Analyzing COVID Data Using the 'SIRD' Method

Ryan Kropp
Electrical and Systems Engineering
Deparement
Washington University
St. Louis, United States of America
kropp.r@wustl.edu

Matthew Walker
Electrical and Systems Engineering
Department
Washington University
St. Louis, United States of America
matthew.r.walker@wustl.edu

Robert Walsh
Electrical and Systems Engineering
Department
Washington University
St. Louis, United States of America
r.j.walsh@wustl.edu

ABSTRACT

I. Introduction

For this project, we had to use our base understanding of the 'SIRD' method to analyze new data and create new models of the data that could be comparable to the actual data in several different scenarios. The 'SIRD' model stands for susceptible, infected, recovered, and dead. It uses proportions to predict how a virus will spread, and in this case, we're using the model to analyze COVID-19 data. The method follows a dynamic model and works via matrix multiplication, where A is our 'SIRD' matrix, which is 4 rows by 4 columns and gets multiplied by our initial conditions, x, to find the next state of the dynamic model. This process repeats continuously as time progresses, and it helps us model the statistics for the population as the disease progresses.

Firstly, for part one we had to find an initial matrix without taking reinfection into account, and then we found another matrix that includes reinfection data. We used educated guesses for our proportions for these two matrices to get a base understanding of the model and what it's supposed to do.

For the next part, we did three different things: We first used the functions given to us to create a new model that got generated around actual COVID data given to us, creating a new vector 'x' with values that we could easily augment. We then set our function equal to the cost so we could minimize the difference between the actual and modeled data. We then had to research how effective different policies could be at minimizing death rates and spread rates, and then lastly took different segments and explained how we could reduce infection and death rates.

Finally, for our last part, we put in our competition data. FINISH THIS

II. Methods

Our goal for this project was to analyze the dynamic SIRD model and apply it to real COVID-19 data. By augmenting the SIRD model, we could realistically predict how the data would change with different governmental policies, such as requiring vaccination and/or mandating masks. We want to use the dynamic model to help reduce the spread and fatality rates of the disease by the greatest amount.

A. The Initial Matrices

B. Using a 'cost' Function to Model Cases and Deaths

For this part, we wanted to find a way to have MATLAB give us our 'SIRD' proportions by essentially creating a function that uses initial conditions and logical infection/death rates to create a new dynamic model.

We created a seven-element column vector 'x' that included initial conditions as well as infection, fatality, and recovery rates. With 'x', we could create our dynamic model.

We did this by basically making a function that takes in data and then creates a linear fit that models infection and death rates for the future based on the inputs. Our inputs here were given to us from the US Census Bureau (**2**), and then we used MATLAB to generate a function that took those inputs and made that linear fit along time t, using the 'lsim' and 'ss' functions provided to us.

