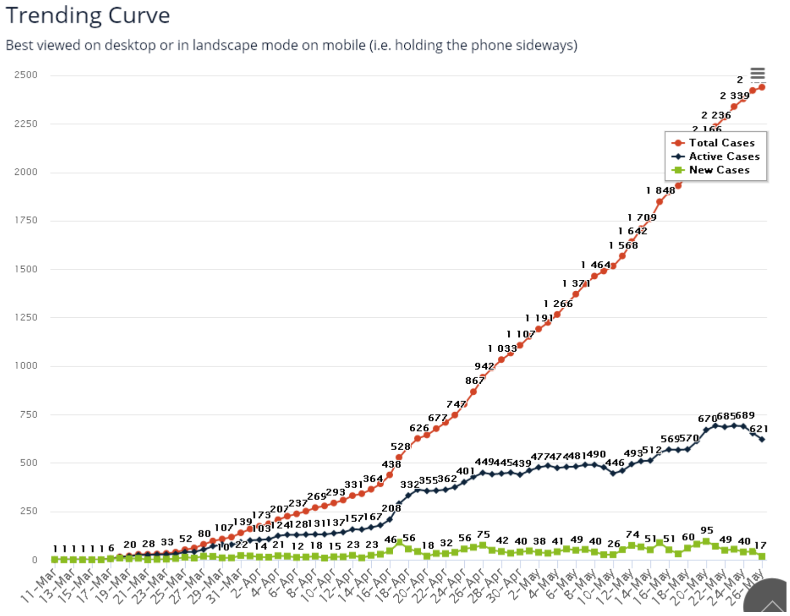
CoronaVirus & Markov Chains



CoronaVirus infections in North Dakota as of May 28, 2020
www.health.nd.gov/diseases-conditions/coronavirus/north-dakota-coronavirus-cases

Markov chains assume a system is of the form

$$X(k+1) = A\ X(k)$$

i.e. there is no memory in the system - meaning X(k+1) does not depend upon states prior to X(k). If the state transition matrix (A) is a constant matrix, then X(k) can be found using eigenvalue and eigenvectors or with z-transforms.

The state transition matrix does not have to be a constant matrix, however. In such cases, numerical solutions (i.e. Matlab) can be used.

A relevant example today is the Corona virus outbreak in the United States. Using what we know from this course, we can simulate the outbreak and analyze different situations, such as

- What impact does social distancing have on the disease spread?
- What impact does wearing masks have on the spread?
- What impact would a vaccine have - which reduces the recovery time to 1/2?

Since we don't have actual numbers for many of the parameters used in this lecture, we'll take a scientific wild-ass guess (SWAG) at them.

System Modeling

Assume there are 4 groups of people:

- H: Healthy and uninfected: People who have not caught the Coronavirus and are susceptible to it.
- I: Infected: People who have the Coronavirus and can spread the disease
- C: Cured: People who had the Coronavirus but have recovered.
- D: Dead: Self explanatory.

Assume that people who are cured

- Cannot catch the Coronavirus again (this is not certain), and
- They cannot spread the Coronavirus any more (this is not certain either)

Define:

- N = Contact Rate: The number of people the average person comes in close contact with each day.
- p = The transmission rate: the probability that someone will catch the disease if they are in close contact with an infected person.
- Cure Rate = 5%: The chance that an infected person will recover on a given day, and
- Death Rate = 0.07% The chance that in infected person will die on any given day.

The cure rate doesn't quite fit into a Markov chain since it appears you are cured 20 days on average (guess?) after catching the Corona virus. To make this fit the Markov chains, assume that the cure rate is a geometric distribution with a mean of 20 days

$$p(cure) = \left(\frac{1}{20}\right)\left(\frac{19}{20}\right)^{k-1} u(k-1)$$

This isn't exactly correct, but the mean matches our guess and it works with Markov chains. The death rate also isn't certain - but initial estimates are that the death rate is 1.4%[1]. So, assume the death rate is 1.4% of the cure rate: 0.07%.

