

1. Consider the simplified Kermack-McKendrick model:

$$\frac{dS}{dt} = -\sigma SI, \quad \frac{dI}{dt} = \sigma SI - \gamma I, \quad (1)$$

with  $S(0) > 0, I(0) > 0, \sigma > 0, \gamma > 0$ .

Define  $\mathcal{R}(S) = \frac{\sigma S}{\gamma}$  and note that  $\mathcal{R}_0 = \mathcal{R}(S(0))$  is the basic reproductive number of the model.

Suppose that  $\mathcal{R}_0 > 1$ , i.e., there is an outbreak of the epidemic. Show that the density of infectives reaches the maximum value exactly at time  $t'$  where  $t' > 0$  is the unique real number such that  $\mathcal{R}(S(t')) = 1$ . Can you see why this should make biological sense? Write out the explicit formula for the maximum density of infectives.

We can think of  $\mathcal{R}(S)$  as the expected number of secondary cases produced by one infective individual if the density of susceptible individuals is fixed at  $S$ . Hence the disease cannot spread if  $\mathcal{R}(S) \leq 1$ .

Recall that  $S(t)$  is a strictly decreasing function of  $t$  and that  $S_\infty < \frac{\gamma}{\sigma}$ , here  $S_\infty := \lim_{t \rightarrow \infty} S(t)$  is the final size of the epidemic. Hence  $\mathcal{R}(S(t))$  is a strictly decreasing function of  $t$ ,  $\mathcal{R}(S(0)) = \mathcal{R}_0 > 1$ , and  $\mathcal{R}(S_\infty) = \frac{\sigma S_\infty}{\gamma} < 1$ . Therefore, there is a unique  $t' > 0$  such that  $\mathcal{R}(S(t')) = 1$ .

Since the function  $V(S, I) = S + I - \frac{\gamma}{\sigma} \log S$  is constant along the solution of (1), the solution lies on the curve

$$I = -S + \frac{\gamma}{\sigma} \log S + C$$

in  $(S, I)$  plane, where  $C = S(0) + I(0) - \frac{\gamma}{\sigma} \log S(0)$ . The function  $S \mapsto -S + \frac{\gamma}{\sigma} \log S + C$  has a unique maximum when  $S = \frac{\gamma}{\sigma}$  (why?), and this is equivalent to  $\mathcal{R}(S) = 1$ , i.e., the maximum happens at the time  $t'$  such that  $\mathcal{R}(S(t')) = 1$ . We have

$$\begin{aligned} I_{\max} &= -\frac{\gamma}{\sigma} + \frac{\gamma}{\sigma} \log \frac{\gamma}{\sigma} + C \\ &= -\frac{\gamma}{\sigma} + \frac{\gamma}{\sigma} \log \frac{\gamma}{\sigma} + S(0) + I(0) - \frac{\gamma}{\sigma} \log S(0) \\ &= S(0) - \frac{\gamma}{\sigma} + \frac{\gamma}{\sigma} \log \left( \frac{\gamma}{\sigma S(0)} \right) + I(0) \\ &= S(0) \left( 1 - \frac{1}{\mathcal{R}_0} + \frac{1}{\mathcal{R}_0} \log \left( \frac{1}{\mathcal{R}_0} \right) \right) + I(0). \end{aligned}$$

2. [an SI epidemic model] Consider the simplified Kermack-McKendrick model, as above. If we assume that  $\gamma = 0$ , so no individual can move out of the  $I$  class. Then once an individual is in the  $I$  class (is infected), it will remain capable of spreading the disease for all future time. The reduced model is then:

$$\begin{aligned} \frac{dS}{dt} &= -\sigma SI, \\ \frac{dI}{dt} &= \sigma SI, \end{aligned} \quad (2)$$

with  $S(0) \geq 0, I(0) \geq 0, \sigma > 0$ . Show that:

(a) There is a unique non-negative solution of (2) defined on  $t \in [0, \infty)$ .

We can see that the vector field  $f(S, I) = (-\sigma SI, \sigma SI)$  is a continuously differentiable function of  $(S, I) \in \mathbb{R}^2$ . Hence there exists a unique solution to the initial value problem (2) defined on some maximal interval  $t \in [0, b)$ . Since  $\dot{S}(t) + \dot{I}(t) = 0$ , and so  $S(t) + I(t) = S(0) + I(0)$  for all  $t \in [0, b)$ , we have  $b = \infty$ . Since  $S(t) = S(0)e^{-\sigma I(t)}$  and  $I(t) = I(0)e^{-\sigma S(t)}$ , we have  $S(t), I(t) \geq 0$  for all  $t \in [0, \infty)$ .

- (b) (1) If  $I(0) = 0$ , then  $I(t) = 0$  and  $S(t) = S(0)$  for all  $t \in [0, \infty)$ .

The above is clearly a solution of (2), and the solution is unique.

- (2) If  $S(0) = 0$ , then  $I(t) = I(0)$  and  $S(t) = 0$  for all  $t \in [0, \infty)$ .

Same as (1).