However, we then needed to create a cost function that essentially sought to minimize the difference between our actual data and the data that those functions generated for us. We decided on a cost function like a chi squared test, but since we were just trying to minimize our values, we disregarded the denominator piece, leaving us with the sum the squared difference between every cell of the actual cases and the modeled cases, which was added to the that same squared difference but for the actual and modeled deaths. Our 'lsim' function now took the cost into account and now also sought to create a line of best fit that was as close to the real data as possible. We then plotted these lines of best fit against the actual data.
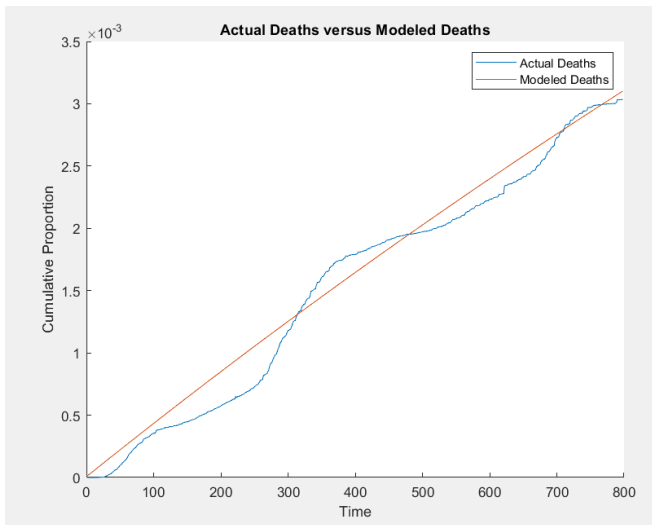
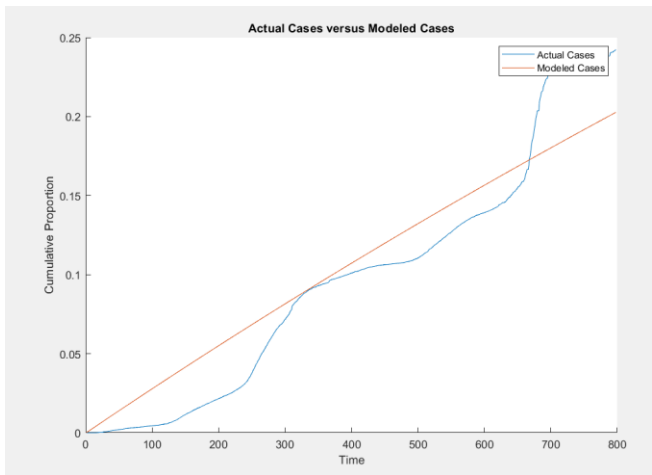Figure 3: Actual Deaths versus Modeled Deaths

Figure 4: Actual Cases versus Modeled Cases

These figures provided excellent results and shows the line of best fit went exactly as planned in both figures. We can assume this dynamic model is solid and can accurately predict cases for the future.

C. Creating a New Model with Ideal Conditions

Next, we wanted to find a new way to prevent death and infection rates using real world data for the effectiveness of each preventative method.

According to the CDC, the Pfizer vaccination prevented COVID-19 hospitalizations by 96% among those 65 and older. (**4**) Since this is the most at-risk population, we can assume the efficacy towards the rest of the population is at least this high. Mortality of the hospitalized infected is close to a fixed 10%, regardless of the at-the-time hospitalization counts. (**3**) If the hospitalizations are reduced by 96%, and we assume almost all fatalities occur in hospitals, then mortality can also be reduced by 96%.

According to the National Library of Medicine, in a 2009 study, masks were proven to be at least 80% effective

against influenza-like diseases. (**5**) Since COVID-19 is an influenza-like disease, we can safely assume there will be an 80% reduction in the COVID-19 infection rate with masking mandates in place, making them another effective measure against the disease.

We can incorporate these values into a new fit, named 'Y_fit_new' in our MATLAB file. Using 'Y_fit_new', we realistically adjust the function's case and death rates based on these masking and vaccination effects. It is within reason to say that on day 0 of the outbreak masks can immediately affect the spread, so we can augment our infection rate to be 20% of the initial value. Vaccine development, however, can't be so quickly implemented. A reasonable guess for when a vaccine could be developed and administered to the bulk of the population is around one year (with the assumption that the world has been exposed to coronavirus strains before,) and for our data we assumed the unvaccinated population to be small enough to ignore (though we know this isn't the case in many regions/countries worldwide.) Our death rate post-vaccination was reduced by the same 96% cited above. So, our case rate was only affected by the mask mandate, but our death rate was affected by both.
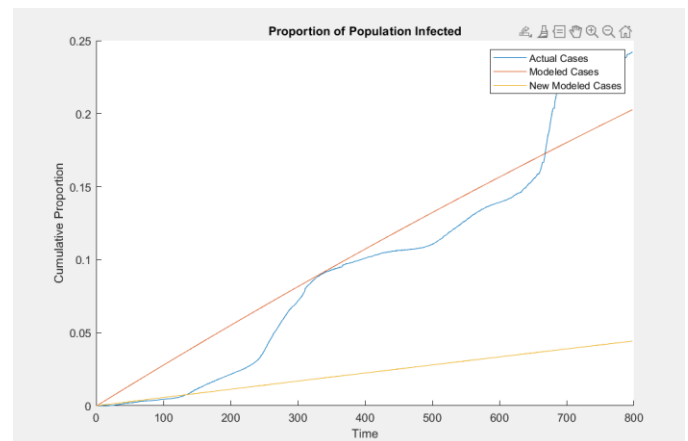

