0.1 Summary

Let us recap what we saw at the end of the last lecture. We moved from well-mixed populations to contact networks, so we add more complexity and the model is more realistic. We also wrote down what is a generic model for SIS dynamics on a general network. We consider the adjiacency matrix and so on. We cannot write down a closed equation for this, because we have two nodes infection probability. The problem is that we do not have an exact expression for this probability. The expression for this probability take into account the probability of three nodes ijk. This is unfeasible for all the models and all the possible graph, in the literature we have just 4/5 nodes. The idea is that we ned a way to approximate this probability, to cut this infinite chain to a certain value. We will use mean-field approximation. We have a quantity that depends on all different quantities but which in this approximation is reduced to its average. We are switching from a many body problem to a one body problem.

Today, we will see different ways to deals with the probability $Prob[\sigma_i = 0, \sigma_j = 1]$.

We start to see a model of SIS for homogeneous network in which all the nodes are equal. There is the same probability of getting infection and we can consider the probabilityies statistically independent. We derived all the equations. The solutions we found are the same of before, the difference is that now it is not exact.

1

Epidemic Spreading on Networks

Today we are going start to analyze what happens when we consider hetheroge-neous networks.

What is the effect of hetherogeneity in the spread of the disease? This assumption $k_i \sim \langle k \rangle$ does not hold, so we cannot assume that all the nodes are equal.

1.1 Degree-based Mean-Field theories (DBMF)

We start with the simple level which is the individual one, in which we consider the individual probability of getting the infection.

Let us start with the paper "Epidemic Spreading in Scale-Free Networks". They provide a model for SIS on scale-free networks. We cannot use the homogeneous approximation, because the network is hethereogeneous. The intuition of this paper are:

- the nodes are not equal. The probability of getting the infection strongly depends on their position (i.e. degree) in the network;

- nodes with the same degree behave in the same way;

We are gonna divide the network in degree classes. We are grouping togheter all the nodes with the same degree. To write down the equation we need to multiply the number of compartments:

$$s_k = \frac{S_k}{N_k}, \qquad \rho_k = \frac{I_k}{N_k}$$

where s_k and ρ_k are the fraction of suscpetible/infected nodes of degree k in the network. We have that N_k represent the number of nodes with degree k in the network. So, we are defined as before the number at degree k of suscpetible and infected in the system.

From that we can write the equation:

$$\frac{}{t}\rho_k(t) = -\mu\rho_k(t) + \beta k \,(1 - \rho_k(t))\Theta_k(t)$$

we have as usual a recovery and infection part. In particular, we have the probability of a contact between a suscpetible of degree k and an infected represented in green. The idea behind it is that we have the probability of being infected $(1 - \rho_k(t))$ and the probability of having contact with an infected $\Theta_k(t)$. We have that:

$$\Theta_k(t) = \sum_{k'} P(k'|k)\rho_{k'}$$

is the probability that a node with degree k as an infected neighbor. We want to sum over all the possibile degree classes and we are gonna see the probability of connecting with one of them, hence this is the probability that another node is infected.

Note that we are making no assumption about the function $P(k'|k)$ which will change with k. We are gonna make an assumption about the structure of this thing. It could be in principle anything, in the sense that it depends on the structure of the network. However, there are some cases in which we can do some assumptions on the structure of the network.

We can assume that network are hethereogenous but I am making connection at random. Hence:

$$P(k'|k) = \frac{k'P(k')}{\sum_{k'} k'P(k')} = \frac{k'P(k')}{\langle k \rangle}$$

where $P(k')$ is the probability of getting a connection at random. Then, we multiply it by k' which is the number of connection that I pick up. Then we normalize over the average degree of the network. Hence, at the end it is the probability that a point in the network points to k'.

Hence:

$$\Theta_k(t) = \frac{\sum_{k'} P(k')\rho_{k'}(t)}{\langle k \rangle} = \Theta(t)$$

In the numerator: we take the probability that a link taken at random points to k', then I am multiplying by the probability of being infected and then I am summing to all the possible degree. We note that this probability does not depend on k anymore. We are just picking up at random, so it should be the same for all the nodes.

