Case Study II: Dynamical Networks and Control *(DUE: MONDAY, NOV 8 @ 4:00PM)*

Textbook Reading: Chapter 9

Introduction

In this case study you will use the skills and methods you have learned in linear algebra and MATLAB to simulate dynamical systems and explore strategies to control them. This case study will link to the ESE topic areas of dynamical systems modeling and control engineering.

A dynamical system model describes the evolution of a system as a function of time. Many dynamical systems models are based on fundamental laws of physics (e.g., a differential equation that describes the velocity of a car), whereas others that are more phenomenological (i.e., they are based on observations and/or high-level intuition). In this case study we will we working with phenomenological models of epidemic spread in populations and networks. Your goals are fourfold:

(i) Implement the SIRD model described in Section 9.3 of the book. This model describes the evolution of a disease within a population.

(ii) Fit this model to data from the ongoing COVID19 pandemic. Specifically, you will attempt to fit your model to disease progression within three populations: St. Louis, Jefferson City, and Springfield, MO. You will then make analytical inferences regarding how the disease is progressing differently between these regions.

(iii) Generalize the SIR model in two ways: (i) Adding a new state to account for individuals in lockdown, and (ii) to describe both within- and across- population dynamics. In other words, you will build a larger network model that attempts to describe the evolution of the disease in all three regions simultaneously, taking into account flow of individuals between regions.

(iv) You will then design an intervention strategy that can include any combination of lockdown, deployment of vaccines, or deployment of treatments. You will predict how your intervention strategy affects viral propagation and outcomes over 2021. You will submit a recommended strategy that will be used in a <u>competition</u> to find the best improvement in outcomes relative to the 'cost' of the intervention (more on this below).

Specific Instructions: Part (i) SIR Modeling

- Read about the SIR model in Section 9.3 of the book. Then, write MATLAB code/scripts that simulate this model *from scratch*. Explore how changing the parameters of the model affects the output. How does the model tend to converge over time? Produce plots that are similar to those depicted in the book. Use Matrix and Vector operations within your code!! Do not use pre-packaged MATLAB functions such as 'lsim' at this stage.

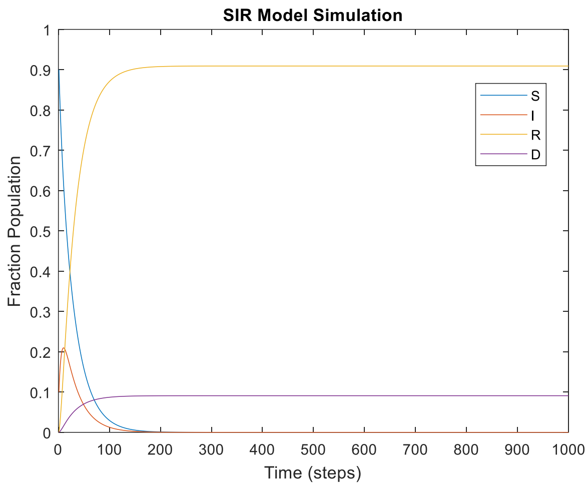

Figure 1: SIRD Model Simulation from Section 9.3.

Specific Instructions: Part (ii) Fitting model to COVID 19 Data

Familiarization with 'ss' and 'lsim'

Now, examine the script 'base_sir.m', which implements the model from Section 9.3 using the MATLAB function 'lsim'. Use MATLAB help to learn more about 'ss' and 'lsim' and what is happening in this script. Here are some key points:

- 'lsim' allows for simulation of a general class of linear dynamical systems, including ones that have external inputs.
- The script first sets up the model using the 'ss' function. It uses the specification of the 'A' matrix and sets a number of parameters. Read about the 'ss' function to learn more. Note especially that we are working in the space of <u>discrete-time models</u>, where time advances in steps (as opposed to continuous time models, which are defined in terms of differential equations).
- Experiment for yourself with base_sir.m and the ss and lsim functions. Compare the output of using these functions with the outcome of your from-scratch effort in Part (i).
- Complete the function 'siroutput_full', contained in the file 'siroutput_full.m' that is meant to take some model parameters and return a simulation output by using 'lsim'.

COVID Data Pull

Load the MATLAB file 'COVIDdata.mat', which contains COVID data for St. Louis, Jefferson City, and Springfield, MO. Familiarize yourself with this data (this is very similar to what you worked with in your previous MATLAB homework). In particular, the relevant data tables are:

COVID_MO – nx5 table containing cumulative COVID cases and deaths for the St. Louis, Jefferson City, and Springfield metropolitan areas.

populations_MO - population of each metro area

Using 'fmincon' to find a model to fit and explain actual COVID19 data

- Once you are comfortable with 'base_sir.m', you will move on to fit models to COVID19 data. This is an example of "parameter optimization". Here, your goal is to find values for the SIRD model rate-constants (i.e., infection rate, fatality rate, recovery rate) and initial conditions (initial values of susceptible, infected, recovered and deceased population fractions). For this, you will use the code scaffolded in 'base_sir_fit.m'. ****<u>Read the comments in that script very carefully!</u>**** In particular, this script uses fmincon (learn more about this function at the MATLAB help), which is a powerful MATLAB function that will find parameters that minimize a given "cost". Your task is to specify this cost via the function 'siroutput', that you must specify. (Note that this function is scaffolded for you in 'siroutput.m'. That file (once you have built it up!) must be in the same folder as base_sir_fit.m for the latter to run.

