Developing mathematical and epidemiological intuition for the basic reproduction number

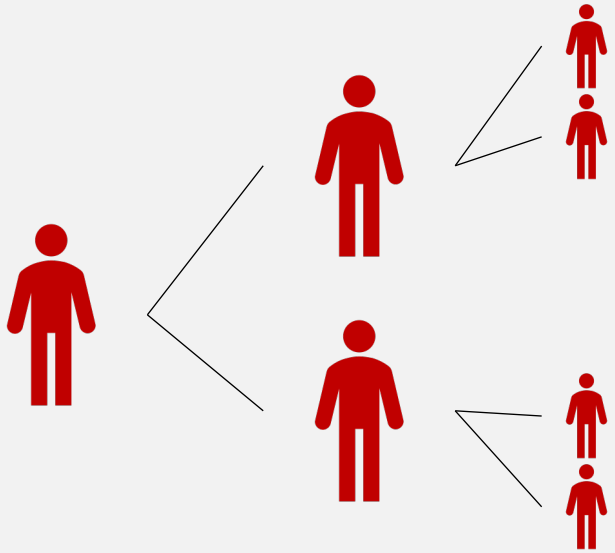
Andrew Brouwer

University of Michigan

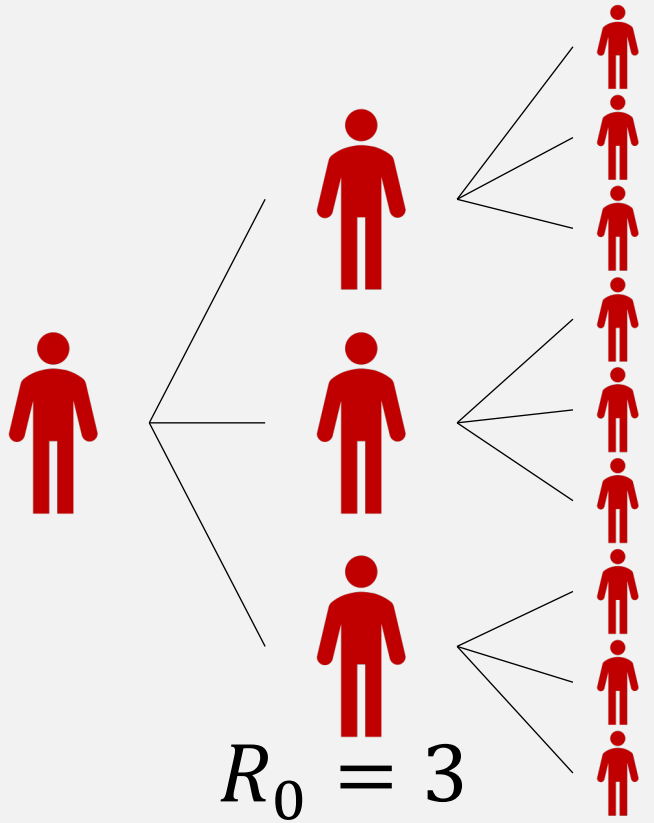
November 10, 2021

Basic reproduction number

- “Expected number of secondary cases arising from a typical primary case in an entirely susceptible population over their infectious period.”



$R_0 = 2$



$R_0 = 3$

Key aspects of the definition

1. Single infectious period in an otherwise susceptible population
   - Technically defined only at the beginning of an outbreak
   - But, R_0 is often used an abstract concept for an infectious disease's epidemic potential
2. *Expected* number of new cases
   - Any specific person will infect more or fewer
3. Over their infectious period
   - There is an interplay between the infectious period, the pathogen's infectiousness, and other aspects of its natural history

The “DOTS”

- The basic reproduction number can be thought of as the product of
  - D: duration
  - O: opportunity
  - T: transmission probability
  - S: susceptibility

Kucharski, 2020

R_0 in the classic SIR model

- In the classic SIR model, R_0 is the contact rate times the average duration of the infectious period

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

$$R_0 = \frac{1}{\gamma} \times \beta$$

Average infectious period

"D"

Transmission rate

"OTS"

Why is the reproduction number important?

- Threshold value for the local stability of the DFE (disease-free equilibrium)
  - In plain language: if R_0 >1, an introduction becomes epidemic; if R_0 <1, it dies out
- Controls the dynamics of the outbreak
  - Larger R_0 → short-lived but explosive outbreaks; reduce R0 to “flatten the curve”
  - Larger R_0 → larger cumulative incidence, i.e., fraction of the population ever infected
  - Larger R_0 → greater immunity needed for herd protection

Calculating R_0

- R_0, or more correctly, R_{eff}, can be calculated from incidence data
  - Not the topic of this lecture
- Mathematical models, which are abstract representations of the disease systems, are used to estimate R_0 as a function of disease-related quantities
- Mathematical representations also allow us to generalize our conception of R_0 when transmission pathways other than person-to-person are involved (e.g., vector, environmental)

The Next Generation Matrix and its geometry

The Next Generation Method

- Most common and rigorous approach to calculating R_0 from a mathematical model
- Developed by Heesterbeek, Diekmann, and colleagues (1990); popularized by Pauline van den Driessche (2002). See references at end of presentation.

The Next Generation Method, part 1

- Define:
  - Let x be the vector of states and x_0 the DFE
  - For each infected compartment i
    - Let f_i be the rate of influx of newly infected people to compartment i
    - Let v_i be the net transfer of individuals out of compartment i by all other means