The method that we are gonna used to solve the differential equation of $\rho_k(t)$ is pretty similar to the ones used before. First of all, we assume that we are in the steady state:

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho_k(t) \to \rho_k \frac{\beta k \Theta}{\mu + \beta k \Theta}$$

The next step is to substitute the expression obtained inside Θ:

$$\Theta_k(t) = \frac{\sum_{k'} P(k')\rho_{k'}(t)}{\langle k \rangle} = \Theta(t) \to \Theta = \frac{1}{\langle k \rangle} \sum_k \frac{k^2 P(k)\beta\Theta}{\mu + \beta k \Theta}$$

The point is: if we want to solve this expression, we need some sort of trick. First of all, what happens is that as usual this expression as different solution. The first one is the trivial solution $\Theta = 0$, but we are interested in the non trivial one. Note that:

$$\Theta = \frac{1}{\langle k \rangle} \sum_k \frac{k^2 P(k)\beta\Theta}{\mu + \beta k \Theta} = f(\Theta)$$

Hence, the solution are the values of Θ were the two values are equal. This is the interception between the line Θ and the function $f(\Theta)$. We want that since Θ is a probability $0 < \Theta \leq 1$. This means that the slope of $f(\Theta)$ should be greater than 1.

The next step, is seeing what means mathematically that we want a slope larger than one:

$$\frac{\mathrm{d}}{\mathrm{d}\Theta}\left(\frac{1}{\langle k \rangle} \sum_k \frac{k^2 P(k)\beta\Theta}{\mu + \beta k \Theta} \right)_{\Theta=0} \geq 1$$

which means

$$\frac{\beta}{\mu \langle k \rangle} \sum_k k^2 P(k) \geq 1 \qquad \to \frac{\beta \langle k^2 \rangle}{\mu \langle k \rangle} \geq 1$$

this is the condition for an endemic state. Since the network has becoming more complex, also the structure for the condition of the endemic state is becoming complex. If we assume that:

$$\frac{\beta \langle k^2 \rangle}{\mu \langle k \rangle} = 1 \rightarrow \beta_c = \frac{mu \langle k \rangle}{\langle k^2 \rangle}$$

which is pretty similar to the one found before but has a term which increase is complexity.

We have to check if it works also for an homogeneous network. This is the first check that we can make. In homogeneous netowrks $\langle k^2 \rangle = \langle k \rangle^2$, recovering:

$$\beta_c = \frac{mu \langle k \rangle}{\langle k^2 \rangle} = \frac{\mu}{\langle k \rangle}$$

this is exactly the expression that we saw before. Things are working well.

Recalling what we say last week, in scale-free networks with $2 < \gamma \leq 3$, we have $\langle k \rangle \rightarrow c$ and $\langle k^2 \rangle \rightarrow \infty$ as $N \rightarrow \infty$. As the netowork is becoming larger, also its variance is becoming larger. This means that:

$$\beta_c = \frac{\mu \langle k \rangle}{\langle k^2 \rangle} \rightarrow 0$$

the epidemic threshold is going to zero. Obviously, this is quite important because if my network is big enough, every disease will spread, no matter its infectivity! When we have disease with a very low infection in a little part of population, they do not disappear because we are in a large network! Hence, we have no more an endemic state and the epidem threshold is really really small for most real epidemic network. This result work as in a thermodynamic limit.

Obviously, this cannot happen in real network. What happens is that we need some finite-size correction. For example, an expression for epidemic thresold when size cannot reach infinity.

If we use scale free distribution, at some point since the degree cannot go to infity, we introduce a cut-off. For instance, for a social network we can reach some number of followers but the system will have a cut off.

We intdroduce an exponential cut-off for this distribution. For instance in an air trasportation network, we see that until a certain point we have a line, then the curve start to change. We cannot have an infinite number of connection. we have a line and then we will see some sort of exponential degree.

We can model this kind of things by adding an expoential term:

$$P(K) \sim k^{-\gamma e^{-k/k_c}}$$

where k_c is a characteristic degree. At some point the term added will becomes the dominant term. What happens? For large k_c and $2 < \gamma < 3$ the epidemic threshold reads as:

$$\beta_c \left(\frac{\mu k_c}{k_{min}} \right)^{\gamma - 3}$$

we will not show the calculation. Tomorrow, we will compare the epidemic treshold in an random network and in a scale-free network to see their differences. This were the result for the SIS.

1.1.1   SIR model

We can write the same equations for the SIR model under the same assumption. We need more equations to take into account also the equation for R.

$$\frac{\mathrm{d}}{\mathrm{d}t} \rho_k^I(t) ...$$

where we have:

$$\Gamma_k(t)$$

which plays exactly the same role of Θ of before. It represents the link from which the infection arrived at the node. We will not show its form. It is little different because of the structure of the SIR. There is a neibhobour that can trasmit the infection to me, but it can recover. Hence, this individual can be suscpetible again. I need to take into account that the disease is coming from one side, because that side of the network for me is forbidden in the sense that I have recovered that connot be infected again.

We found:

$$\beta_c = \frac{\mu \langle k \rangle}{\langle k^2 \rangle - \langle k \rangle}$$

and the important things is that $\beta_c^{SIS} \neq \beta_c^{SIR}$. This is the first time in the course that the epidemic treshold for these two models is not the same.