- Some suggestions when working to specify 'siroutput.m':
  - The 'cost' that you return in this function should depend on the error between the true data and the output of the model.
  - Use aspects of 'base_sir.m' as needed!
- Try fitting models to the St. Louis COVID data in the following different ways:
  - Many experts speak of different phases or waves of the disease. Find different models for different phases of propagation. In other words, try and identify time-points at which propagation dynamics appear to change. A coarse visual adjudication will suffice at this stage. Isolate those ranges of time and fit different models for different waves. Compare/contrast the different fits. Do your models fits speak to differences in viral propagation dynamics between these phases?

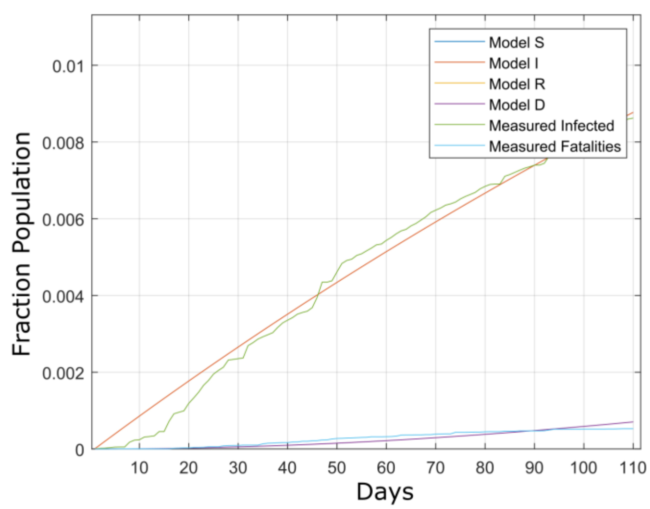

Figure 2: Example of a model fit to the first 110 days of St. Louis City COVID data.

Specific Instructions: Part (iii) Model generalization

During the course of the pandemic there have been two strategies that have had significant impact on viral propagation: lockdowns and vaccinations. In this part of the Case Study, you will generalize the model to be able to describe these scenarios. First, assume that vaccinated individuals can be aggregated into the recovered population. How would you modify the model to reflect vaccinations? Second, add a fifth state to the model to account for individuals in lockdown (let's call this state 'L' and the augmented model, with vaccinations and lockdowns, the SLIRD model). Repeat your model fitting and analysis, but for the SLIRD model applied to each of the 3 metropolitan areas separately. During 2021, do your model fits suggest any differences in vaccination dynamics between regions?

Now, repeat this analysis but for a larger 3 region network model (i.e., with St. Louis, Kansas City and Missouri integrated into a single model). This network model should contain "interaction/transport" terms that model the exchange of individuals between regions.

Specific Instructions: Part (iv) Model interpretation

Examine the model(s) you have obtained in Part (iii) above. What do your parameters tell you about similarities/differences in viral propagation dynamics between regions? Does your model predict and/or identify known historical events, such as the rise of the Delta variant, the adoption of more stringent distancing/masking requirements, etc. Are there any other interpretations you can glean from the models you have constructed?

Specific Instructions: Part (v) Intervention Design, Prediction and Competition

A key aspect of dynamical systems modeling in engineering is that it provides a way to analyze the underlying factors that govern a given time-varying phenomenon and to *predict* how the phenomenon would continue to evolve in the future.

Prediction:

Simulate your SLIRD model for St. Louis forward in time to predict how the pandemic may evolve in the future. Use a nominal end date 365 days into the future.

Policy design:

In your model, different public health measures (lockdowns, masking, vaccines, treatments, etc.) can be thought of in terms of increasing or decreasing specific entries/coefficients within the SLIRD transition matrix.

The goal of this last part of the case study is to provide <u>policy recommendations</u> using your model.

Specifically, assume that benefits are given by:

$$J_{benefit} = \ 10(\|\Delta I\|) + 10(\|\Delta D\|)$$

Here, ΔI and ΔD are vectors that denote the difference in infected and deceased population fractions between the baseline and controlled models. In other words, the i^{th} entry of ΔD is the difference between the fraction of the population that is deceased on day i in the baseline model and the fraction of the population that is deceased on day i in the model with intervention.