- Then $\frac{dx_i}{dt} = f_i(x) - v_i(x)$

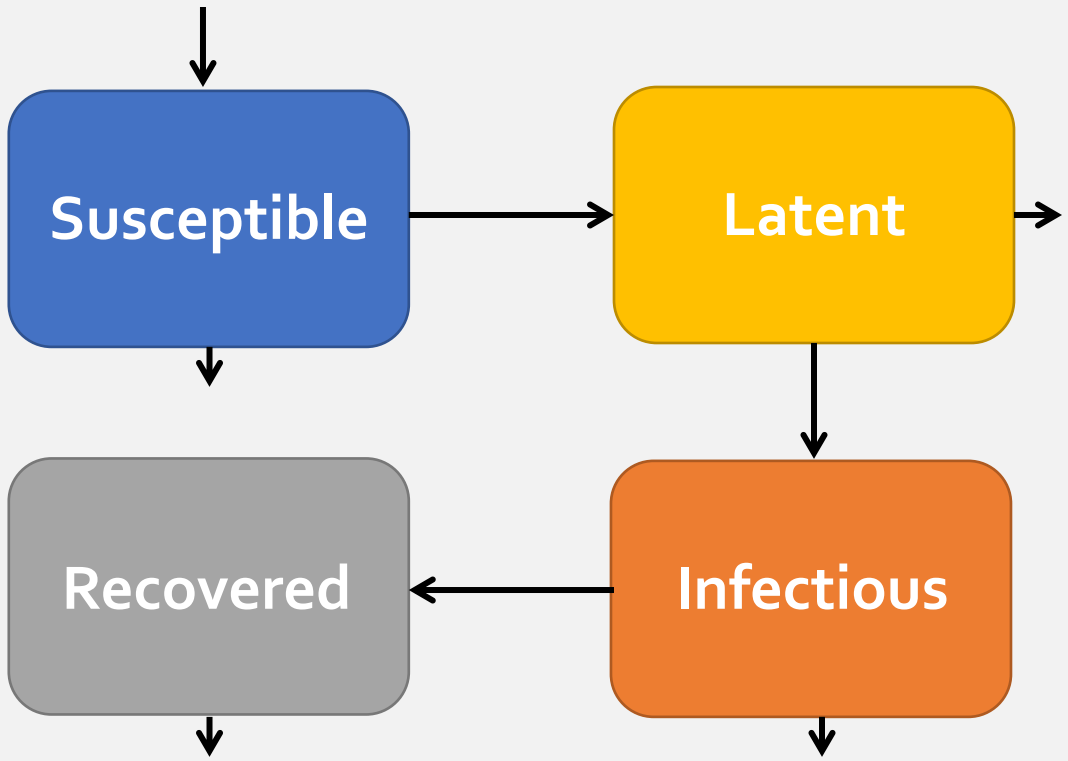
SLIR model with demography

$$\frac{dS}{dt} = \mu - \beta IS - \mu S$$

$$\frac{dL}{dt} = \beta IS - \sigma L - \mu L$$

$$\frac{dI}{dt} = \sigma L - \gamma I - \mu I$$

$$\frac{dR}{dt} = \gamma I - \mu R$$



SLIR model with demography

$$\frac{dS}{dt} = \mu - \beta IS - \mu S$$

infected

$$\frac{dL}{dt} = \underbrace{\beta IS}_{\text{new}} - \sigma L - \mu L$$

$$\frac{dI}{dt} = \sigma L - \gamma I - \mu I$$

$$\frac{dR}{dt} = \gamma I - \mu R$$

$$f(x) = \begin{bmatrix} \beta IS \\ 0 \end{bmatrix}$$

$$v(x) = \begin{bmatrix} (\sigma + \mu)L \\ (\gamma + \mu)I - \sigma L \end{bmatrix}$$

The Next Generation Method, part 2

- Define:
  - Let F and V be the Jacobians (matrix of 1st derivatives) of f and v evaluated at the disease-free equilibrium
  - That is
    - $F_{ij} = \left.\frac{df_i}{dx_j}\right|_{x=x_0}$
    - $V_{ij} = \left.\frac{dv_i}{dx_j}\right|_{x=x_0}$

e.g., $F = \left[\begin{matrix} \frac{df_L}{dL} & \frac{df_L}{dI} \\ \frac{df_I}{dL} & \frac{df_I}{dI} \end{matrix}\right]_{x=x_0}$

SLIR model with demography

$$\frac{dS}{dt} = \mu - \beta IS - \mu S$$

infected

$$\frac{dL}{dt} = \underbrace{\beta IS}_{\text{new}} - \sigma L - \mu L$$

$$\frac{dI}{dt} = \sigma L - \gamma I - \mu I$$

$$\frac{dR}{dt} = \gamma I - \mu R$$

$$x_0 = (1,0,0,0)$$

$$f(x) = \begin{bmatrix} \beta IS \\ 0 \end{bmatrix}$$

$$v(x) = \begin{bmatrix} (\sigma + \mu)L \\ (\gamma + \mu)I - \sigma L \end{bmatrix}$$

$$F = \begin{bmatrix} 0 & \beta \\ 0 & 0 \end{bmatrix}$$

$$V = \begin{bmatrix} \sigma + \mu & 0 \\ -\sigma & \gamma + \mu \end{bmatrix}$$

The Next Generation Method, part 3

- The matrix $\mathrm{F}V^{-1}$ is called the next generation matrix

- The basic reproduction number is the spectral radius (largest eigenvalue) of the next generation matrix.

The Next Generation Method, part 3

- The matrix $\mathrm{F}V^{-1}$ is called the next generation matrix

- The basic reproduction number is the spectral radius (largest eigenvalue) of the next generation matrix.

Sorry, what? Why?

Understanding the Next Generation Method

- The matrix V^{-1} is a matrix of residence times.
- Imagine a newly infected individual entering compartment k:

> The (j, k) entry of V^{-1} is the average length of time this individual spends in compartment j during its lifetime, assuming that the population remains near the DFE and barring reinfection.
>
> -van den Driessche & Watmough, 2002

Understanding the Next Generation Method

- The matrix V^{-1} is a matrix of residence times.
- Imagine a newly infected individual entering compartment k:

> The (j, k) entry of V^{-1} is the average length of time this individual spends in compartment j during its lifetime, assuming that the population remains near the DFE and barring reinfection.
>
> -van den Driessche & Watmough, 2002

Starting in L, you leave when you recover or die.

Starting in I, you never go to L.

$$V^{-1} = \begin{bmatrix} \frac{1}{(\sigma+\mu)} & 0 \\ \frac{\sigma}{(\sigma+\mu)(\gamma+\mu)} & \frac{1}{(\gamma+\mu)} \end{bmatrix}$$

Starting in L, you go on to I, unless you die first.

Starting in I, you recover or die.

Understanding the Next Generation Method

- The matrix F is a matrix of rate of new infections.

  The (i,j) entry of F is the rate at which infected individuals in compartment j produce a new infection in compartment i.

  -van den Driessche & Watmough, 2002

Understanding the Next Generation Method

- The matrix F is a matrix of rate of new infections.

The (i,j) entry of F is the rate at which infected individuals in compartment j produce a new infection in compartment i.

-van den Driessche & Watmough, 2002

Only people in compartment I make new infected people, and those new infected people all start in L.

$$F = \begin{bmatrix} 0 & \beta \\ 0 & 0 \end{bmatrix}$$

Latent people are not (yet) infectious.

Understanding the Next Generation Method

- The matrix $\mathrm{F}V^{-1}$ is called the next generation matrix

  Hence, the (i, k) entry of the product $\mathrm{F}V^{-1}$ is the expected number of new infections in compartment i produced by the infected individual originally introduced into compartment k.

  -van den Driessche & Watmough, 2002

Understanding the Next Generation Method

- The matrix $\mathrm{F}V^{-1}$ is called the next generation matrix

> Hence, the (i, k) entry of the product $\mathrm{F}V^{-1}$ is the expected number of new infections in compartment i produced by the infected individual originally introduced into compartment k.
>
> -van den Driessche & Watmough, 2002

$$FV^{-1} = \begin{bmatrix} \dfrac{\beta\sigma}{(\sigma+\mu)(\gamma+\mu)} & \dfrac{\beta}{(\gamma+\mu)} \\ 0 & 0 \end{bmatrix}$$