1.2 Individual Based Mean-Field theories (IBMF)

Before we were assuming that all the nodes with the same degree were equal. Now, we are gonna studying the indivudual based mean-field theories for individuals. We are not consider a single network, but an average over all the possible network we can create given that degree distribution.

Assuming that all the nodes with the same degree are equal, we are not looking at the single network but at their average. This is what in physics is called **annealed networks**. In the opposite when I called **quenched networks**, I am consider a particular realization of one network. The idea is: instead of considering the average we are considering a particular network. This is de difference in doing a degree bases or an indivudual based.

Let write down the equation for the quenced mean field. We are gonna use the discrete time formulation of equation because it is a bit simpler (we can write also the same things with differential equations).

We are consider $\rho_i(t)$ as the probability of a node of being infected at time t. What is my probability of being infected at time $t + 1$ is:

$$\rho_i(t+1) = \rho_i(t)(1 - \mu) + (1 - \rho_i(t))q_i(t)$$

which is the probability of being infected and not get cured and the second is the normal term when I am getting the infection (the probability of being suscpetible and the probability of getting the disease). We have that $q_i(t)$ is the probability of node i of getting infetced by at least one neighbour.

We need an expression for $q_i(t)$. The basic idea is that:

$$q_i(t) = 1 - \prod_{j=1}^{N}[1 - \beta A_{ij}\rho_j(t)]$$

We have the node i in green, and the infected neighbour in red. We see that $\beta A_{ij}\rho_j(t)$ is the probability of getting infetced by node j (at least). We have that $[1 - \beta A_{ij}\rho_j(t)]$ is the probability of NOT getting infected by node j. I have to repeat this calculation for all my infected nehbours, hence $\prod_{j=1}^{N}[1 - \beta A_{ij}\rho_j(t)]$ is the probability of NOT getting infected by any neighbor. Finally, $q_i(t)$ represents the probability of getting infected by at least one neighbor. Hence, the probability of getting infected is 1 minus the probability of not getting infected by any neighbor.

Hence, once we have this things, first of all we can solve it numerically, which means for instance by iteration. In this case we are gonna have an equation for each

of the node. We have 2^N equation where N is the size of the system. The equation for $\rho_i(t+1)$ can be solved by iteration. I am gonna repeat this equation many times until I reach the steady state.

The difference with the degree based mean field theories is that actually A_{ij} we are including all the adjiacency matrix, while before only the average.

We can also solve analytically the system at teh steady state to estimate the epidemic threshold. I am assuming that if I am in the epidemic treshold, what happens is that ρ is very small for all the nodes. If $\rho_i^* = \varepsilon_i^* \ll 1$, we can use an approximation for q_i^*:

$$q_i^* = 1 - \prod_{j=1}^{N}[1 - \beta A_{ij}\varepsilon_j^*] \sim \beta \sum_{j=1}^{N} A_{ij}\varepsilon_j^*$$

Substituting we have:

$$\mu\varepsilon_i^* = \beta(1 - \varepsilon_i^*) \sum_{j=1}^{N} A_{ij}\varepsilon_j^*$$

We have a linear system where the interaction is represented by the adjiacency matrix:

$$\frac{\mu}{\beta}\varepsilon_i^* = \sum_{j=1}^{N} A_{ij}\varepsilon_j^*$$

This linear system has solution only if $\frac{\mu}{\beta}$ is an eigenvalue of the adjiacency matrix A_{ij}. This is why last week we say that the spectrum of the adjiacency matrix was important. Hence:

$$\beta = \frac{\mu}{\Lambda_i}$$

Since, we are interested in the smallest possible value of β for which we have a solution, we take the largest eigenvalue of the adjiacency matrix A:

$$\beta_c = \frac{\mu}{\Lambda_{max}}$$

This is the expression for the epidemic treshold, and it is a general results that it is valid not only by this approxmation but for a generic case for a generic network.

What is the relation between the epidemic treshold for DBMF and IBMF? We have that:

$$\beta_c^{DBMF} = \frac{\mu \langle k \rangle}{\langle k^* \rangle}, \qquad \beta_c^{IBMF} = \frac{\mu}{\Lambda_{max}}$$

The relation amogn them for scale-free networks is:

$$\Lambda_{max} \sim max(\sqrt{k_{max}}, \langle k^* \rangle / \langle k \rangle)$$

Specifically:

$$\beta_c \sim \begin{cases} \mu/\sqrt{k_{max}} & \gamma > 5/2 \\ \mu \langle k \rangle / \langle k^2 \rangle & 2 < \gamma < 5/2 \end{cases}$$

We can conclude that IBMF is more accurate than DBMF.