Then, the number of new cases each day will be

- The number of infected people at day k, I(k), times
- The number of people each person comes in contact each day (N), times
- The chance of infecting someone (p), times
- The change that the person contacted was uninfected
- Minus the number of people who are cured (5% of the infected people),
- Minus the number of people who die (1.4% of 5% of the infected people)

or

$$\delta I(k) = I(k)\left(N \cdot p \cdot \left(\frac{H(k)}{H(k)+I(k)+C(k)}\right) - 0.05 - (0.014)(0.05)\right)$$

$$= (a - 0.05 - 0.0007) I(k)$$

where

---

[1] https://www.worldometers.info/coronavirus/coronavirus-death-rate/

$$a = N \cdot p \cdot \left(\frac{H(k)}{H(k)+I(k)+C(k)}\right)$$

The number of new cured people is

$$\delta C(k) = 0.05 \cdot I(k)$$

The number of new dead people is

$$\delta D(k) = 0.014 \cdot 0.05 \cdot I(k)$$

Cured and Dead are absorbing states (once there, you cannot leave).

This results in the Markov chain model being

$$\begin{bmatrix} H(k+1) \\ I(k+1) \\ C(k+1) \\ D(k+1) \end{bmatrix} = \begin{bmatrix} 1 & -a & 0 & 0 \\ 0 & 1+a-0.05-0.0007 & 0 & 0 \\ 0 & 0.05 & 1 & 0 \\ 0 & 0.0007 & 0 & 1 \end{bmatrix} \begin{bmatrix} H(k) \\ I(k) \\ C(k) \\ D(k) \end{bmatrix}$$

Note that all of the columns equal 1.000 (required for this to be a valid Markov chain).

Simulation #1: March 1, 2020

Assume most of the population is healthy. This results in

$$a = N \cdot p \cdot \left(\frac{H(k)}{H(k)+I(k)+C(k)}\right) \approx Np$$

From a mathematical standpoint, we don't really care how you arrive at Np: only the net result matters.

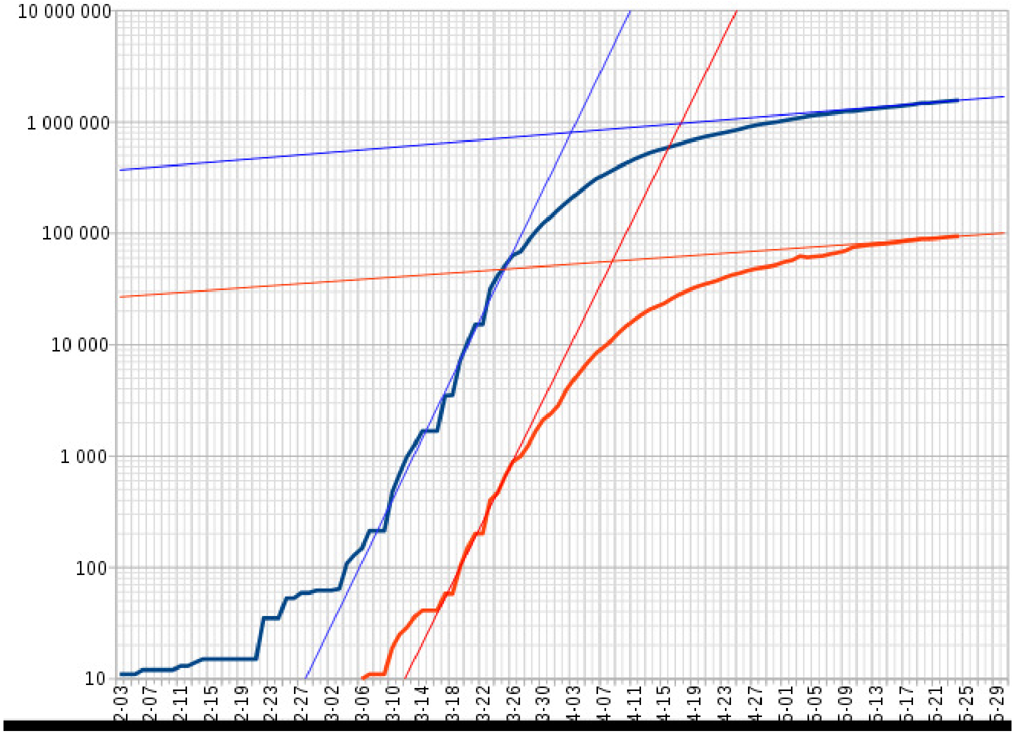
'a' determines the growth rate of the disease