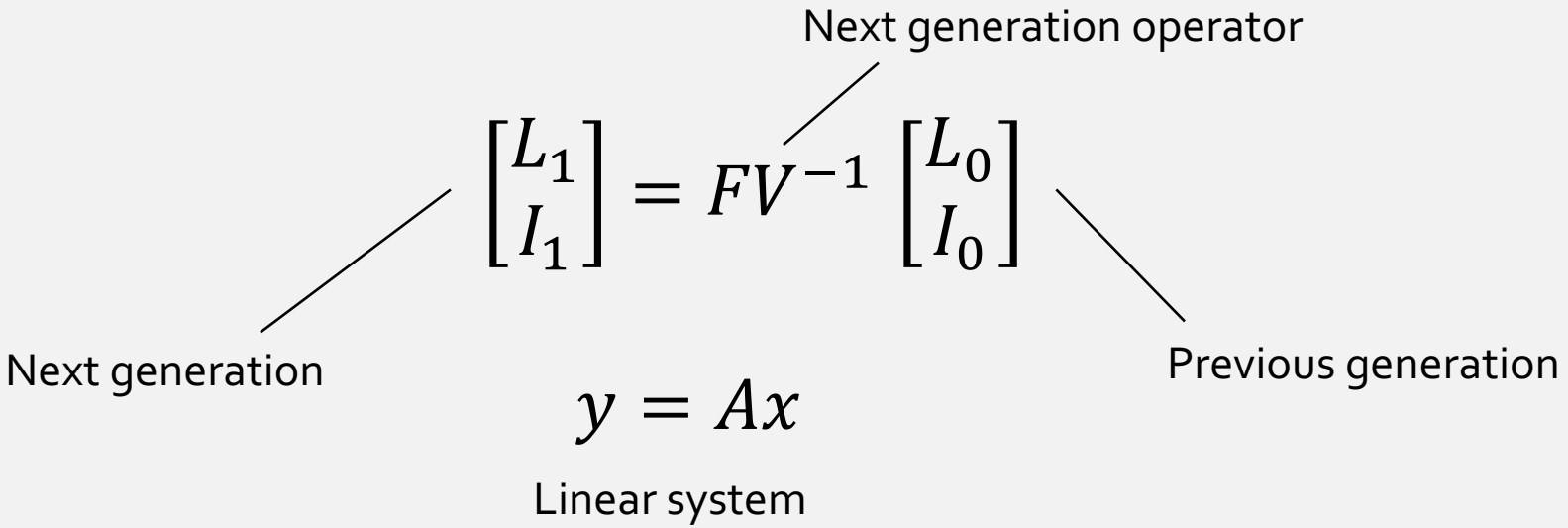
Infected people make new exposed, not new infectious.

Understanding the Next Generation Method

- Okay, we can interpret the entries of the next generation matrix, but why the spectral radius?

Geometry of the NGM

- Near the disease-free equilibrium, the process of making the next generation is approximately linear.

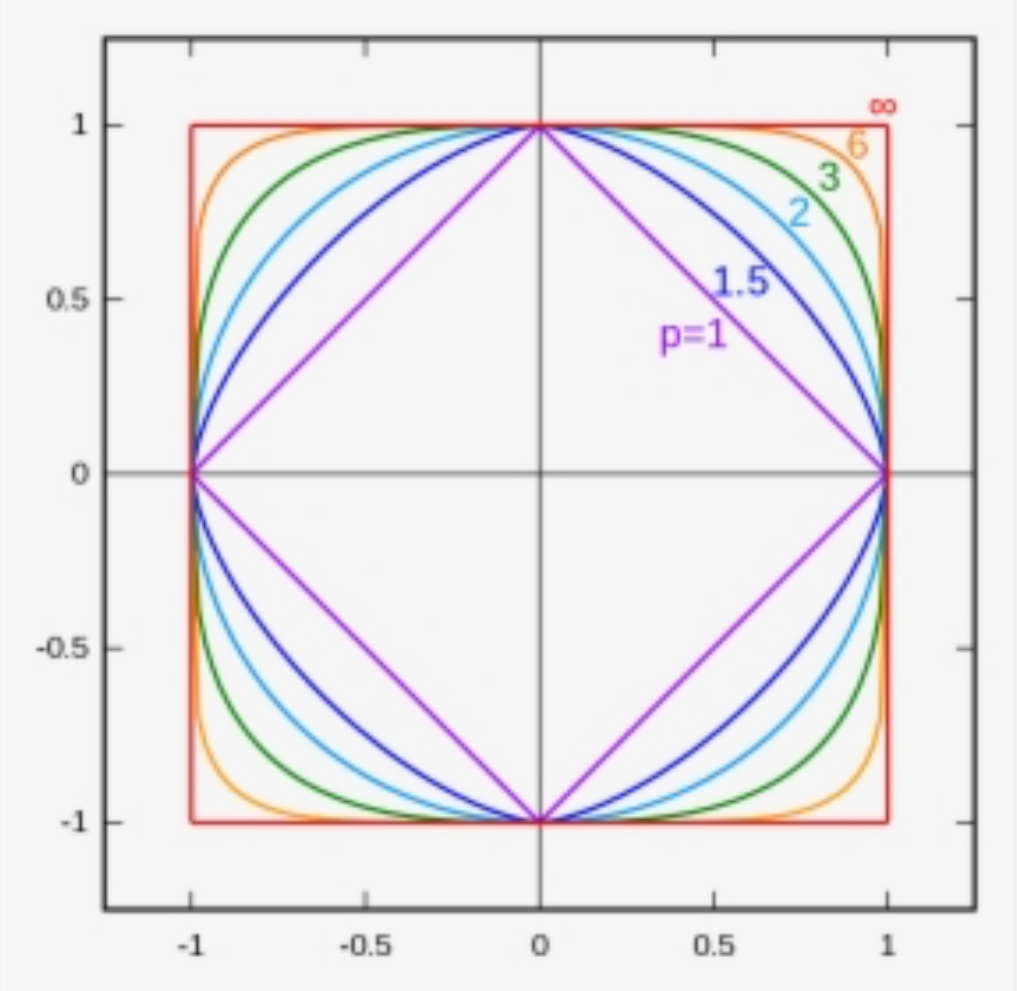


- We can use the geometry of linear systems!

Geometry of the NGM

- Consider the unit circle in the 1-norm
$$\{x: |x| = 1\}$$

- The 1-norm is natural choice because we are partitioning a fixed population into different compartments



