Modeling Diseases, Mutations, and Vaccines as Interacting Species

Furkan Toprak
Texas A&M University
November 2020

Abstract

The purpose of this paper is to demonstrate how SIR models can be represented as Lotka-Volterra models in order to analogize S, I, and R as competing species in an ecosystem. Furthermore, we later discover the flexibility of this analogy as we expand the SIR and generalized Lotka-Volterra models to include vaccinations and mutations of the modeled epidemic. The content of this paper consists of several instances of theoretical proof and practical demonstration. Iteratively, we will first expand and explain our mathematical models and secondly demonstrate the legitimacy of our model with some Python code output. We find that vaccines can be analogized to species that do not directly interact with each other but depend on the same prey, I. Similarly, we also find that mutations can be seen as species that do not interact with each other but depend on the same prey, S. We lastly conclude that generalized Lotka-Volterra models are sufficient enough to model the complexities of epidemics with mutations and vaccinations; even more notably this translation highlights that susceptible populations, vaccinated populations, infected populations, and resistant populations can be seen as species that are competing within the same ecosystem.

Introduction

SIR models help visualize the spread, growth, and outcomes of epidemics. SIR models emulate the dynamics of a population by classifying the population within three exclusive categories: Susceptible, Infected, and Recovered (henceforth abbreviated as S, I, and R and mathematically represented as S(t), I(t), and R(t), respectively). The simplest nonlinear models represent the interactions between the three categories using a system of ordinary differential equations; namely the Kermack-McKendrick Model is represented below:

- $\frac{dS}{dt} = -\beta SI$
- $\frac{dI}{dt} = \beta SI - \gamma I$
- $\frac{dR}{dt} = \gamma I$

  where β represents the infection rate and γ represents the recovery rate.

The novelty of this model comes from the two terms βSI and γI. Simply put, SI represents the number of all interactions that involve 1 member of S and 1 member of I. The term βSI, therefore, represents all probabilistically *expected* infections at each time step. The slope of S at every time step, therefore, represents the expected number of infections deducted from the set of members S. Additionally, the term γI and the slope of R represent the expected number of

recovered infected members. Lastly, the slope of I is the expected growth of the set I at each time step; explicitly it is the difference between the number of new infected members and the number of recovered infected members.

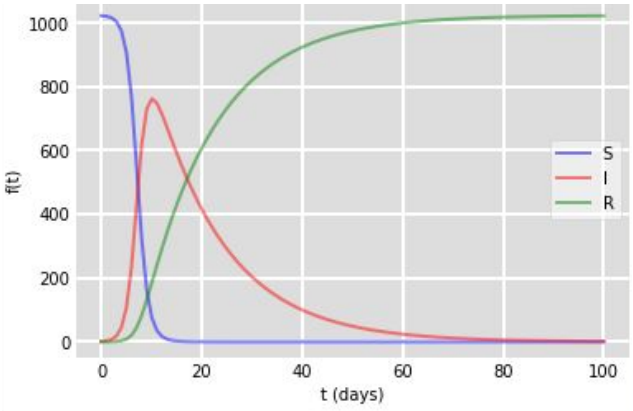

Figure 1: An example of Kermack-McKendrick Model, where $N = 2^{10}$, $\beta = 0.001, \gamma = \frac{1}{14}$.

For the sake of simplicity and tangibility, we will adopt the Kermack-McKendrick Model as our exemplary SIR model, which we will manipulate and modify in our explorations below. It is important to understand the nature of the basic SIR model, and specifically the Kermack-Mckendrick Model, in order to appreciate its fundamental accuracy when portraying epidemics.

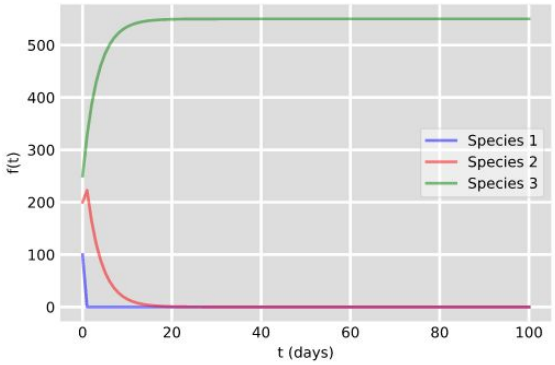

Figure 2: An example of a 3-species Lotka-Volterra system, where the total population of the ecosystem is 550 members.