$$I(k+1) = (1+a-0.05-0.0007)I(k)$$

or

$$I(k) = (1+a-0.05-0.007)^k \; I(0)$$

To determine this growth rate, look at a graph of the disease spread:



Disease spread of CoronaVirus in the United States
Blue = cases, orange = deaths. Source: Wikipedia

In early March, the disease was growing at a rate of

- 10x increase every 9 days
- 2x increase every 3 days

This gives

$$(1+a-0.05-0.007)^9 = 10$$

$$1+a-0.05-0.007 = 1.29155$$

or

$$a = 0.34815$$

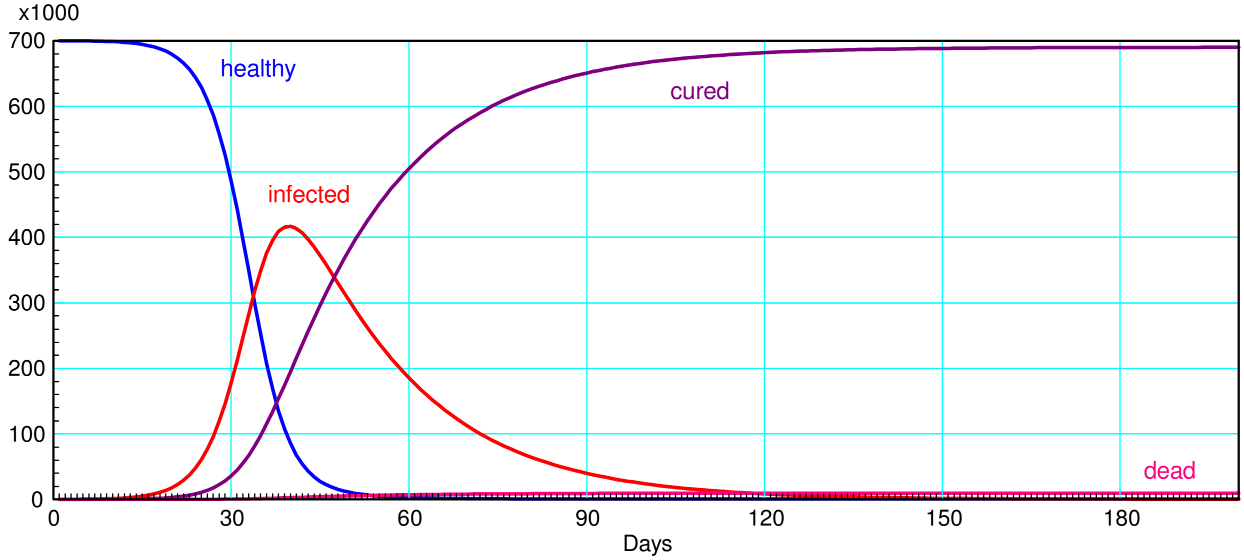
$$Np = 0.34815$$

Again, mathematically, how you come up with this number doesn't matter. One reasonable way to do this is

- N = 10 *the average person comes in contact with 10 people each day*
- p = 0.0348 *the chance of catching Coronavirus is 3.48%*

Using these numbers, you can simulate what will happen in the state of North Dakota assuming an initial condition of

- $H(0) = 700{,}000$ *the population of North Dakota*
- $I(0) = 100$ *start with 100 people infected as of March 1, 2020*



Markov chain simulation with Np = 0.34815
Prediction of what would have happened if no preventive measures were taken

This model predicts that if nothing had changed as of March 1st (and the entire state was interacting at the same rate and there were no people entering or leaving, etc.),

- The peak in the number of people infected would have been in late April,
- 416,920 people would have been infected, and
- 9658 people will eventually die (about 1.4% of 700,000)

While this model is simplistic in the extreme, it does match up fairly well with more sophisticated models. It also points out some of the problems with not taking action:

- If 416,920 people were infected at the same time, our hospitals would be over whelmed.
- If hospitals could not see any new patients, the death rate would have gone up substantially
- etc.