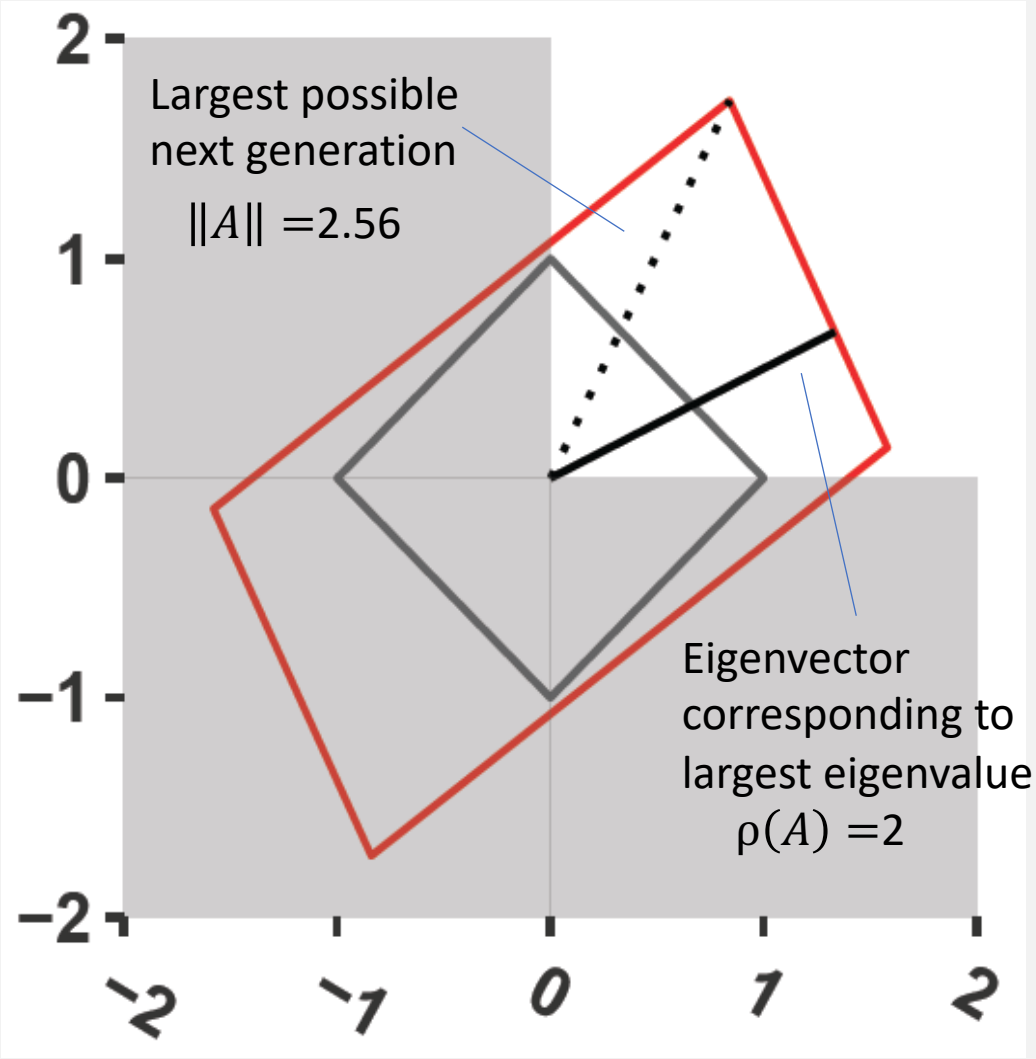
Geometry of the NGM

- Apply a linear transformation to a circle

$$A = \begin{bmatrix} 1.58 & 0.84 \\ 0.14 & 1.72 \end{bmatrix}$$

$$\{Ax: \ |x| = 1\}$$

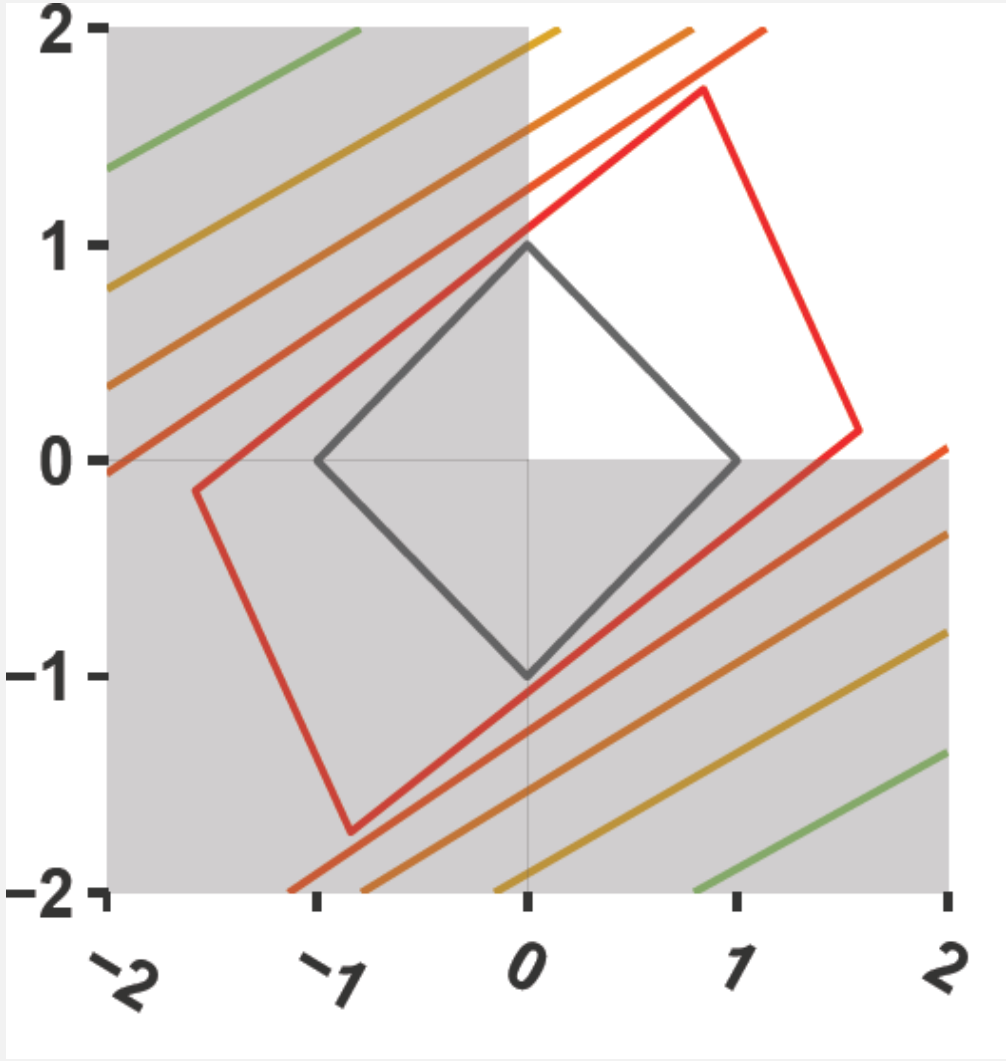
The size of the next generation depends on the partition of the previous generation.