Generalized Lotka-Volterra equations are models that represent interactions between species and allow both positive and negative among any of the N species. The set of ordinary differential equations between N species are represented as[1]:

[1] Remien, C. H., Eckwright, M. J., & Ridenhour, B. J. (2018). Parameter Identifiability of the Generalized Lotka-Volterra Model for Microbiome Studies. doi:10.1101/463372

$\frac{dN_i}{dt} = \sum_j f_{ij}(N_i,\ N_j)\ +\ \sum_k g_{ik}(N_i,\ X_k)$.

F, the community matrix, is an NxN matrix representing the $O(N^2)$ network of interactions between any two species in the ecosystem of N species. Conceptually, the first term corresponds to the growth of N_i when interacting with N_j. The second term has to do with interactions of species with exogenous variables, which do not exist in SIR models, and will be set to zero.

Conceivably, we can see how generalized Lotka-Volterra equations can be used to represent epidemics in the same manner SIR models can. In this paper, we will use this fact to analogize epidemics as three interacting species: S, I, and R. We will then expand our SIR (Kermack-McKendrick) model to additionally analogize mutations of diseases (where each mutation causes a new strain of the disease) and vaccinations (where each vaccine inhibits only one strain) to adding more species to our system of generalized Lotka-Volterra equations.

Limitations

In order to conceptually facilitate and explore mutations and vaccines, the SIR models are constrained such that a member that is in set R is indeed resistant and not dead (removed from the population) and no member may be infected by two mutations at once. Otherwise, the effect of adding new strains could not be properly compared due to overlapping of multiple diseases and a rapidly diminishing population. Lastly, the total population is static in all models. Therefore, any advances to the flexibility of this model may result from additional research.

Moreover, we will be introducing vaccines and mutations. We will assume widely that members can only be infected with one mutation at a time, to "avoid unresolvable interaction terms"[2].

Modelling Simple SIR Models as Generalized Lotka-Volterra Equations

In order to convert our SIR model as a system of Generalized Lotka-Volterra Equations, we must establish matrices N, F, and G. As aforementioned,

$G = 0;\ N = [S(t),\ I(t),\ R(t)]$; R(0) = 0, I(0) = 1, S(0) = n - I(0) - R(0)

Lastly,

$F(X, Y) = [[0,\ \beta XY,\ 0],\ [\beta XY,\ -\gamma X,\ 0],\ [0,\ \gamma Y,\ 0]]$. We have thus converted the Kermack-McKendrick model into a generalized Lotka-Volterra Equation. Additionally, we have characterized the following:

- S can be analogized to a species that derives its growth non-linearly from the size of itself and I, times an infection constant (β).

[2] Rundell, W. (2020). [E-mail to the author].

- I can be analogized to a species that derives its growth from itself, the number of susceptible people times an infection constant (β), and the recovery rate (γ).
- R can be analogized to a species that derives its growth from I, without regard to its own size.

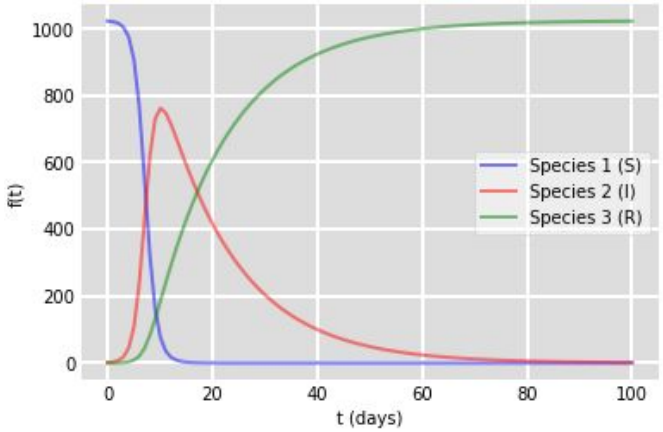

Figure 3: An example of a 3-species Lotka-Volterra system, modeled to replicate the configuration of Figure 1.

Modelling Nuance: Vaccines

Vaccines help slow the spread of epidemics by reducing the number of susceptible members. To demonstrate the effectiveness of vaccination regimes, see Figures 4.1, 4.2, 4.3.