The main thing to take from this is that

- Markov chains can be applied to more than just electric circuits, and
- They also allow you to ask various questions in simulation.

The latter helps policy makers understand the impacts of different policies.

Matlab Code:

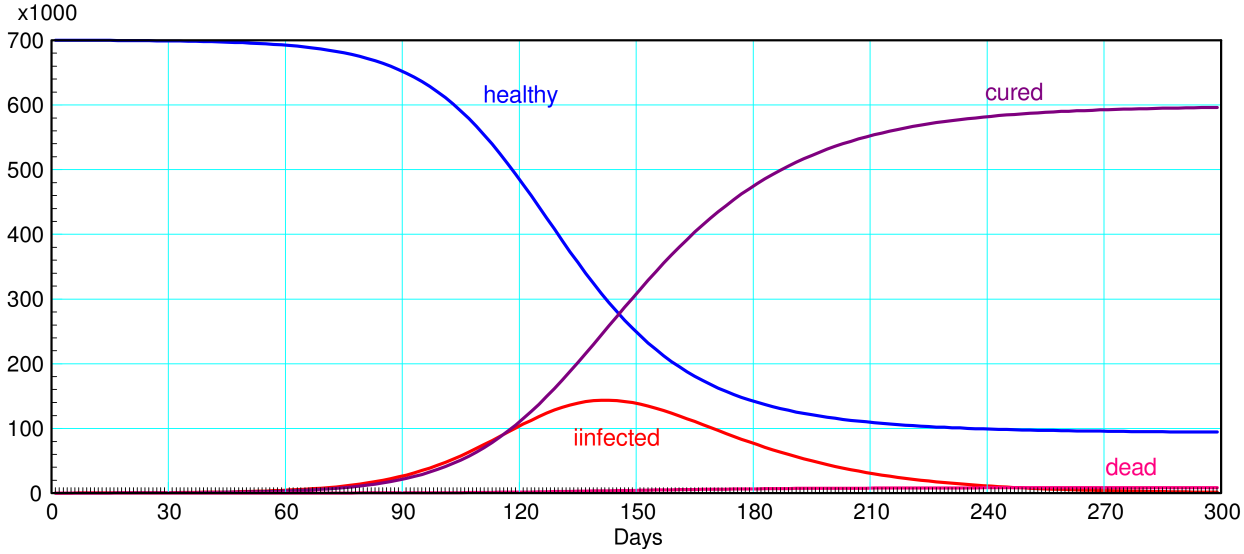
```
% States X
% X(1) = Uninfected
% X(2) = Infected
% X(3) = Cured
% X(4) = Dead;
X = [700000;100;0;0];

CureRate = 0.05;
DeathRate = 0.014 * CureRate;
Infectivity = .034815;
N = 10;              % interactions per person
% note:  a = N * Infectivity

Y = [];

for i=1:200

    NewInfections = X(1) * X(2) * Infectivity * N / sum(X(1:3)) ;
    Cures = CureRate * X(2);
    Deaths = DeathRate * X(2);

    X(1) = X(1) - NewInfections
    X(2) = X(2) + NewInfections - Cures - Deaths;
    X(3) = X(3) + Cures;
    X(4) = X(4) + Deaths;

    Y = [Y ; X'];
end

plot(Y)
```

Simulation #2: Increase Social Distancing

With these models, you can start to ask questions

- What impact would it have if N were reduced 3x? (self isolation, closing schools, etc), or
- What impact would it have if p were reduced 3x? (people start wearing face masks)

This is easy to simulate: just divide either N or p by 3 (doesn't matter which one).