Geometry of the NGM

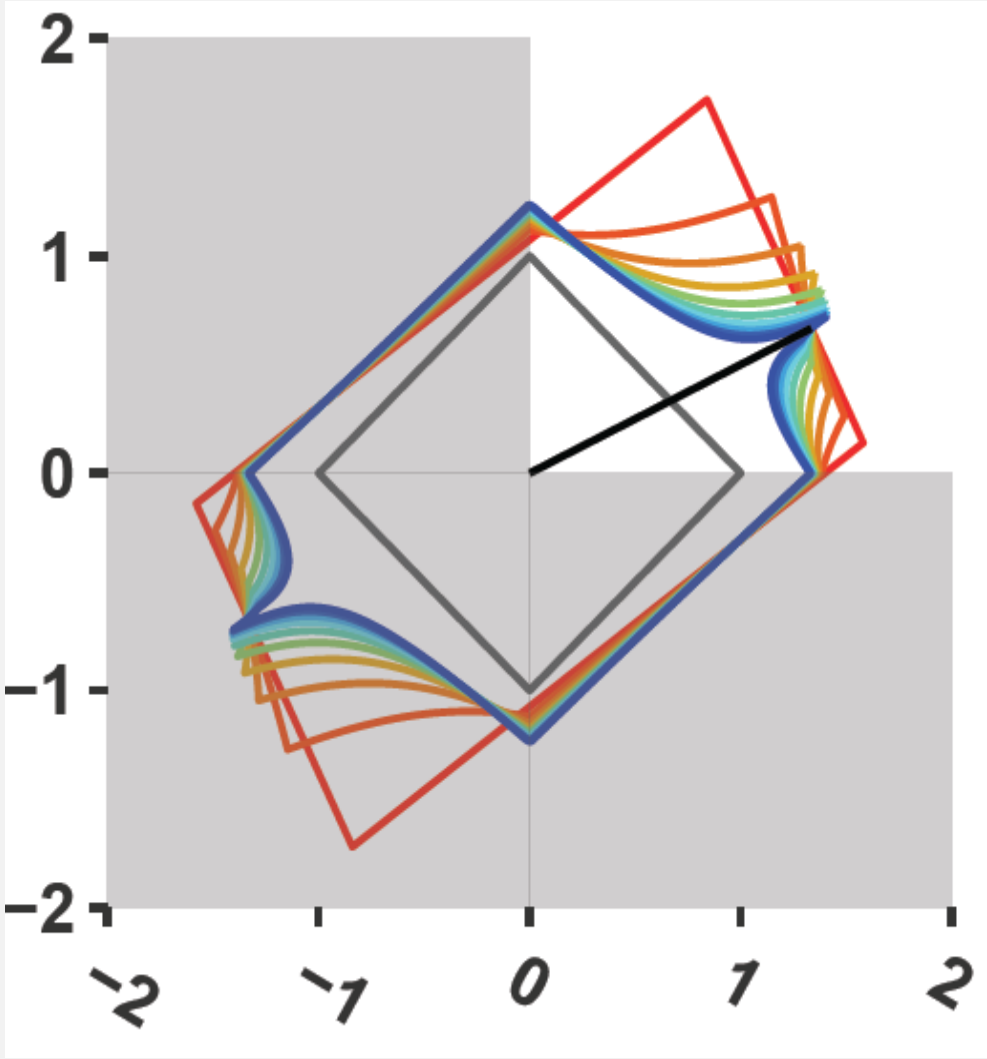
- After multiple generations, $A^n x$, the shape gets larger and becomes exaggerated. This tells us about long term behavior.

- But, we don't care about long-term, linear behavior, per se, which will quickly deviate from the true non-linear behavior. We want *average* behavior.



Geometry of the NGM

- Scale our shape so that we get average generation size, $\sqrt[n]{|A^n x|}$
- As n increases, we see $\sqrt[n]{\|A^n\|}$ converging to $\rho(A)$.
- The size of the average next generation will be $\rho(A)$ because (almost every) initial condition converges to lie along the dominant eigenvector.



Geometry of the NGM

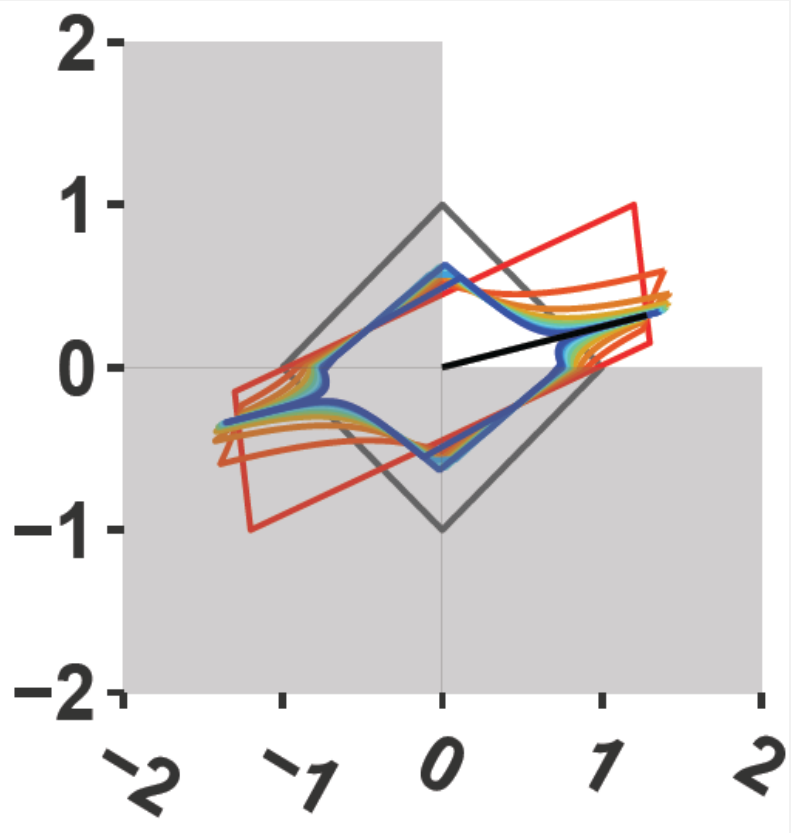
- The eigenvector corresponding to the largest eigenvalue of A also has a useful interpretation. It is the stable distribution of the infectious compartments.

$$\rho(A)\nu = A\nu$$

$$2\begin{bmatrix}2/3\\1/3\end{bmatrix} = \begin{bmatrix}1.58 & 0.84\\0.14 & 1.72\end{bmatrix}\begin{bmatrix}2/3\\1/3\end{bmatrix}$$

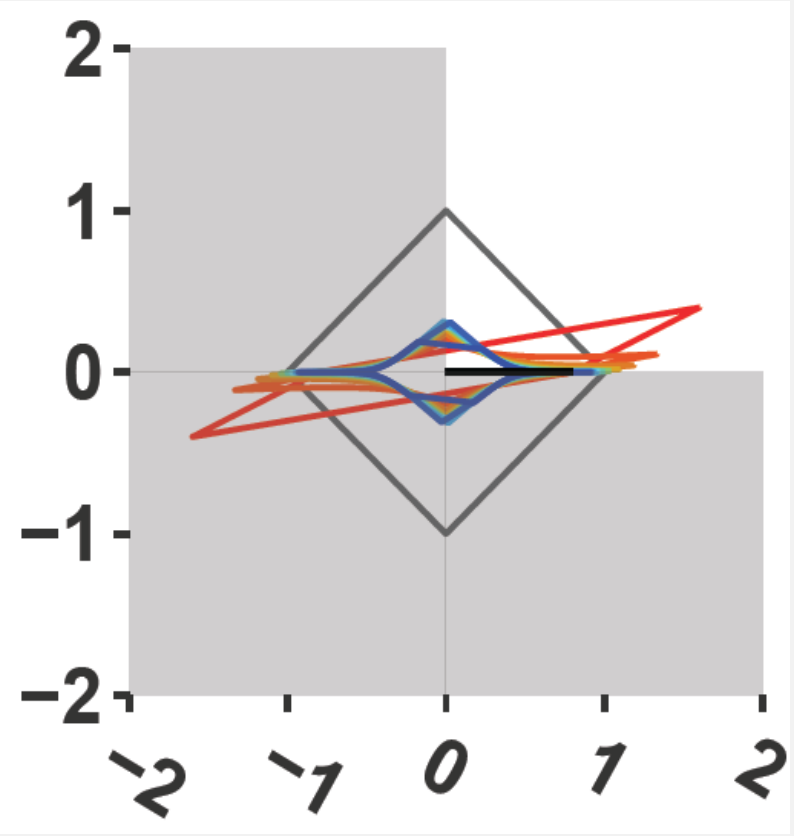
- Note that this stable distribution is unlikely to manifest in reality because an epidemic quickly leaves the vicinity of the DFE where this analysis is valid.