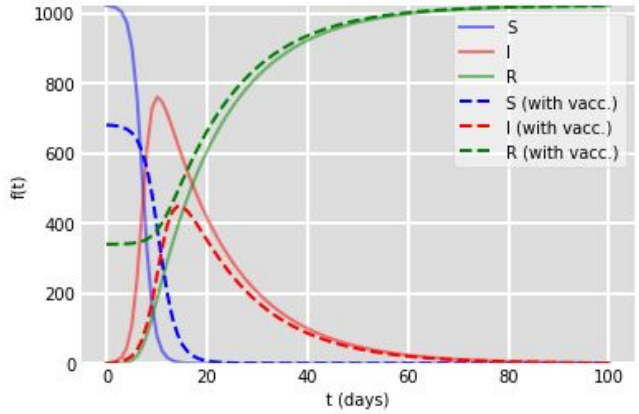

Figure 4.1: A version of Figure 1 where 1/3rd of the original population is originally vaccinated.
Discussion: This demonstrates how an initially vaccinated body of members causes epidemics to end earlier and with less infections. Moreover, infection peaks occur later and less rapidly under vaccination regimes.

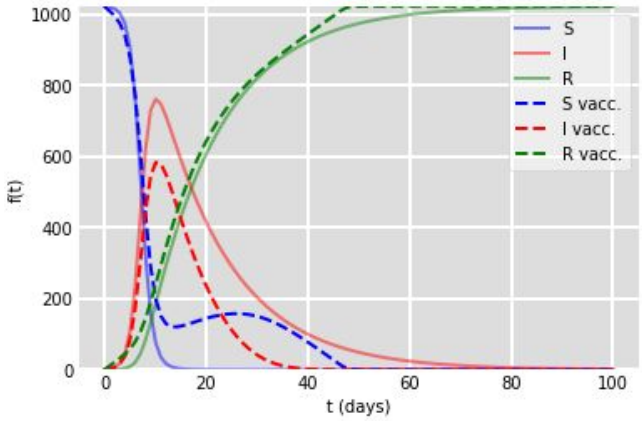

Figure 4.2: A version of Figure 1 where none of the original population is originally vaccinated but 1 percent of the population gets vaccinated every day. Discussion: This demonstrates how the effectiveness of an already-vaccinated population is much greater than a rapid vaccination regime.

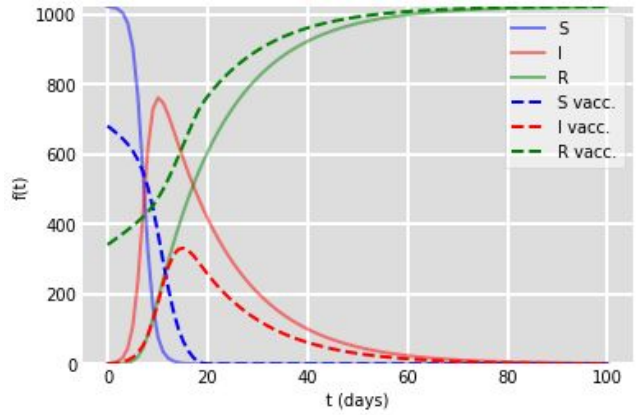

Figure 4.3: For the sake of being thorough, this figure represents the benefit of both of the aforementioned techniques.

To introduce a gradual vaccination to our SIR model, $\frac{dS}{dt}$ should be decreased and $\frac{dR}{dt}$ should be increased by the same amount. To assume an initial vaccination rate to our SIR model, it is sufficient to decrease S_0 and increase R_0 by the same amount. The effect of the initial benefit gained from υ vaccines, in fact, can be modeled by decreasing S(t) by V(t) and increasing R(t) by V(t) where $V(t) = \sum_{i=1}^{\upsilon} n_i(t)$ and $n_i(t)$ is the number of members whose earliest vaccination was vaccine i (in order to not count duplicates) at time t.

Let it be noted that we are omitting, for brevity's sake, the fact that vaccinations are not effective forever and must be periodically re-applied. However, if one were to model this nuance, they would apply some delayed decay to R and growth to S by a term that would be dependent on $V(t-\tau)$, where τ is the lifetime of a vaccine.

