Simulating Gene Regulatory Networks to Estimate Stochastic Transition Variation

J. Lopez[1] and J. Arbelo[2]
Faculty Mentor: O. Akman[3], Ph.D.
1. National University of Colombia, 2. University of Puerto Rico - Rio Piedras , 3. Illinois State University

Abstract

Genes are often characterized as segments of DNA that provide a blueprint for cells and organisms to properly function. One of their main purpose is to effectively help control the different processes and regulations that take place within every organism. Interestingly, there are no current biological methods that can easily reveal the details of gene interactions. This inspired a group of mathematicians and biologists to use distinct methods to make such calculations. Out of the many methods used, one of the most effective was the employment of Markov Chains to simulate gene expression data to infer networks of interactions, which we dub as gene regulatory networks (or GRNs).

In this work, our job was to not only to extend the concepts found in the paper *An Investigation of Gene Regulatory Network State Space Variability*, by Liesman et al. (2021) but to adopt the ideas found in *Modeling Stochasticity and Variability in Gene Regulatory Networks*, a research paper written by Murrugarra et al. (2012), and *Dynamics of Gene Regulatory Networks with Stochastic Propensities*, a paper written by by Dr. Akman et al. (2018), and adapt them into a practical sense by making various simulations. The simulations were made using a statistical program known as R. By the end of the experiment, the R program shows that the propensities found in Gene Networks are not in fact seen as constant, as stated in the Murrugarra paper, rather they can take on different values that can be changed with a certain probability. The code was made public through GitHub repository where is presented the special case of three genes which we are going to present here.

Introduction

In 2012, the paper *Modeling Stochasticity and Variability in Gene Regulatory Networks* was published by the University of Nebraska and written by David Murrugara, alongside a group of mathematicians and biologists. This paper laid the foundation for the types of models that could be best associated with gene networks and the probability of a change state gene. One of the main systems this document proposes is the representation of a gene turning on and off, with a 1 or 0 respectively. This change can be a product of internal noise like biological uncertainty or external noise like experimental noise produced by measurements, or latent variables related with environmental conditions. However, while the research paper was very precise with what it was trying to accomplish, it assumed that the propensities, or the percentage at which the gene would turn on and off, of each gene could be seen as constant. This realization is what inspired Dr. Akman, and Liesman et al, to write a paper where the propensities were seen as changing rather than constants. Our focus for this research is to update the idea of stochasticity in gene networks, more importantly we will focus on the idea of internal noise and external noise modeling GRN state transition (edge) variability using a Beta distribution (this is because the range is between 0 and 1 and has multiple shapes).

Methods

The method of this research consists in making a code in R, where there are an n number of genes, that are limited to two states (on or off). In this instance, we used 3 different genes: Gene A, Gene B and Gene C. These genes will be attributed a random probability for activation and degradation as a starting point. Afterwards, applying the functions, found in the Murrugara paper, for the activation and degradation of genes, we will calculate the probability concerning the change of the state of each gene. Subsequently, the code will generate random variances for the beta distribution, to be used to calculate the parameters for both alpha and beta. Later, we will run 100 simulations for each beta distribution with original propensities as mean and random variance, to then calculate the probability of a new change in state accompanied by an interval generated by variance simulations in which that probability can take values. The variance can be seen in the graph 3, where the smaller the variance the bluer it appears, while the bigger the variance is, the greener it appears.

Update functions

$$\pi_{i,x}\left(x_i \to x_i\right) = \begin{cases} 1 - p_{i,act} & x_i < f_i(x) \\ 1 - p_{i,degr} & x_i > f_i(x) \\ 1 & x_i = f_i(x) \end{cases}$$

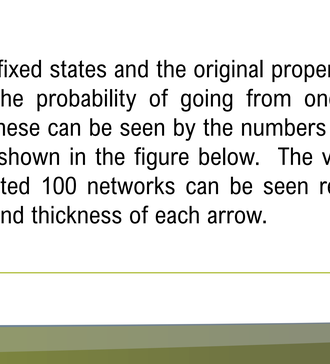
$$\pi_{i,x}\left(x_i \to f_i(x)\right) = \begin{cases} p_{i,act} & x_i < f_i(x) \\ p_{i,degr} & x_i > f_i(x) \\ 1 & x_i = f_i(x) \end{cases}$$

$$\pi_{i,x}\left(x_i \to y_i\right) = 0 \quad \forall\, y_i \notin \langle x_i, f_i(x) \rangle$$

Results

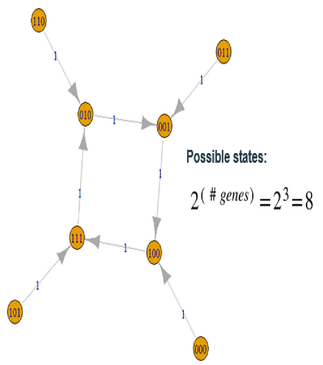
The code gives us a graph that represents the interaction between genes. The same shows the activation and degradation of the genes using colors.

Genes interaction



The code first calculates the fixed states, where the number states depends on the number of genes.

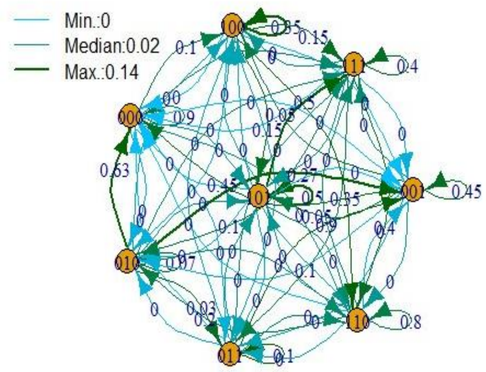
Fixed states



Initial state			Next state		
A	B	C	A	B	C
0	0	0	1	0	0
0	0	1	1	0	0
0	1	0	0	0	1
0	1	1	0	0	1
1	0	0	1	1	1
1	0	1	1	1	1
1	1	0	0	1	0
1	1	1	0	1	0

Using the fixed states and the original propensities, we calculate the probability of going from one state to another. These can be seen by the numbers colored in dark blue shown in the figure below. The variance of the simulated 100 networks can be seen reflected in the color and thickness of each arrow.

Simulations network



Conclusion

By generating a large number of simulations of the propensities, we can then approximate the probability of going from one state to another as a normal distribution, taking the original probability as the mean and the estimated variance of the simulations. This would mean that the transition probabilities can be found within the interval $\hat{p} \pm z_{\frac{\alpha}{2}} * \sigma_{\hat{p}}$ with an uncertainty of $\alpha * 100\,\%$. Therefore, we can indeed state that the propensities do not have to be seen as constants, rather can be seen as constantly changing values.

Public code can be found in the next repository: https://github.com/JuliethLopez/Stochastic-gene-regulatory-networks

References

- Murrugarra, D., Veliz-Cuba, A., Aguilar, B., Arat, S., & Laubenbacher, R. (2012). Modeling stochasticity and variability in Gene Regulatory Networks.
- Akman, O., Comar, T., L.Harris, A., Hrozencik, D. and Li, Y. (2018). Dynamics of gene regulatory networks with stochastic propensities.
- Liesman, Sara; Akman, Olcay; and Hrozencik, Daniel (2021) "An Investigation of Gene Regulatory Network State Space Variability," *Spora: A Journal of Biomathematics*: Vol. 7, 86–103.