Additional examples



$$B = \begin{bmatrix} 1.30 & 1.20 \\ 0.15 & 1.00 \end{bmatrix}$$

Eigenvalues: 1.6, 0.7



$$C = \begin{bmatrix} 0.80 & 1.60 \\ 0.00 & 0.40 \end{bmatrix}$$

Eigenvalues: 0.8, 0.4

Developing symbolic interpretations of the NGM

Interpreting the NGM

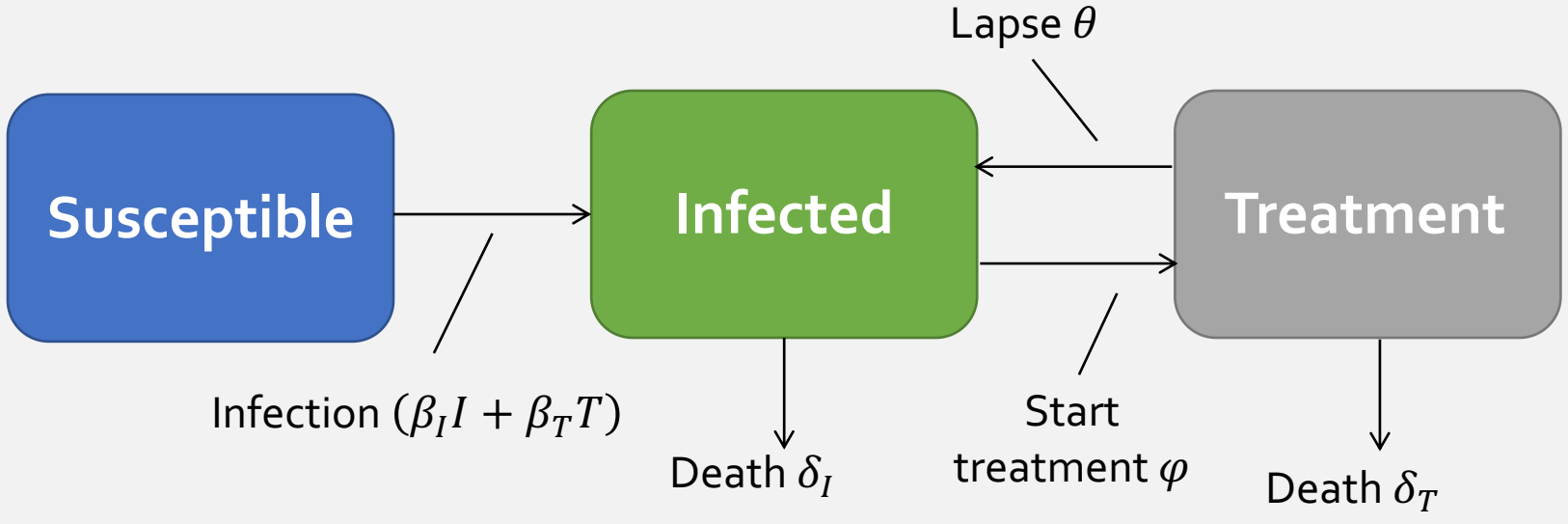
- We now have a better understanding why the spectral radius is the right measure.

- But we still may struggle with interpreting the NGM matrix in terms of our parameters, especially in higher dimensional cases.

Example: Treatment compliance

- Infected individuals can go on and off a treatment that reduces their infectivity and mortality rate

$$\dot{S} = -\mathrm{S}(\beta_I I + \beta_T T)$$
$$\dot{I} = \mathrm{S}(\beta_I I + \beta_T T) - (\varphi + \delta_I)I + \theta T$$
$$\dot{T} = \varphi \mathrm{I} - (\theta + \delta_T)T$$



Example: Treatment compliance

$$f(x) = \begin{bmatrix} \mathrm{S}(\beta_I I + \beta_T T) \\ 0 \end{bmatrix} \qquad \mathrm{F} = \begin{bmatrix} \beta_I & \beta_T \\ 0 & 0 \end{bmatrix}$$

$$\mathrm{v}(x) = \begin{bmatrix} (\varphi + \delta_I) I - \theta T \\ (\theta + \delta_T) T - \varphi \mathrm{I} \end{bmatrix} \qquad \mathrm{V} = \begin{bmatrix} \varphi + \delta_I & -\theta \\ -\varphi & \theta + \delta_T \end{bmatrix}$$

$$V^{-1} = \begin{bmatrix} \dfrac{\theta + \delta_T}{(\varphi + \delta_I)(\theta + \delta_T) - \varphi\theta} & \dfrac{\theta}{(\varphi + \delta_I)(\theta + \delta_T) - \varphi\theta} \\ \dfrac{\varphi}{(\varphi + \delta_I)(\theta + \delta_T) - \varphi\theta} & \dfrac{\varphi + \delta_I}{(\varphi + \delta_I)(\theta + \delta_T) - \varphi\theta} \end{bmatrix}$$

Example: Treatment compliance

$$f(x) = \begin{bmatrix} \mathrm{S}(\beta_I I + \beta_T T) \\ 0 \end{bmatrix} \qquad \mathrm{F} = \begin{bmatrix} \beta_I & \beta_T \\ 0 & 0 \end{bmatrix}$$

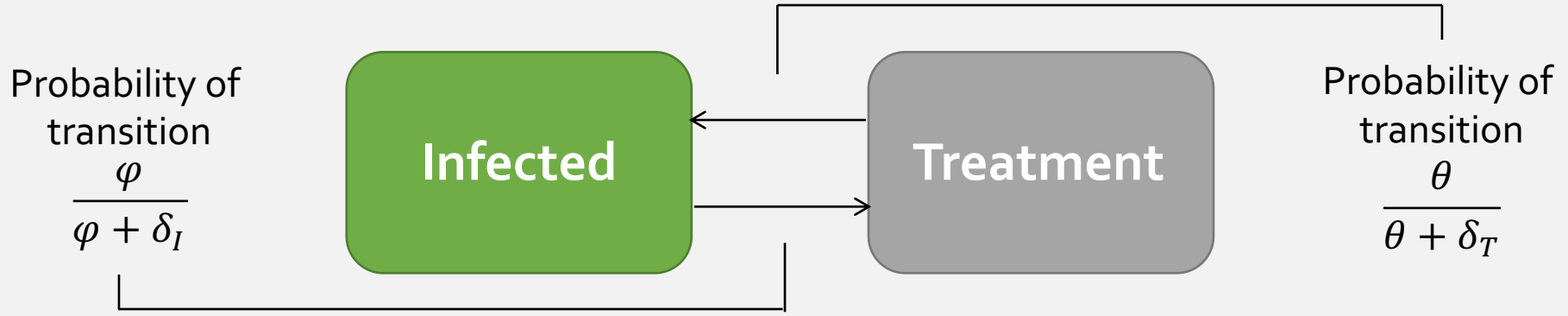
$$\mathrm{v}(x) = \begin{bmatrix} (\varphi + \delta_I) I - \theta T \\ (\theta + \delta_T) T - \varphi \mathrm{I} \end{bmatrix} \qquad \mathrm{V} = \begin{bmatrix} \varphi + \delta_I & -\theta \\ -\varphi & \theta + \delta_T \end{bmatrix}$$

$$V^{-1} = \begin{bmatrix} \dfrac{\theta + \delta_T}{(\varphi + \delta_I)(\theta + \delta_T) - \varphi\theta} & \dfrac{\theta}{(\varphi + \delta_I)(\theta + \delta_T) - \varphi\theta} \\ \dfrac{\varphi}{(\varphi + \delta_I)(\theta + \delta_T) - \varphi\theta} & \dfrac{\varphi + \delta_I}{(\varphi + \delta_I)(\theta + \delta_T) - \varphi\theta} \end{bmatrix}$$

What? These are average times spent in the compartments? How do we interpret $(\varphi + \delta_I)(\theta + \delta_T) - \varphi\theta$?