Markov chain simulation with Np = 0.34815 / 3
Prediction of how the Coronavirus would spread if the Np were reduced by 3x

Note that by dropping either term by 3x

- The maximum number of people who are infected drops from 416,920 down to 143,710 (reducing the stress on our health care system)
- The peak is pushed out to 140 days (4.5 months),
- The disease eventually dissipates with 83,000 people never having been infected, and
- The final death toll is 7427 - slightly less than simulation #1

Simulation #3: Increase in social distancing and wearing masks

Next,

- What impact would it have if N were reduced 3x? (self isolation, closing schools, etc), and
- What impact would it have if p were reduced 3x? (people start wearing face masks)

for a net results of Np reducing 9x.

This is more akin to what actually happened in mid March with schools going to remote learning, bars, health clubs being closed, and people being asked to wear face masks and social distance.

By reducing Np 9x, the growth rate drops to less than one - meaning that the disease should dissipate exponentially.

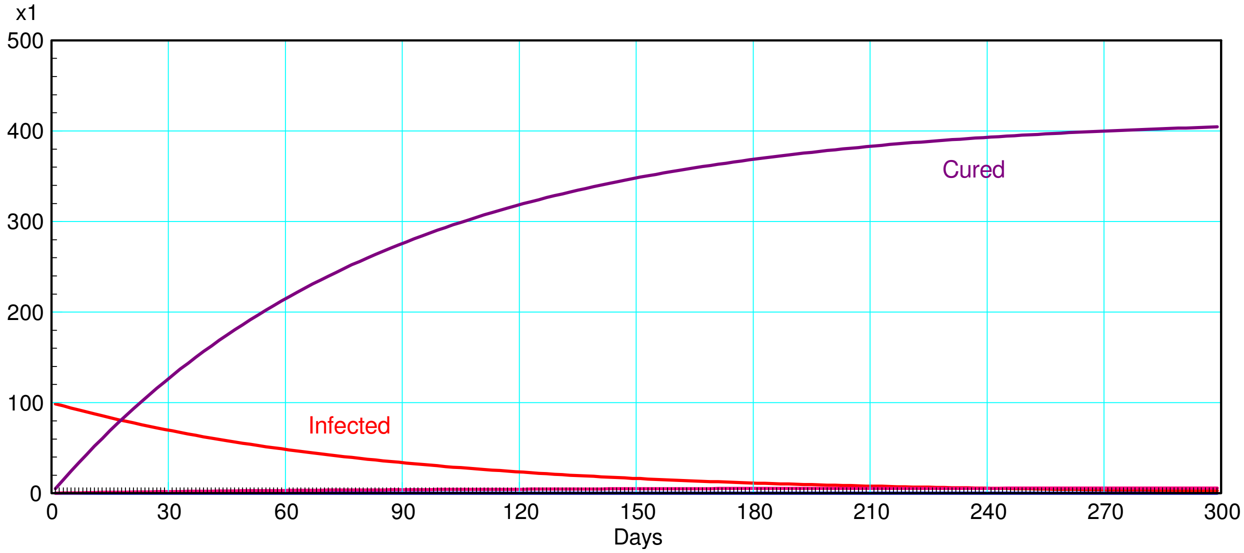
$$y(k+1) = (1 + Np - 0.05 - 0.0007)\ y(k)$$

$$y(k+1) = (1 - 0.012)\ y(k)$$

resulting in the disease never taking hold. The result is

- The total number of people infected is 412
- The total number of deaths = 5

The total number of deaths in this case is 140 people infected is only 400. This is really the goal: if each person infects less than one other person, then the disease will dissipate and never become a problem.



