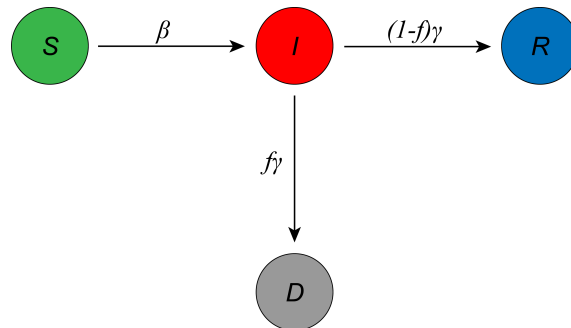Today we're going to analyze a simple compartment model used in epidemiology, the SIR (Susceptible, Infected, Recovered) model. This forms the basis for many other, more complicated models. The following is a flow diagram for disease progression:

The ordinary differential equations (ODEs) are:

$$\frac{dS}{dt} = -\beta SI$$

$$\frac{dI}{dt} = \beta SI - \gamma I$$

$$\frac{dR}{dt} = (1 - f)\gamma I$$

$$\frac{dD}{dt} = f\gamma I$$

Where:

S = Susceptible population
I = Infected population
R = Recovered population
β = disease transmission rate
γ = recovery rate
f = fraction of infected that die

Another important factor is $\mathcal{R}_0$, the reproductive ratio. If it is 1, it means 1 infected person typically infects one other person. If this number is below zero, you do not see a rise in cases, whereas the opposite is true if $\mathcal{R}_0 > 1$.

If this was a linear set of ODEs, we would be able to solve explicitly for S, I, R, etc.

$$\frac{dy}{dt} = \begin{bmatrix} S \\ I \\ R \end{bmatrix}$$

$$y(t) = \sum v_i e^{\lambda_i t}$$

However, nonlinear equations are tricky. One way to get a feel for them is to plot a phase diagram, where the axes are variables, and you plot little arrows at different points in space to represent the velocities, or flows, of those variables. This plot of arrows is typically called a "quiver" plot.

No Intervention
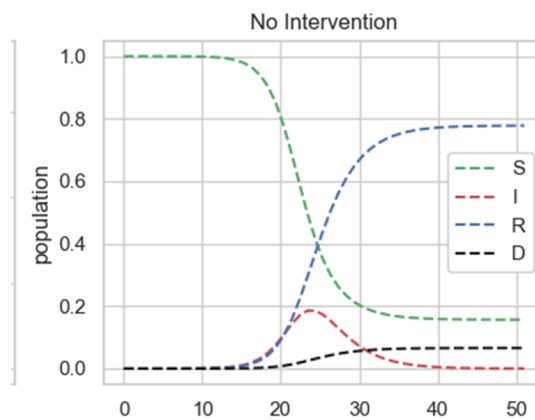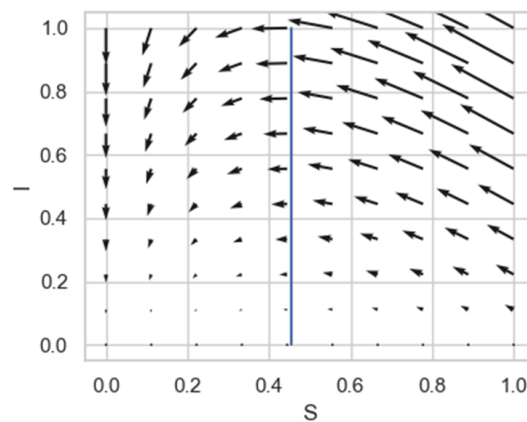
Even though you don't *explicitly* know $S(t)$ and $I(t)$, you can get a feel for how things are going to behave. The susceptible population decreases, and the infected initially rise, and then fall back down again. There is a line plotted here where $\frac{dI}{dt} = 0$. That is the nullcline for I. You can solve for this line by setting $\frac{dI}{dt} = 0$, and then solving for S.

We can also make approximations at critical points in the model.

When $\beta SI \gg \gamma I$:

$$\frac{dI}{dt} \approx \beta SI,$$

$$I(t) \approx e^{\beta St}$$

When $\beta SI = \gamma I$

$$\frac{dI}{dt} = 0$$

When $\beta SI \ll \gamma I$

$$\frac{dI}{dt} \approx -\gamma I,$$

$$I(t) \approx e^{-\gamma t}$$

But you would also like to know the time dependent versions of the populations, i.e. you would like to integrate the ODEs, but as I mentioned earlier, for most nonlinear ODEs, there simply isn't an analytical solution, so you have to solve them numerically.