Expected number of visits

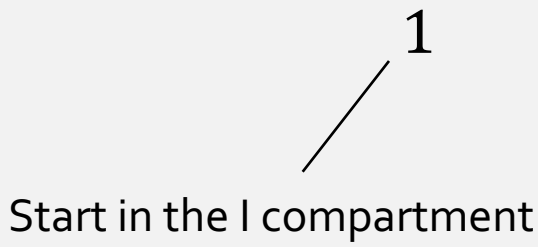
- How many times, on average, will one not be on treatment? To answer this, first ask, what is the probability that one relapses once they start treatment.



- The probability of jumping to treatment and back is $\frac{\varphi\theta}{(\varphi+\delta_I)(\theta+\delta_T)}$.

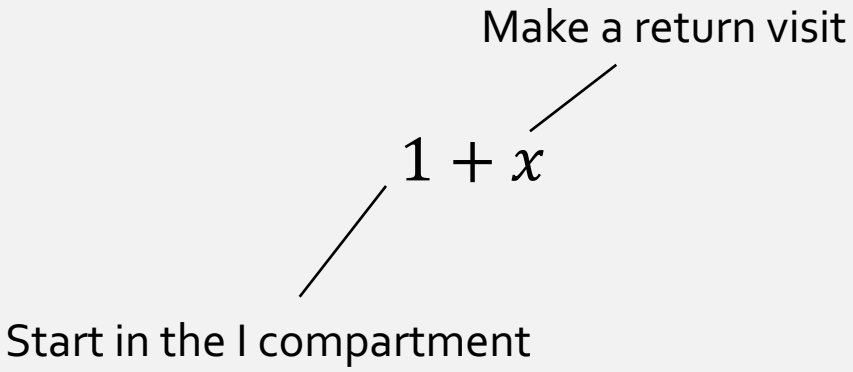
Expected number of visits

- Let us count the number of visits times the probability of making the visit. This will give us the expected number.
- First define $x := \frac{\varphi\theta}{(\varphi+\delta_I)(\theta+\delta_T)}$



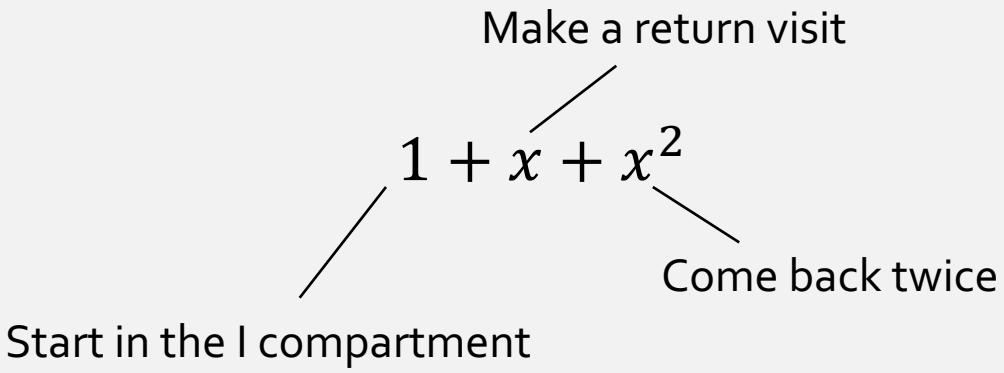
Expected number of visits

- Let us count the number of visits times the probability of making the visit. This will give us the expected number.
- First define $x := \frac{\varphi\theta}{(\varphi+\delta_I)(\theta+\delta_T)}$



Expected number of visits

- Let us count the number of visits times the probability of making the visit. This will give us the expected number.
- First define $x := \frac{\varphi\theta}{(\varphi+\delta_I)(\theta+\delta_T)}$



Expected number of visits

- Let us count the number of visits times the probability of making the visit. This will give us the expected number.
- First define $x := \frac{\varphi\theta}{(\varphi+\delta_I)(\theta+\delta_T)}$

Make a return visit

$$1 + x + x^2 + x^3 + \cdots = \frac{1}{1-x}$$

Come back twice

Start in the I compartment

Aside

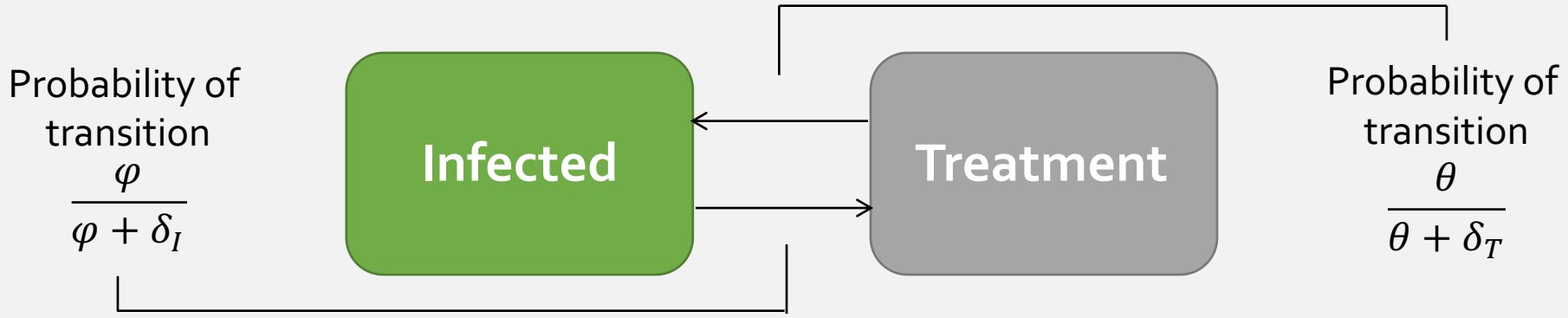
- Is it reasonable that probability of relapse is the same every time one goes on treatment?
- Well, whether it is or not, it’s tacitly baked into the model assumptions.
- As always, it’s very important to understand the tacit assumptions of your model.

Expected number of visits

- By our formula, the expected number of visits to I is
$$\frac{1}{1-x} = \frac{1}{1-\frac{\varphi\theta}{(\varphi+\delta_I)(\theta+\delta_T)}} = \frac{(\varphi+\delta_I)(\theta+\delta_T)}{(\varphi+\delta_I)(\theta+\delta_T)-\varphi\theta}$$
- Each visit to I lasts, on average, $\frac{1}{\varphi+\delta_I}$.
- So, one expects to spend
$$\frac{(\varphi+\delta_I)(\theta+\delta_T)}{(\varphi+\delta_I)(\theta+\delta_T)-\varphi\theta} \times \frac{1}{\varphi+\delta_I} = \frac{\theta+\delta_T}{(\varphi+\delta_I)(\theta+\delta_T)-\varphi\theta}$$
much time in the compartment over one's infectious lifetime.
- This is ${V^{-1}}_{1,1}$. Other entries can be derived similarly.