Let us find the corresponding generalized Lotka-Volterra model: Now,

$$G = 0;\ N = \left[S(t),\ I(t),\ R(t),\ V_1(t), \ldots,\ V_\upsilon(t)\right]$$

$$V(0) = R(0) = \sum_{i=1}^{\upsilon} V_i(0)\ ;I(0) = 1,\ S(0) = n - I(0) - \sum_{i=1}^{\upsilon} V_i(0)$$

and F is equivalent to the matrix represented in Table 1.

	1 (S)	**2 (I)**	**3 (R)**	**4 (V_1)**	**...**	**5(V_υ)**
1 (S)	0	$-\beta XY$	0	$-min(\theta_{j-3} * n,\ S)$	...	$-min(\theta_{j-3} * n,\ S)$
2 (I)	βXY	$-\gamma X$	0	0	...	0
3 (R)	0	γY	0	$min(\theta_{j-3} * n,\ S)$	...	$min(\theta_{j-3} * n,\ S)$
4 (V_1)	$min(\theta_{i-3} * n,\ Y)$	0	0	0	...	0
⋮	⋮	⋮	⋮	⋮	⋱	0
5(V_υ)	$min(\theta_{i-3} * n\ ,\ Y)$	0	0	0	...	0

Table 1: The F matrix of the new model. The left column represents the row index, i, whereas the top dashed row represents the column index, j. The red additions represent the expansion to the community matrix when we incorporate υ vaccines. Above, $X = N_i, Y = N_j$. θ_k represents the vaccination rate of the population using vaccine *k*.

The legitimacy of this method is demonstrated below in Figure 5:

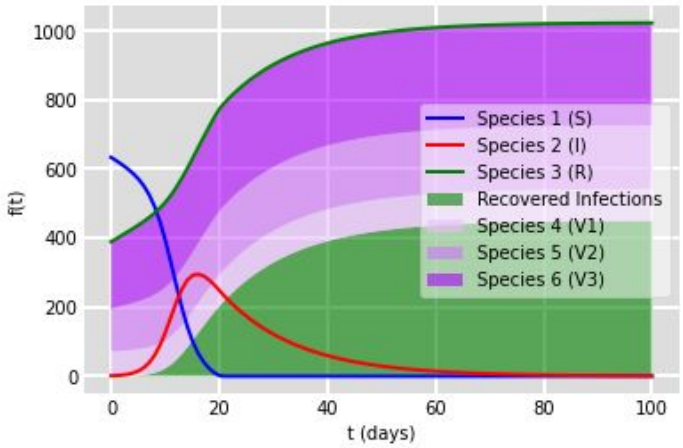

Figure 5: A demonstration of how vaccines can be modeled as species when using generalized Lotka-Volterra equations. Discussion: For more meaningful interpretation, the composites of R (at least numerically) which are the number of infected and recovered members, vaccine 1

recipients, vaccine 2 recipients, and vaccine 3 recipients are stacked. As can be seen here, species 4, 5, and 6 are cooperative and inhibit S.

Modelling Nuance: Mutations

Our SIR model can be expanded by adding ω more infection classes such that our system is now:

- $\frac{dS}{dt} = -S\sum_{i=1}^{\omega}\beta_i I_i$
- $\frac{dI_i}{dt} = S\beta_i I_i - \gamma_i I_i$
- $\frac{dR}{dt} = \sum_{i=1}^{\omega}\gamma_i I_i$

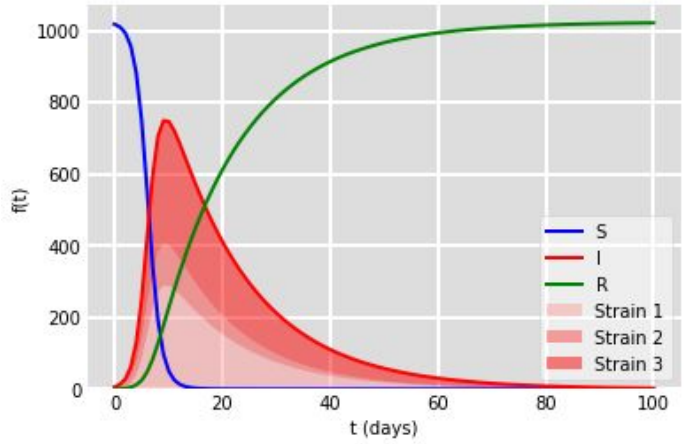

Figure 6: An example of a Kermack-McKendrick model with 3 strains of I.