- (3) If  $S(0) > 0$  and  $I(0) > 0$ , then  $S(t)$  decreases monotonically toward 0 as  $t \rightarrow \infty$  and  $I(t)$  increases monotonically toward  $S(0) + I(0)$  as  $t \rightarrow \infty$ . This means that every individual in the population eventually becomes infective.

If  $S(0), I(0) > 0$ , then we must have  $S(t), I(t) > 0$  for all  $t \in [0, \infty)$  (why?). Since  $\dot{S}(t) = -\sigma S(t)I(t) < 0$  and  $\dot{I}(t) = \sigma S(t)I(t) > 0$ ,  $S(t)$  is a strictly decreasing function of  $t$  and  $I(t)$  is a strictly increasing function of  $t$ . Let  $S_\infty := \lim_{t \rightarrow \infty} S(t)$ . Suppose, to get a contradiction, that  $S_\infty > 0$ . Then  $\dot{I}(t) = \sigma S(t)I(t) \geq \sigma S_\infty I(t)$ , and so  $I(t) \geq I(0)e^{\sigma S_\infty t}$ . This implies that  $I(t) \rightarrow \infty$  as  $t \rightarrow \infty$ , but this is a contradiction since  $0 < I(t) \leq S(t) + I(t) = S(0) + I(0)$ . Hence  $S_\infty = 0$ , i.e.,  $S(t) \rightarrow 0$  as  $t \rightarrow \infty$ . Since  $S(t) + I(t) = S(0) + I(0)$  for all  $t \in [0, \infty)$ , we must have  $\lim_{t \rightarrow \infty} I(t) = S(0) + I(0)$ .

Incidentally, if we let  $N = S(t) + I(t)$ , then we can rewrite the differential for  $I$  as

$$\frac{dI}{dt} = \sigma(N - I)I.$$

This scalar differential equation can be solved explicitly (by, e.g., separation of variables technique), and we have

$$I(t) = \frac{I(0)N}{I(0) + (N - I(0))e^{-\sigma Nt}}.$$

- (c) Do you think this is a good model? Why or why not?

This is an open-ended question. We may argue that the assumption that infectives stay infected forever is not a good assumption, since every individual must die eventually. Also, on the longer time scale, we have to take birth and death of the population into account. Recall that we choose to ignore birth and (non-disease related) death in the simplified Kermack-McKendrick model because of the assumption that the epidemic dies out within a short period of time (the mean infectious period is  $1/\gamma$  for the simplified Kermack-McKendrick model).

3. Recall the equation for the final size,  $S_\infty$ , of the simplified Kermack-McKendrick model,

$$S_\infty - \frac{\gamma}{\sigma} \log S_\infty = S(0) + I(0) - \frac{\gamma}{\sigma} \log S(0). \quad (3)$$

Suppose  $S(0), I(0), \sigma, \gamma > 0$  are given. Show that there are two values of  $S_\infty > 0$  that satisfies (3), and that the final size is the smaller of the two.

Let  $f(x) = x - \frac{\gamma}{\sigma} \log x$ ,  $x > 0$ . Then  $f'(x) = 1 - \frac{\gamma}{\sigma x}$  and  $f''(x) = \frac{\gamma}{\sigma x^2}$ . Hence  $f(x)$  is concave upward and has a unique minimum at  $x_{\min} = \frac{\gamma}{\sigma}$ . Let  $y_{\min} = f(x_{\min})$ . It also follows from the concavity of  $f(x)$  that, for every  $y > y_{\min}$ , the equation  $f(x) = y$  always has exactly two solutions  $x_1$  and  $x_2$  with  $0 < x_1 < x_{\min} < x_2$ .

Since  $S(0) + I(0) - \frac{\gamma}{\sigma} \log S(0) > S(0) - \frac{\gamma}{\sigma} \log S(0) \geq y_{\min}$  (why?), there are two  $S_\infty > 0$  that satisfies (3), with one being strictly smaller than  $\frac{\gamma}{\sigma}$  and the other one being strictly larger than  $S(0)$ . But since the final size must be less than  $\frac{\gamma}{\sigma}$  (and also since  $S_\infty$  must be less than  $S(0)$ ), it has to be the smaller of the two.

4. There was an outbreak of an infectious disease in a city of 1 million inhabitants. The average length of infectious period of the disease was 7 days. At the end of the epidemic, there were 70,000 people that didn't contact the disease. Assume that the outbreak can be describe by the simplified Kermack-McKendrick model, and that the initial number of infectives was so small that it can be neglected. Find the basic reproductive number,  $\mathcal{R}_0$ , for this outbreak.

From the final size equation (3) with  $S(0) = 1000000$ ,  $S_\infty = 70000$ , and  $I(0) = 0$ , we have

$$70000 - \frac{\gamma}{\sigma} \log 70000 = 1000000 - \frac{\gamma}{\sigma} \log 1000000.$$

This implies  $\frac{\gamma}{\sigma} \approx 349721$ , and so  $\mathcal{R}_0 = \frac{\sigma S(0)}{\gamma} \approx 2.86$ . Note that we do not really need the information about the average length of infectious period.