Figure 5: Plot of actual and originally modeled cases, now with the new modeled cases after taking the masking mandate into account from day one.
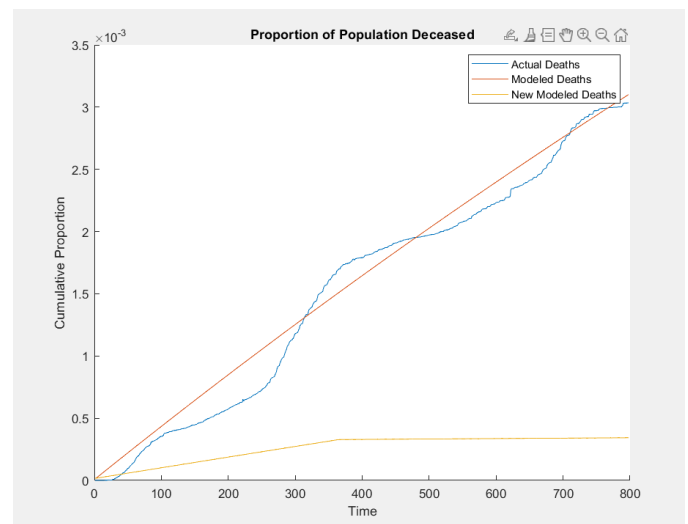
Figure 6: Plot of the actual and originally modeled deaths, now with the new modeled deaths after taking the masking mandate into account from day one and the new vaccination into account on the 365th day.

In Figure 5, we plotted the actual cases against the previously and newly modeled cases, and in Figure 6 we did the same for deaths, where the vaccine-induced mortality reduction is delayed 365 days. We got remarkably positive results, showing how effective these policies can be with ideal implementation

D. 3.5

E. 4.1/4.2

III. RESULTS AND DISCUSSION

IV. CONCLUSION

REFERENCES

[1] Boyd, S., & Vandenberghe, L. (2018). *Introduction to Applied Linear Algebra: Vectors, Matrices, and Least Squares*. Cambridge: Cambridge University Press. doi:10.1017/9781108583664

[2] Bureau, US Census. "County Population Totals: 2020-2021." *Census.gov*, 1 Mar. 2022, https://www.census.gov/data/datasets/time-series/demo/popest/2020s-counties-total.html.

[3] "Covid-19 Hospital Data in Mortality among Confirmed Covid-19 Encounters by Week." *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, 12 Sept. 2022, https://www.cdc.gov/nchs/covid19/nhcs/hospital-mortality-by-week.htm.

[4] "Effectiveness of COVID-19 Vaccines in Preventing Hospitalization among Adults Aged ≥65 Years - Covid-Net, 13 States, February–April 2021." *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, 12 Aug. 2021, https://www.cdc.gov/mmwr/volumes/70/wr/mm7032e3.htm.

[5] *National Center for Biotechnology Information*, U.S. National Library of Medicine, https://www.ncbi.nlm.nih.gov/.

[6] "MATLAB Deep Learning Toolbox." *Deep Learning Toolbox Documentation*, https://www.mathworks.com/help/deeplearning/index.html.

[7] The New York Times. "Coronavirus (Covid-19) Data in the United States." *GitHub*, 2021, https://github.com/nytimes/covid-19-data.