Graph-theoretic interpretation

- Write the adjacency matrix A of the weighted, directed graph of the infected compartments such that $A_{m,n}$ is the probability of moving from compartment n to compartment m



$$\mathrm{A} = \begin{bmatrix} 0 & \frac{\theta}{\theta + \delta_T} \\ \frac{\varphi}{\varphi + \delta_I} & 0 \end{bmatrix}$$

Graph-theoretic interpretation

- Then

$$M = I + A + A^2 + A^3 + \cdots = (I - A)^{-1}$$

is the matrix who (i, j) entry is the expected number of visits to compartment i if you start in compartment j.

- So, we can write V^{-1} as the product of waiting times and this matrix of expected visits

Graph-theoretic interpretation

Average time spent in a visit to I

Probability of going T to I

Waiting times matrix

Expected visits matrix

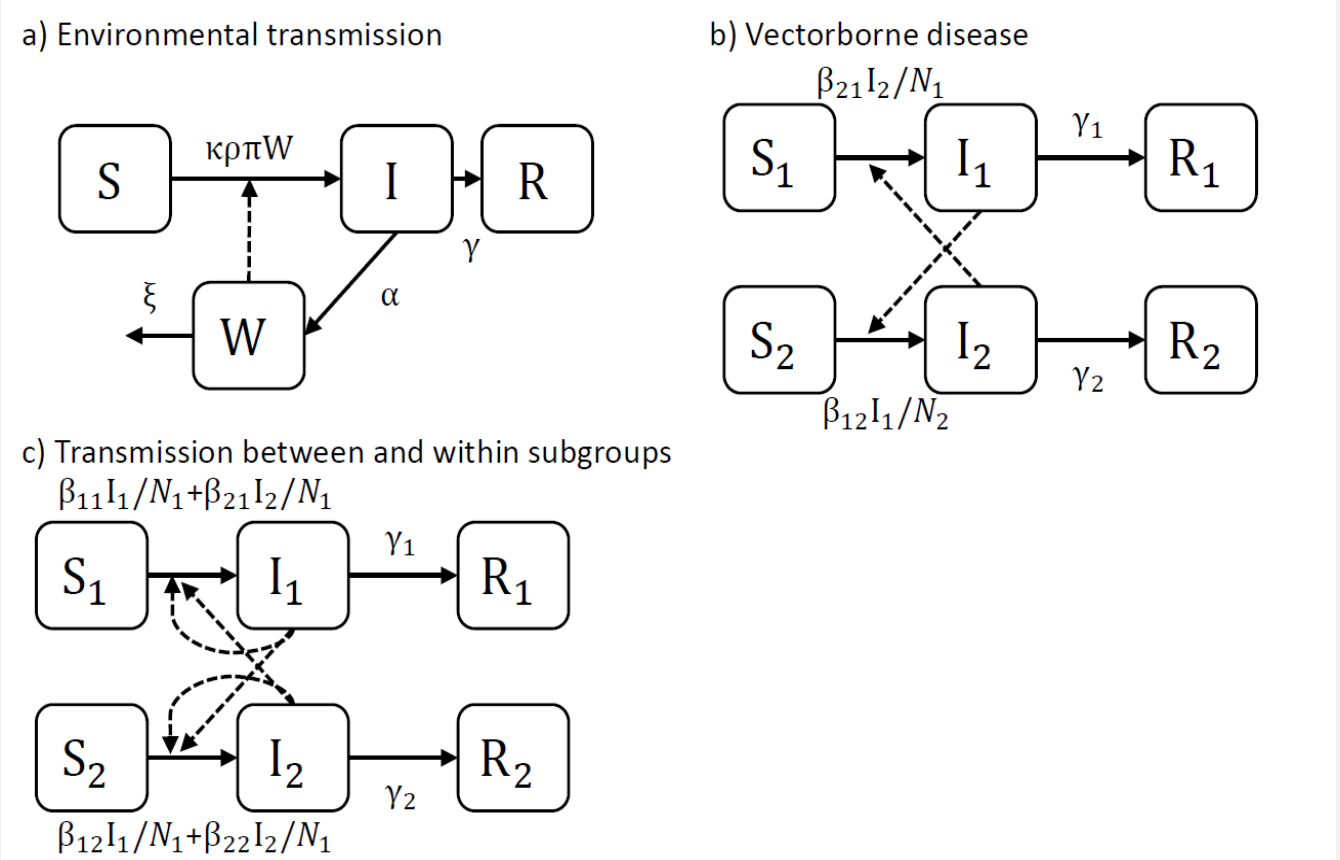
$$V^{-1} = \begin{bmatrix} \frac{1}{\varphi + \delta_I} & 0 \\ 0 & \frac{1}{\theta + \delta_T} \end{bmatrix} \times \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & \frac{\theta}{\theta + \delta_T} \\ \frac{\varphi}{\varphi + \delta_I} & 0 \end{bmatrix} \right)^{-1}$$

Average time spent in a visit to T

Probability of going I to T

Exercises

- Environmental transmission
- Vectorborne transmission
- Between- and within-subgroup transmission



Take-aways and suggestions

- A geometric interpretation of the next generation matrix helps explain why the basic reproduction is the spectral radius
- Stepping through interpretations for the entries of F, V^{-1}, and $\mathrm{F}V^{-1}$ can help make understanding R_0 easier
- R_0 depends on what we decide counts as a new infection. This is an epidemiological not mathematical decision.
- Calculating and interpreting R_0 for simpler models than your target model can build intuition

Further reading

R0 and the next generation matrix

- O. Diekmann, J. A. P. Heesterbeek, J. A. J. Metz, On the definition and the computation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations, Journal of Mathematical Biology 28 (4) (1990) 365-382.
- O. Diekmann, J. A. P. Heesterbeek, Mathematical epidemiology of infectious diseases: model building, analysis and interpretation, John Wiley & Sons, 2000.
- P. Van Den Driessche, J. Watmough, Reproduction numbers and sub- threshold endemic equilibria for compartmental models of disease transmission., Mathematical Biosciences 180 (2002) 29-48.
- O. Diekmann, J. A. P. Heesterbeek, M. G. Roberts, The construction of next-generation matrices for compartmental epidemic models, Journal of the Royal Society, Interface 7 (47) (2010) 873{885.

Further reading

General reading

- R. M. Anderson, R. M. May, Infectious Diseases of Humans: Dynamics and Control, Oxford University Press, Oxford, 1992.
- H. W. H. Hethcote, The mathematics of infectious diseases, SIAM review 365 42 (4) (2000) 599-653
- M. J. Keeling, P. Rohani, Modeling infectious diseases in humans and animals, Princeton University Press, 2011.
- A. J. Kucharski, The Rules of Contagion, Prole Books Ltd, London, 2020.

Type and target reproduction numbers

- M. G. Roberts, J. A. P. Heesterbeek, A new method for estimating the eort required to control an infectious disease, Proceedings of the Royal Society of London. Series B: Biological Sciences 270 (1522) (2003) 1359-64.
- Z. Shuai, J. A. P. Heesterbeek, P. van den Driessche, Extending the type reproduction number to infectious disease control targeting contacts between types, Journal of Mathematical Biology 67 (5) (2013) 1067{82.