**Instructions:**

- Please turn in a single PDF file.
- Please share a link to your code, but do not attach the actual code.
- Handwritten math (scanned and included in a PDF) is fine, but please watch out for 10MB+ file sizes!

1. The goal of this problem is to develop flexibility with your Forward Euler code, and to learn a bit about the effect of step size on the accuracy of the solution.

   a. Using your Forward Euler method, simulate the solution to the *normalized* SIS model discussed in class (Week 3) using $\beta = 3$ and $\gamma = 2$, and with $(s_0, i_0) = (0.99, 0.01)$. Create three plots ranging from $t = 0$ to $t = 25$. On the first, simulate using a step size $\Delta t = 2$. On the second, use $\Delta t = 1$. On the third, use $\Delta t = \frac{1}{2}$. In each plot, show only your solution's $I(t)$ in a red solid line, labeled as "Forward Euler", and then also plot the analytical solution from class in a black dashed line, labeled as "Analytical." Please also set the y-axis range to $[0, 0.5]$.

   Solution: These are the plots for the given step sizes:
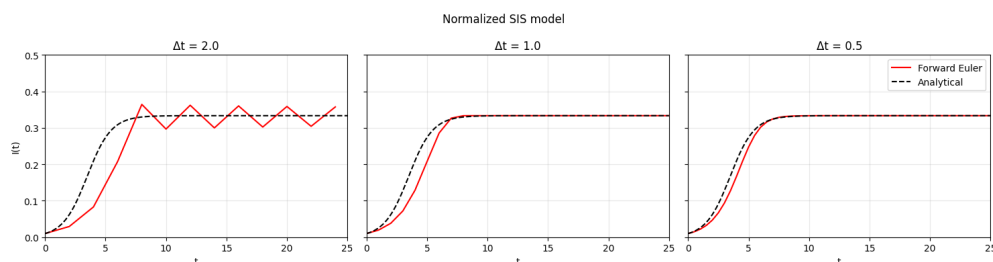
   b. Comment on what you see in your three plots. How does the step size affect our solution?

   Solution: What I notice here is that the smaller the step size, the more accurate the simulation is. The largest step size of 2 showed a curve with much more variance than the analytical solution. With each decrease in step size, the curves converge more and more towards each other.

   c. Define the maximum absolute error for a simulation using a particular Δt as

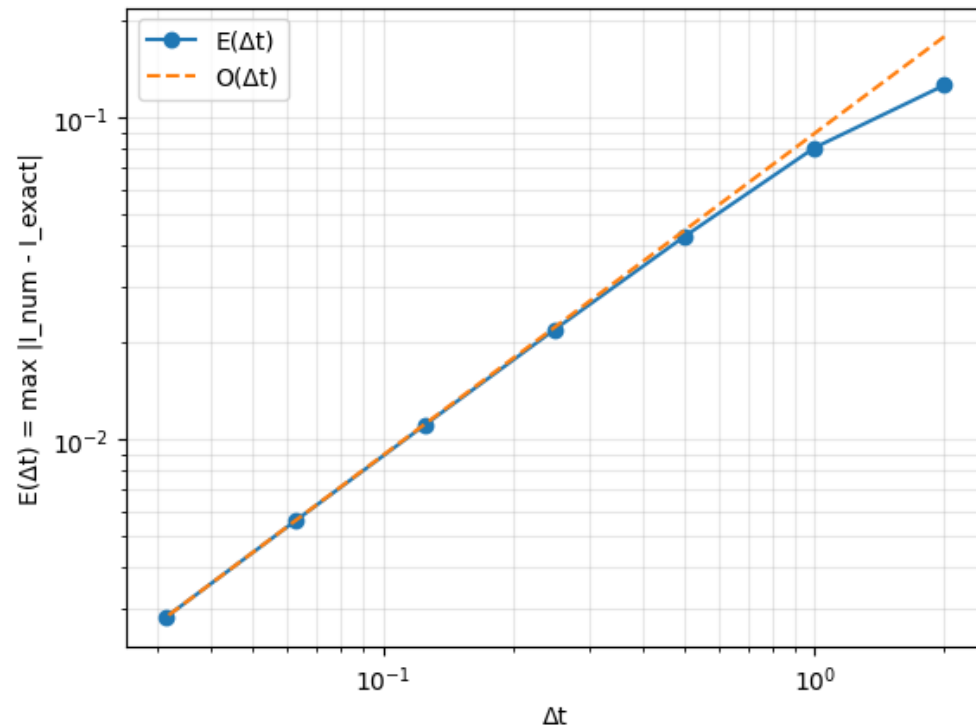 $$E(\Delta t) = \max_t \left| I_{\text{Euler}, \Delta t}(t) - I_{\text{analytical}}(t) \right|.$$