The equivalent of this expansion to the generalized Lotka-Volterra model would look like the following:

$$G = 0;\ N = \left[S(t),\ I_1(t), ...,\ I_\omega(t),\ R(t)\right]$$

$$R(0) = 0\ ;I(0) = \sum_{i=1}^{\omega} I_i(0),\ S(0) = n - I(0) - R(0)$$

and F is equivalent to the matrix represented in Table 2.

	1 (S)	**2 (I_1)**	**...**	**1 + ω (I_ω)**	**2 + ω (R)**
1 (S)	0	$-\beta_{j-1}XY$	...	$-\beta_{j-1}XY$	0
2 (I_1)	$\beta_{i-1}XY$	$-\gamma_{j-1}Y$	...	0	0
⋮	⋮	⋮	⋱		0
1 + ω (I_ω)	$\beta_{i-1}XY$	0	...	$-\gamma_{j-1}Y$	0

2 + ω (R)	0	$\gamma_{j-1}Y$	...	$\gamma_{j-1}Y$	0

Table 2: The F matrix of the new model. The left column represents the row index, i, whereas the top dashed row represents the column index, j. The red additions represent the expansion to the community matrix when we incorporate ω strains. Above, $X = N_i, Y = N_j . \beta_k$ represents the infection rate of mutation *k* and γ_k represents the recovery rate of members with mutation *k*. The correctness of this model is demonstrated in Figure 7.

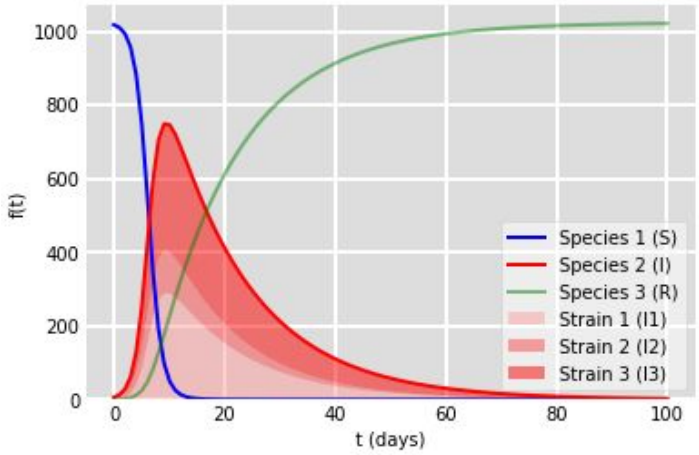

Figure 7: A demonstration of how strains of the same disease can be modeled when using generalized Lotka-Volterra equations. Discussion: This representation highlights how the different strains of the same infectious disease can be analogized to species competing for the same prey (S).

Methodology

For all code, see https://github.com/FurkanToprak/EpidemicModels.git

Conclusion

In conclusion, we have demonstrated how SIR models can be converted into generalized Lotka-Volterra equations. In addition to this, we have demonstrated that both traditional SIR models and the Lotka-Volterra representation can be expanded to include the existence of υ vaccines and ω mutations. In fact, one could trivially combine the two models so that $\upsilon = \omega$ and vaccine *i* is effective only against mutation *i*. While this combination would not result in a finer resolution or higher level of analysis, it would allow for a greater historical view of a case study epidemic and give the capacity to predict mutations and project responses to the new mutations.

In this way, we can model a epidemiological history as an ecosystem of $2 + \omega + \upsilon$ competitive species where the species are constituted from S, ω strains, υ vaccines, and R.

References

Nowak, M. A. (2006). *Evolutionary Dynamics: Exploring the Equations of Life*. Cambridge, MA: The Belknap Press of Harvard University Press.

Remien, C. H., Eckwright, M. J., & Ridenhour, B. J. (2018). Parameter Identifiability of the Generalized Lotka-Volterra Model for Microbiome Studies. doi:10.1101/463372

Rundell, W. (2020). Subject for MATH 308H Final Project [E-mail to the author].