Markov chain simulation with Np = 0.34815 / 3
Note that the Y axis is times 1 (vs 1000 in previous graphs)
Prediction of how the Coronavirus would have spread if Np were reduced 9x

Simulation #4: Vaccine

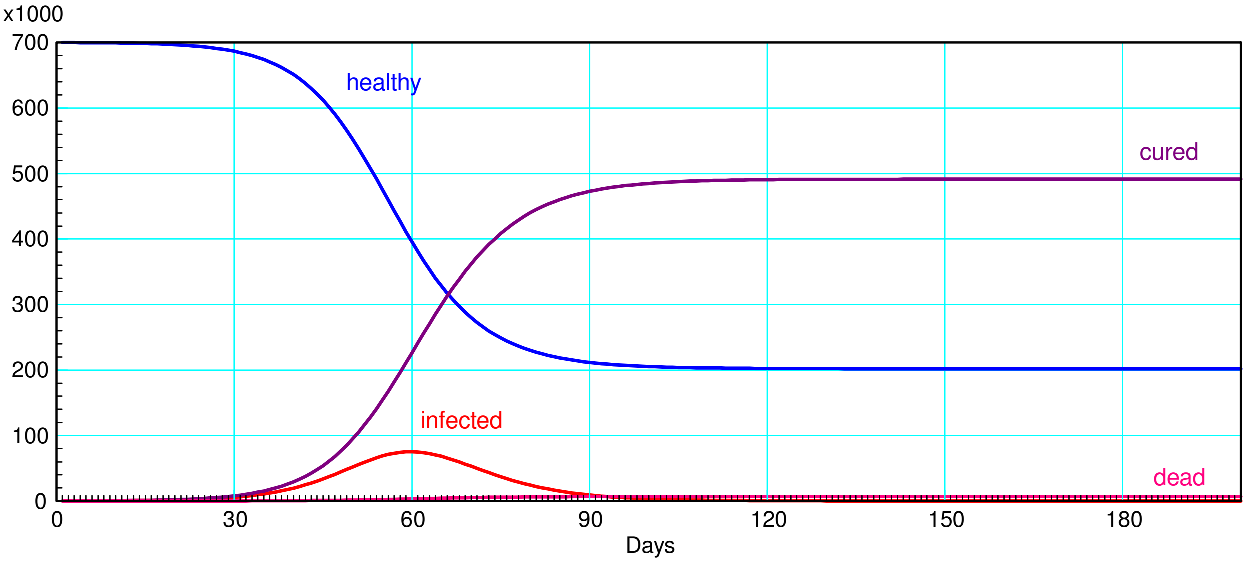
What happens if a vaccine is found? This can be modeled by reducing the time a person has the disease (20 days on average in these simulations) to, say, 5 days. To incorporate this into the model, change the cure rate from 0.05 (mean = 20 days) to 0.2 (mean = 5 days).

If you repeat

- Without social distancing (N = 10), and
- Without wearing masks (p = 0.0348)

you wind up with

- A smaller but still sizable peak in the number of infected people (83,000)
- With a peak that's 2 months out (60 days), and
- 6922 deaths before the Coronavirus dissipates.



final values 201.0856 0.0000 494.4915 6.9229

The previous simulation assumes that once a vaccine is found, we revert to the March 2020 policies of

- No social distancing,
- No wearing of masks
- Everything is open.

This simulation suggests that, even with a vaccine which reduces the time that you are ill by 4x, you still could wind up with a large number of infected people and a large number of deaths.

