Appendix D

Computational Projects

These projects are designed around the computational implementation of one or more mathematical models and the use of that implementation to explore the properties of the model, and are intended to be done individually, rather than in groups. The projects should be treated as a computational experiments, with results written in the style of a lab report that is aimed at a third party with little background in the subject. The results thus need to be fully explained with such an audience in mind. A recommended structure would include:

Title: Project title, author and date.

Abstract: A *brief* summary of the key questions, results and conclusions. In general this should not be longer than a half page, double spaced.

Introduction: A description of the problem(s) under consideration, and any relevant background information. This section should include both a summary of the biological problem(s) being considered, as well as the mathematical modeling frameworks that will be used.

Results and Discussion: An integrated presentation of the results of each "experiment" and any relevant discussion. This should be written in narrative prose (that is, complete sentences that walk the reader through everything that was done), with any generated graphics and code inserted at the appropriate location.

Conclusion: A brief summary of the key insights obtained, and perhaps any unanswered questions that remain.

References: A list of all resources used in completion of the project. If a resource is used in generating a specific result or in support of specific point made in the introduction or discussion, this should be noted with a citation at the appropriate place in the report text. A specific format for citations/references is not required, but the format used should be consistent.

Note that the projects were created with the use of MATLAB in mind, and in my classes this is expected. However, the projects are not dependent on this choice, and could be handled equally well with other programming languages.

D.4 Making predictions with a simple SIR model

Viral respiratory infections are a particularly common class of infectious disease that jump between species relatively often, leading to epidemics in populations with minimal immunity to the novel form. COVID-19 was the most recent of these. With this class of infectious agent, individuals generally develop some degree of immunity after an infection (although the immunity may wane over time), and thus is it common to model these using an SIR model, schematically described by:

$$S \xrightarrow{\beta} I \xrightarrow{\gamma} R$$

with a difference equation model written as:

$$\begin{array}{rclcrcl} \hat{S}_{i+1} & = & \hat{S}_i & - & \beta\hat{S}\hat{I} & & \\ \hat{I}_{i+1} & = & \hat{I}_i & + & \beta\hat{S}\hat{I} & - & \gamma\hat{I} \\ \hat{R}_{i+1} & = & \hat{R}_i & & & + & \gamma\hat{I} \end{array}$$

Here, $\hat{S}$, $\hat{I}$ and $\hat{R}$ refer to fractional populations, which are most commonly reported in the epidemiological literature in units of *cases per 100,000*. β is a parameter that describes the likelihood of new infections, and γ is the rate of recovery. We can express β and γ in terms of the basic reproductive factor, r_0, and the average length of infectivity, τ, as: $\gamma = \frac{1}{\tau}$ and $\beta = \frac{r_0}{\tau}$. For COVID-19, estimates early in the pandemic suggested a basic reproductive factor of about 3.0, with an average length of infectivity of about 10 days.

- Implement an SIR model of COVID-19 in Matlab, using parameters consistent with our understanding early during the COVID-19 pandemic. Run a simulation of the expected disease dynamics if 4 infected individuals were to arrive in a city of 10 million with no prior immunity, and plot $\hat{S}$, $\hat{I}$ and $\hat{R}$ as a function of time on a single graph, and address the following:
  - When does the peak (highest total number of currently infected individuals) occur (relative to the arrival of the first infected individuals), and how many active cases are there per 100,000 people at the peak?
  - When does the highest number of *new daily infections* occur, and and how many new daily cases are there per 100,000 people at this time?
  - By the end of the epidemic, what fraction of the population has been infected?
  - If 10% of cases require in-hospital care, how many hospital beds per 100,000 people would be needed at peak demand, and how many daily hospital admissions (per 100,000) would be expected at peak growth?
  - If the mortality of the illness was 1.5% (consistent with early data for COVID-19), how many people in the city would have died by the end of the epidemic?
- The global average annual mortality rate is about 837 per 100,000 (and the mortality rate in the US is very close to the global average). The US has about 270 hospital beds per 100,000, quite similar to the number in Canada and the UK); Japan and South Korea have about 1270 beds per 100,000 and some developing countries have

as few as 50 beds per 100,000. Given these reference values, discuss why the results from your simulation would lead to major worries about COVID-19 spreading to new places? Keep in mind that mortality rates for an infectious diseases typically assuming adequate treatment, and increase dramatically if there are medical facilities can not meet demand.

- Repeat your simulation three times: (i) with the basic reproductive factor (r_0) doubled; (ii) with the average infective time (τ) halved; and (iii) with the initial number of cases increased to 400 (two orders of magnitude). For each case, plot $\hat{S}$, $\hat{I}$ and $\hat{R}$ as a function of time. Briefly explain how r_0, τ and I_0 seem to affect the course of the epidemic.

- Many responses to managing a potential epidemic effectively reduce the basic reproductive number. Repeat the simulation three more times, now with (i) $r_0 = 1.5$ (reduced by a factor of 2); (ii) $r_0 = 1.0$ (reduced by a factor of 3); (iii) $r_0 = .5$ (reduced by a factor of 6). For each case, plot $\hat{I}$ and $\hat{R}$ as a function of time, omitting $\hat{S}$ for clarity. Discuss how these simulations differ from each other and from what was seen above. How helpful would each factor of reduction of r_0 be in reducing the impact of the infectious disease?

- Vaccination (or other sources of prior immunity) can be incredibly effective in helping to prevent infectious disease, but getting a large fraction of the population immunized for a novel infectious disease is a challenge. Repeat your simulation to consider the cases where (i) 25%, (ii) 50%; or (iii) 75% of the population is already immune (whether due to vaccination or past exposure to a similar virus). Discuss what you observe, including an assessment of which prior immunity rates seem to provide "herd immunity", in which no epidemic occurs.

- A minimal vaccination level required for herd immunity can be analytically derived from the SIR model; show this derivation and calculate this thresh hold for the given model parameters. Are your simulation results consistent with this?

- Calculate the minimal vaccination level required for herd immunity for a modestly reduced r_0 of 1.5, and discuss what this suggests about the possible synergy of responses strategies. In particular, consider how strategies that may not be sufficient to prevent an epidemic individually may be able to do so when used in combination.