Further, suppose costs are given by

$$J_{costs} = 100(\|\Delta L\|^2) + \ 800(1-\gamma)(\|\Delta I\|^2) + 800(1-\beta)(\|\Delta D\|^2)$$

The parameters γ and β denote the *mean relative change* in the infection and fatality rates, respectively. For example, if your baseline model has an infection rate of 0.05 and your intervention model has an average infection rate of 0.04, then $\gamma = \frac{0.04}{0.05} = 0.8$. Thus, the more you 'dial up/down' these rates over time, the more cost is incurred. Note that the costs are quadratic in the norm of the difference vectors.

Now, you will construct a policy recommendation in terms of how to incentivize lockdowns/distancing, vaccines, and/or treatments in order to maximize benefits relative to costs. Your recommendation will include aspects such as: 1) how much upregulation in population lockdowns you wish to enact, 2) how much downregulation in infection rate you wish to enact, and 3) how much reduction in fatality rate you wish to enact. Note that downregulating these parameters should be accompanied with appropriate re-parameterization of the entire model for consistency (i.e., so that all columns in the SLIRD transmission matrix add to one).

The relative benefit of your policy can be assumed as:

$$J_{relative} = J_{benefit} - \alpha J_{cost} - Wobble$$

The coefficient α can be interpreted as a weighting on the cost (e.g., some cities may be more willing to tolerate lockdowns, etc.). The $Wobble$ of a policy as the extent to which is fluctuates over time. For example, you might imagine that constantly and rapidly dialing up and down coefficients would lead to irritation amongst citizens. We will define the $Wobble$ as the maximum standard deviation of all model coefficients over time.

Build your policy in the function ‘sirpolicy’. This shell of this function is provided, but the contents are entirely up to you. This function should take the current values of S, L, I, R, D, and the current model state transition matrix. The function should return a new model state transition matrix to govern the next time step evolution. If the new matrix and the current matrix are the same, then the policy is to maintain status quo.

Try and construct a *policy* that achieves a high relative benefit over the date range 1/1/2021-10/1/2021 in St. Louis City. Explore how your policy deals with different values of α, with the goal of finding one that performs well over a range of α values.

<u>*Make sure that your returned model is consistent*</u> (i.e., all columns sum to one and all rates are positive).

Be sure to discuss the interpretations of your policy (does it correspond to lockdown, vaccination, treatment, etc.).

Competition: We will establish a number of “judges’ models” for 1/1/2021 – 10/1/2021: some that reflect mild and under-control pandemic dynamics, and some that reflect difficult to control pandemic dynamics that will include several unexpected events. We will then apply your policy function to the judges’ models and examine its performance according to the relative benefit measure above, obtaining a net relative benefit score over scenarios.

<u>Keep in mind:</u> Maintaining status quo may incur cost! This is because the status quo you maintain may reflect considerable intervention over the underlying judges’ model.

Reporting Requirements

The report is a key part of this Case Study and should include presentation of your methods and results for each of the above parts. In particular, consider and discuss the following:

A. Modeling validity: To what extent is your model accurate (i.e., fits the data)? Does your model seem valid, or does it make predictions/assumptions that are not realistic and/or are overly strong? It is often said that all models are wrong but some are useful. Discuss your model in these terms.
B. Policy design: What are the ethical implications of your policy design? Does the above policy framework make sense, or would you deviate from this and why? How does the outcome of your constructed policy compare with what actually happened in 2021, thus far? Try and interpret any similarities or differences.

What to Turn In

1. Your built-up files. Write your own scripts as necessary to make figures similar to Figure 2, above, which compare your model outputs to the actual data. Use the matlab `publish` command as appropriate to highlight key steps or outputs, and prepare a pdf or doc for submission (similar to what you have done on MATLAB homeworks), in addition to your code. **Make your own decisions regarding what steps are most salient and need to be highlighted.**

2. The following MATLAB files.

    a. A file ‘model.mat’ that contains your baseline SLIRD transmission matrix as of 1/1/2021.
    b. ‘sirpolicy.m’ that contains your sirpolicy function, which will be used for competition scoring.

3. A 5-7 page report illustrating your simulations and design choices. **Use the templates provided with the case study materials.** Make sure that your report clearly states/presents:

a. All modeling choices
b. Plots and/or Figures illustrating model outputs
c. Conclusions about the design strategies based on the quantitative findings presented in the report.
d. You do not need to provide the entire model with all numerical values, etc.

4. A signed honor code document, as distributed with the case study materials.

Rubric

1. Correctness of MATLAB code (30%)
    a. Simulation of SIR model, SIR network model and intervention model
2. Presentation (20%)
    a. Plots and Figures should be clear
    b. Written report should be well-organized and clearly written
3. Programming style (20%)
4. Study design (30%)
    a. Design of interventions
    b. Discussion of modeling choices
    c. Discussion of modeling validity
    d. Discussion of ethical implications