   Write a function that runs the appropriate simulation, computes the analytical solution, and returns E without plotting. Share a link to your code for this problem.

   Solution: My code for problem 1C

   d. Create a plot on log-log axes showing $E(\Delta t)$ vs Δt for values

   $$\Delta t \in \{2, 1, \tfrac{1}{2}, \tfrac{1}{4}, \tfrac{1}{8}, \tfrac{1}{16}, \tfrac{1}{32}\}$$

Solution:

e. Comment on what you observe in this plot, and comment on cases when you would want a larger or smaller step size, and why? Imagining yourself in an advisory position in your community, can you think of any scenario where there is a connection between the step size of your simulation and the ethics of your advice?

Solution: The maximum absolute error $E(\Delta t)$ tends to decrease as the step size Δt decreases. This further supports what I saw earlier that smaller step sizes lead to more accurate simulations.

2. The goal of this problem is to get you thinking about the constraints on population contact structure and contact matrices, as well as sensitivity analyses.

   a. As one who is interested in modeling disease transmission on college campuses, you hire two teams to measure contact patterns on a nearby campus. The first team, led by Dan Pemic, tells you that there are 200 faculty and 1800 students, with a contact matrix of

$$C_{\text{Pemic}} = \begin{pmatrix} 3.1 & 43.5 \\ 4.7 & 25.0 \end{pmatrix}$$

   The second team, led by Flynn Uenza, tells you that there are 210 faculty and 1750 students, with a contact matrix of

$$C_{\text{Uenza}} = \begin{pmatrix} 3.0 & 44.5 \\ 4.8 & 25.1 \end{pmatrix}$$

   Whom do you trust more, Dan Pemic or Flynn Uenza? To answer this question, consider the self-consistency (or lack thereof) of each dataset. Explain your reasoning in words and include any calculations used to arrive at your conclusions.

   Solution: This can be checked by seeing if the number of contacts from faculty to students is about equal to the number of contacts from students to faculty. Mathematically this is as follows:

$$n_F c_{F \to S} \approx n_S c_{S \to F}$$

   Pemic:
   $n_F = 200$
   $n_S = 1800.$

   Totals: $200 \cdot 43.5 = 8700$ vs. $1800 \cdot 4.7 = 8460.$

   Mismatch $= 8700/8460 - 1 \approx 2.83\%.$

   Uenza:
   $n_F = 210$
   $n_S = 1750.$

   Totals: $210 \cdot 44.5 = 9345$ vs. $1750 \cdot 4.8 = 8400.$

   Mismatch $= 9345/8400 - 1 \approx 11.25\%.$

   Given this information, and comparing the error rates, I would say that Pemic's data is more self-consistent and by extension more trustworthy as well.

3

b. A straightforward fix to self-consistency issues is to "symmetrize" the rates. First, we compute the implied total number of intergroup contacts from faculty to students, and then compute the same from students to faculty. After averaging those two counts, divide by the appropriate population size to get per-person rates. Use this approach to symmetrize the two contact matrices.

Solution: Symmetrize only off-diagonals using:

$$A = \frac{n_F c_{F \to S} + n_S c_{S \to F}}{2}$$

Then:

$$c'_{F \to S} = \frac{A}{n_F}$$

$$c'_{S \to F} = \frac{A}{n_S}$$

Keep diagonals unchanged.