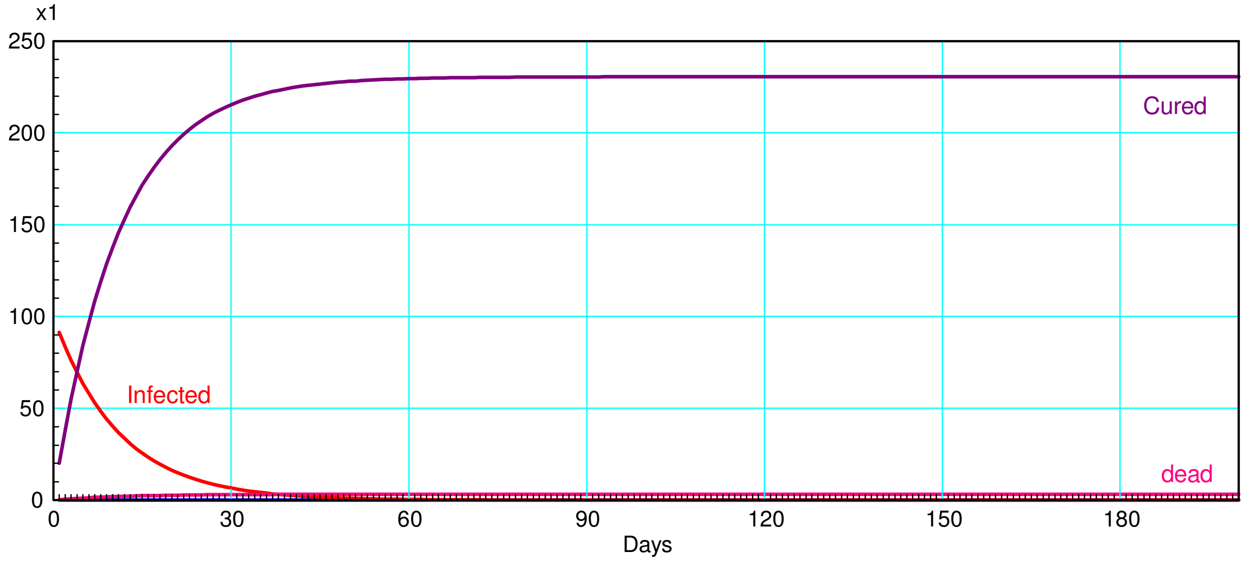
So, suppose instead that you combine

- A vaccine that reduces the time to overcome the illness from 20 days to 5 days, and
- People are still asked to wear masks (and they do so - another big assumption), reducing the infectivity rate by 3x.

These two combined in the previous simulation result in the following graph. Note that

- The infectivity rate drops to less than 1 - resulting in the number of people being infected dropping exponentially,
- The total number of people infected becomes just 240, and
- Only 80 people die from the disease.

This suggests that it might be wise to continue to wear masks and/or social distance, even after a vaccine is found.



696.7002 0.0000 5.7197 0.0801

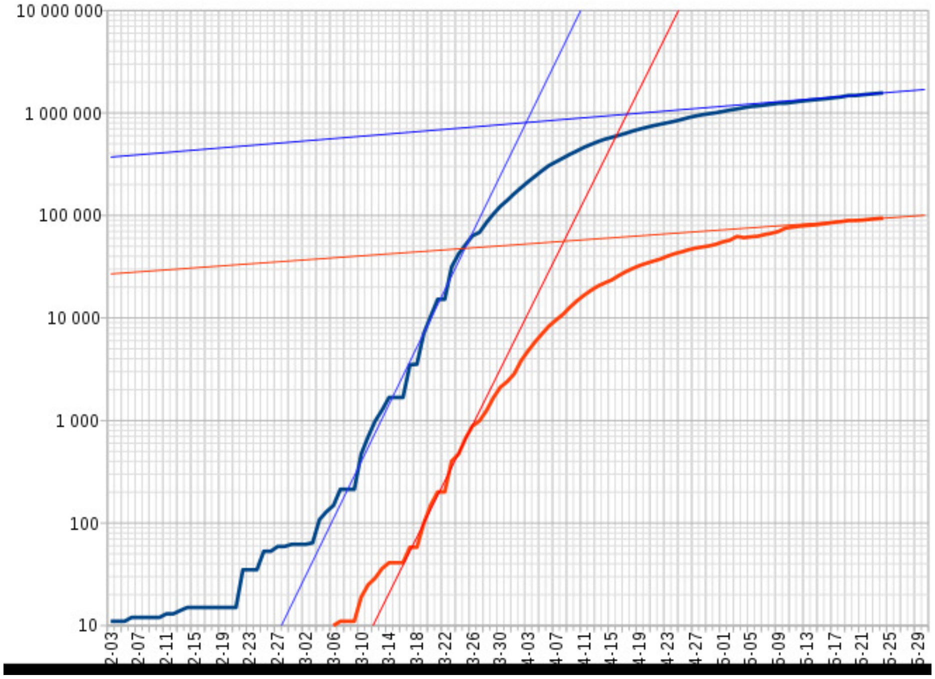
Simulation #5: May 1, 2020

Finally, lets model the situation we're in as of May 1, 2020. At this point

- Schools are closed (sort of) with classes being held via remote learning
- Bars, restaurants, health clubs are closed,
- Nonessential workers are asked to self isolate,
- Travel is strongly discouraged - except to buy essentials such as groceries, and
- People are asked to wear face masks in public.

