When is a network complex? Connectance as a driver of degree distribution

T. Poisot

Introduction

- complexity in networks
- degree distribution vs. connectance
- "physical" argument

Statistical argument

Assuming and ecological network made of n species, and assuming undirected interactions with no self-edges (*e.g.* no cannibalism), there can be at most $M = n(n-1)/2$ interactions in this network, in which case it is a complete graph (the results presented below hold for both directed graphs, and graphs in which self-edges are allowed). This maximal number of links, M_n, represent the whole space of possible links. With this information in hand, it is possible to know the total number of possible networks given a number l of interactions.

If we term S_n the set of all possible M_n edges in a n-node network, then the number $G_{n,l}$ of possible networks with l links is the number of l-combinations of S_n, meaning that $G_{n,l} = C_l^{M_n}$, or

$$G_{n,l} = \frac{M_n!}{l!(M_n - l)!}$$

Note that this number of possible networks include some graphs in which nodes have a degree of 0, and that in most ecological studies, will be discarded. Finding out the number of graphs in which some nodes have a degree of 0 is similar to finding out how many networks exist with l links between $n-1$ nodes. If one node is removed from the network, there are C_{n-1}^{n} possible combinations of nodes. For each of these, there are $G_{n-1,l}$ possible networks configurations. Note that these networks will also include situations in which *more* than one species has a degree of 0. Calling $R_{n,l}$ the number of networks with n nodes and l edges in which all nodes have at least one edge attached, we can write

$$R_{n,l} = G_{n,l} - C_{n-1}^{n} \times G_{n-1,l}$$

Based on these informations, we can make two predictions.

Prediction 1: $R_{n,l}$ and $G_{n,l}$ will be maximized when l is close to $M_n/2$. In other words, the maximal number of possible networks occurs when connectance is intermediate.

Prediction 2: $R_{n,l}$ will become asimptotically closer to $G_{n,l}$ when l is close to M_n. In other words, there is only one way to fill a network of n nodes with M_n interactions, and in this situation there is no possibility to have nodes with a degree of 0.

We now illustrate these predictions using networks of 10 nodes, with a number of edges varying from 10 to M_{10} (*i.e.* 45 edges).

- probability to generate a suitable network

Simulations

In the previous part, we show mathematically that connectance (the number of realized *vs.* possible interaction), relative to the network size, determined the size of the *network sapce*, *i.e.* how many possible network combinations exist. Based on this, we can therefore predict that the degree distribution will be contingent upon network connectance. Specifically, we expect that the variance of the degree distribution, which is often used [@fortuna_nestedness_2010], will display a hump-shaped relationship with connectance. The mean, kurtosis, and skewness of the degree distribution should all vary in a monotonous way with connectance.

In the simulations below, we use a network of 25 nodes, filled with 30 to M_{25} interactions.

- unipartites
- explain the procedure
- give results for average, variance, kurtosis, skewness

Practical consequences

- null models
- swap