This is a common problem in biology, and in science in general. For example, many ODEs for chemical interactions have nonlinear terms, like the βSI in the equations above.

To solve the ODEs numerically, a commonly used algorithm is the Runge-Kutta (RK) method, which is an average of several nested Euler-based approaches.

Fourth-order Runge-Kutta (RK4):

$$\frac{d\mathbf{y}}{dt} = f(t, y), \quad \mathbf{y}_0 = \begin{bmatrix} S_0 \\ I_0 \\ R_0 \end{bmatrix}$$

$$y_{n+1} = y_n + \frac{1}{6}h(k_1 + 2k_2 + 2k_3 + k_4),$$

$$t_{n+1} = t_n + h$$
$$k_1 = f(t, y),$$

$$k_2 = f(t_n + \frac{h}{2}, y_n + h\frac{k_1}{2}),$$

$$k_3 = f(t_n + \tfrac{h}{2}, y_n + h\tfrac{k_2}{2}),$$

$$k_4 = f(t_n + h, y_n + hk_3),$$

In the example code, the maximum population is set to "1". To determine how a larger population would behave, you can simply multiple all calculated populations by the population of interest. For example, NYC has ~8M people. The parameters in the example code have been chosen to approximate the conditions in NYC from March 8, 2020 to March 20, 2020 when a shelter-in-place order was instituted.

1. Using the β, γ, and f provided in the sample code, explore how this particular model would behave.
    a. Quiver Plot: Plot a quiver plot to see how the S and I populations will change.
    b. Integrate: Write an RK4 algorithm to predict how S, I, R and D.
    c. Plot how these populations will behave in 12, 50, and 100 days.
    d. Plot the cumulative number of infected and dead. For infected, this will be the sum of I, R, and D.

Use:

$$\mathbf{y}_0 = \begin{bmatrix} S_0 \\ I_0 \\ R_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 23/8M \\ 0 \end{bmatrix}$$

2. On March 20th, 2020, a shelter-in-place order was announced. This essentially decreased the susceptible population (S) that was accessible to infected individuals (I). To model this, do the following:
    a. Perform the 12 day modeling you did in 1c.
    b. Perform another round of modeling for 38, or 88 days, but start these simulations with the following populations:

$$\mathbf{y}_0 = \begin{bmatrix} S_0 \\ I_0 \\ R_0 \end{bmatrix} = \begin{bmatrix} 0.42 * S_0 \\ I_{12} \\ R_{12} \end{bmatrix}$$

In other words, start I and R with the results you got from the 12th day of the prior simulation. This will mimic the effect of social distancing.

c. Plot the cumulative number of infected and dead. For infected, this will be the sum of I, R, and D.

How do the models compare? Does cutting down the population by ~50% cut the infections down by 50%? More? Less?

It is important to note that this is a very simple model. A more accurate model of COVID-19 would be:
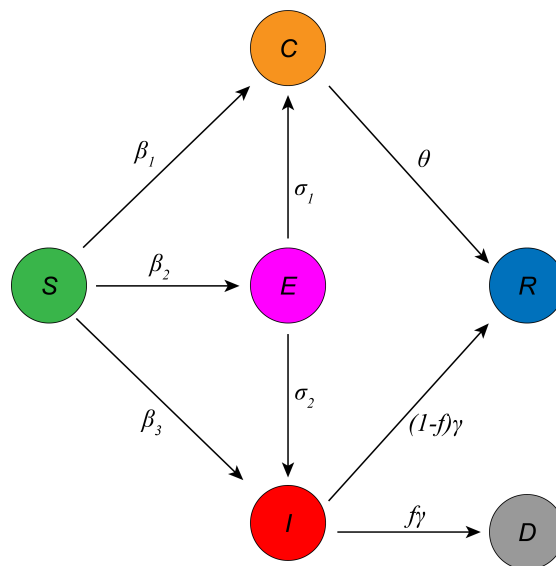
Where,

C = Carriers
E = Exposed

Also, unlike chemicals, humans are not well-mixed and diffusing, which the continuous ODE assumes. A more accurate model would be stochastic, and involve spatial and temporal differences in the populations, as well as the infection parameters. With these models, rather than use single parameters, you draw them from a distribution, and draw your populations from distributions to represent spatial and temporal dynamics of a region.