The impact of these changes shows up rather clearly in the number of cases of Coronavirus reported in the U.S.

- Prior to May 1, 2020, the disease was doubling every 3 days
- After May 1, 2020, the disease was still growing, but the doubling time increased to 56 days



Growth of CoronaVirus in the United States as of May 28, 2020
Blue = cases, orange = deaths. Source: Wikipedia

This implies that the growth rate (a) dropped to

$$(1 + a - 0.05 - 0.0007)^{56} = 2$$

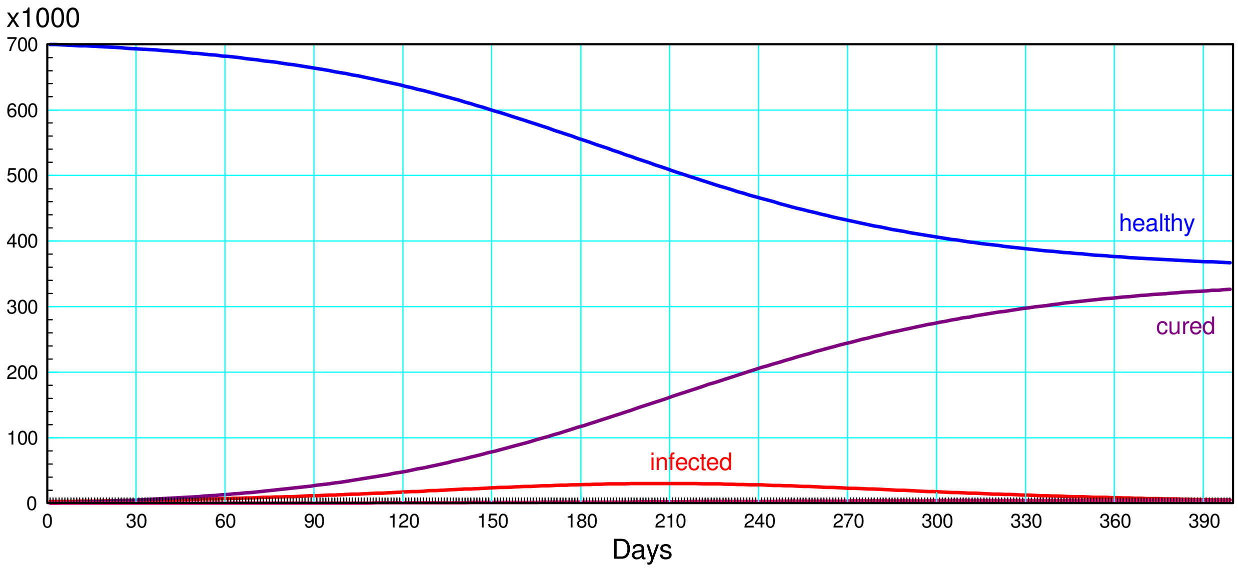
$$a = 0.06298$$

Using this growth rate in our simulation, we can predict how the disease will spread in North Dakota, making some (incorrect) simplifying assumptions:

- Everyone in the state has an equal probability of interacting with everyone else in the state, and
- North Dakota has the same infectivity rate as the United States (0.06298)

Note with this simulation

- The maximum number of people infected is 16,540
- The maximum happens in December, 2020,
- The total number of deaths in North Dakota is 3067, and
- The disease is around for over a year.



Prediction for how the CoronaVirus will spread assuming Np = 0..06298

Disclaimer:

The simulations presented are very simplistic and are in no way intended to challenge the models made by the Center for Disease Control or to influence policy decisions made at the state or local levels. These models, simplistic as they are, are reasonably consistent with other (more sophisticated) models.

The point behind this lecture is that Markov chains are useful: they allow you to analyze systems ranging from tennis matches to electronic circuits to disease studies. They're kind of fun to play with too....