that they do not all have the same degree and therefore are not exposed to the same risks (or really, the same transition rates). One approach to account for that is to create $k_{\max}$ compartments, where $k_{\max}$ is the highest degree in the network. Infectious nodes of degree k are therefore not simply in the I compartment, but in the I_k compartment. As we will see next week, we are now using a **configuration model** to describe a dynamical process on the ensemble of all networks with the same degree distribution, assuming nodes of equal degree follow the same rules.

W can write an equation to follow the fraction $\tilde{I}_k$ of nodes in a given compartment

$$\frac{\tilde{I}_k}{d\tilde{t}} = \lambda k(p_k - \tilde{I}_k)\theta - \tilde{I}_k \ , \tag{10}$$

where p_k is the total fraction of nodes of degree k and therefore the carrying capacity for $\tilde{I}_k$ which can only be between 0 and 1. The moment closure is now θ which we interpret as the probability that a random edge around a susceptible node leads to an infectious node. We can calculate it like so

$$\theta = \frac{\sum_k k\tilde{I}_k}{\sum_k kp_k} \equiv \frac{\sum_k kI_k}{\sum_k kN_k} \tag{11}$$

where we follow the same logic as before and calculate the ratio of stubs connected to infectious nodes (numerator) to the total number of stubs (denominator).

With this set of equations, we can easily get the total fraction of infectious nodes by summing over all possible degrees for infectious nodes:

$$\tilde{I} = \sum_k \tilde{I}_k \ . \tag{12}$$

More complex mean-field models

More complex mean-field models exist, for example to account for group structure or degree correlations. Let me know if you have a structure that interest you and we can brainstorm how to write down a model for it. The truth is that if you can describe a network in words, you can probably do math using that description! We will see more on that next week when we talk about models *of* networks.