Pemic:

$$A = \frac{200 \cdot 43.5 + 1800 \cdot 4.7}{2} = \frac{8700 + 8460}{2} = 8580$$

Then:

$$c'_{F \to S} = \frac{8580}{200} = 42.9$$

$$c'_{S \to F} = \frac{8580}{1800} = 4.766\bar{6}$$

$$C^{\text{sym}}_{\text{Pemic}} = \begin{pmatrix} 3.1 & 42.9 \\ 4.7667 & 25.0 \end{pmatrix}$$

Uenza:

$$A = \frac{210 \cdot 44.5 + 1750 \cdot 4.8}{2} = \frac{9345 + 8400}{2} = 8872.5$$

Then:

$$c'_{F \to S} = \frac{8872.5}{210} = 42.25$$

$$c'_{S \to F} = \frac{8872.5}{1750} = 5.07$$

$$C^{\text{sym}}_{\text{Uenza}} = \begin{pmatrix} 3.0 & 42.25 \\ 5.07 & 25.1 \end{pmatrix}$$

These now are the symmetrized contact matrices.

c. How different are these two symmetrized matrices, really? Answer the question by computing the ratio of R_0 under Pemic's data to R_0 under Uenza's data, assuming SIR models with otherwise identical parameters.

Solution: With identical parameters, R_0 is proportional to the dominant eigenvalue in the contact matrix. In other words, R_0 is the ratio of the two dominant eigenvalues of the two symmetrized matrices.

Given a 2×2 matrix:

$$\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix} \right)$$

The dominant eigenvalue is:

$$\lambda_{\text{max}} = \frac{(a+d) + \sqrt{(a-d)^2 + 4bc}}{2}$$

Plugging in values for Pemic:

$$a = 3.1$$
$$b = 42.9$$
$$c = 4.7667$$
$$d = 25.0$$
$$\lambda_{\text{max}} \approx 32.061$$

Doing the same for Uenza:

$$a = 3.0$$
$$b = 42.25$$
$$c = 5.07$$
$$d = 25.1$$
$$\lambda_{\text{max}} \approx 32.389$$

.

Therefore:

$$\frac{R_{0:\text{Pemic}}}{R_{0:\text{Uenza}}} \approx \frac{32.061}{32.389} \approx 0.990$$

So the two R_0 values are very close to each other. They only differ by about 1%.

3. (Grad / EC): The goal of this problem is to get you to think about additional flavors of models that build on the SIR model backbone, and practice writing down flow diagrams and systems of differential equations. For each of the following situations please (i) draw a flow diagram with the SIR backbone in black, (ii) include any modifications in a second color, and (iii) write out your differential equations using the same color scheme.

   (a) Suppose that we want to model the possibility that the natural history of infection means that a person is infected *but not infectious* before becoming infected-and-infectious. Let the typical *latent* period—the time between exposure and infectiousness—last for q days. Use the letter E for this new compartment. Draw a flow diagram for the non-normalized system, and write a set of corresponding differential equations.

   (b) Suppose that we want to model Hospitalization, with the following assumption: Infected folks either recover directly *or* they are hospitalized first and then recover. Let the direct recovery rate be γ, and suppose that there are 4 direct recoveries for every 1 hospitalization. Let the typical duration of a hospitalization be δ days. Use the letter h for the hospitalized compartment. Draw a flow diagram for the normalized system, and write a set of corresponding differential equations. You should assume that folks in the hospital do not come into contact with anyone else during their hospital stay.

   (c) Suppose that we want to model an infectious disease that afflicts a growing population of bacteria, such that the infection follows an SIR model, the bacteria grow according to a logistic growth model with intrinsic growth rate α and carrying capacity K. Let all bacteria reproduce, but suppose that susceptible and recovered bacteria produce susceptible progeny, while infected bacteria produce infected progeny. The carrying capacity is shared by all three bacteria.