

GTDMO - 3rd Assignment

Group 6

November 25, 2020

1 Introduction to the Problem

We can interpret the spread of the epidemics as a stochastic process because the individuals of the population N can assume 3 different states, such as:

- S = Susceptible: not infected and not immune (can get infected)
- I = Infected: the individual has already contracted the virus
- R = Removed: the individual has recovered and become immune, or dead.

At each temporal step the individual can either remain in the same state or move to the following one:

- susceptible individual becomes infected with probability 0.09
- each infected individual is removed with probability 0.27
- each removed individual returns back to be susceptible with probability 0.09

Supposed that we are considering a population $N=10000$, our aim is to find the number of individuals in the 3 states after a high number of time steps.

2 Problem Interpretation

This problem can be represented as a Markov Chain as we have 3 different possible states, and each individual can move from a state to another with one time step, moving forward in time.

As initial distribution we have that each individual belongs to the susceptible state (S). We represent it as an initial state vector:

$$s = [1 \quad 0 \quad 0] \quad (1)$$

Given the transition probabilities from one state to the other, our transition matrix is:

$$\begin{matrix} & S & I & R \\ \begin{matrix} S \\ I \\ R \end{matrix} & \begin{pmatrix} 0.91 & 0.09 & 0 \\ 0 & 0.73 & 0.27 \\ 0.09 & 0 & 0.91 \end{pmatrix} \end{matrix} \quad (2)$$

Taking into account this Markov Chain, we can now make some considerations about it:

- As our set of values of the variable is finite, $S = \{S, I, R\}$, we can define our Markov Chain as having a finite state space, which means that all the set of values of the random variables is a finite set $\{v_1, v_2, \dots, v_n\}$.
- As the transition matrix does not change over time, we can define the Markov Chain to be homogeneous.

We can define a Markov Chain as homogeneous if and only if the transition probabilities are independent of the time t , that is, there exist constants $P_{i,j}$ such that:

$$P_{i,j} = Pr[X_t = j | X_{t-1} = i] \quad (3)$$

holds for all times t .

To find the number of individuals in the 3 states after a high number of time steps, we should check that the Markov Chain has a stationary distribution. To do so, we should verify if the Markov Chain is irreducible and aperiodic.

A Markov Chain is called irreducible if starting from any state s_i there is a positive probability ($P_{i,j}^{(m)} > 0$) to reach every other state s_j in a finite m number of steps. We can say that in this case every state is *accessible* and belong to only one communication class.

By this definition we can state that the Markov Chain we are taking into consideration is irreducible it clearly has this property.

For what it concerns the periodicity, we can define the period $d(k)$ of a state k of a homogeneous Markov Chain with transition matrix P as:

$$d(k) = gcd\{m \geq 1 : P_{i,j}^{(m)} > 0\} \quad (4)$$

if $d(k) = 1$, then we call the state k *aperiodic*. A Markov chain is aperiodic if and only if all its states are aperiodic.

Therefore, for an aperiodic Markov Chain with finite state space, it exists a positive integer N so that

$$P_{i,i}^{(m)} > 0 \quad (5)$$

for every $m \geq N$.

Our Markov chain is clearly aperiodic since the number of steps needed for a state s_i to return to state s_i is not periodic.

We can finally state that any aperiodic and irreducible finite Markov chain has precisely one stationary distribution and since our Markov chain appears to have all these properties we can, therefore, assume that it has a stationary distribution, i.e. a probability distribution that remains unchanged as time progresses. In other words, it shows the long term behaviour of the Markov Chain.

a row vector π is said to be a stationary distribution for the Markov Chain, if it satisfies the following conditions:

- the π_k are nonnegative real numbers such that $\sum_{k=0}^{n-1} \pi_k = 1$.
- $\pi P = \pi$.

3 Solution

A Markov chain with state space S and transition matrix P can be represented by a labeled directed graph $G = (V, E)$ where edges E are given by transitions with nonzero probability. The edge (u, v) is labeled by the probability $P_{u,v}$.

$$E = \{(u, v) | P_{u,v} > 0\} \quad (6)$$

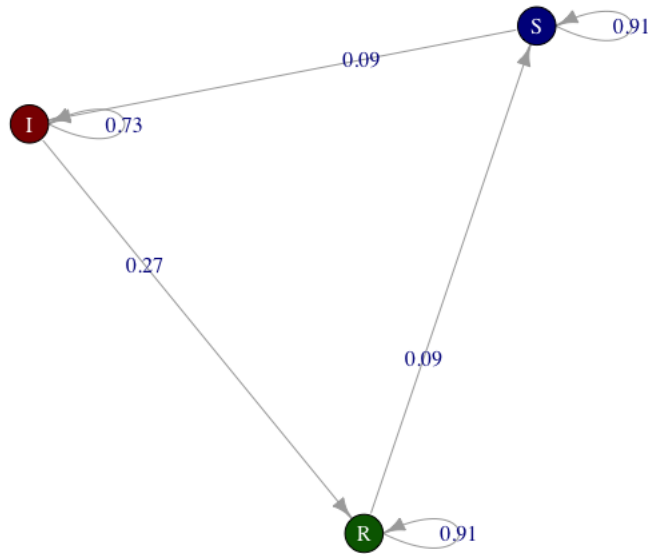


Figure 1: Labeled directed graph

We can establish a long term spread of the virus in the population in the three states S, I and R using the Markov Chain modelling. Therefore we develop the model using R in the following steps:

- Step 1: Create variables
Generate a random number in the states s , with discrete distribution p with the method of the inverse transform where:

s = vector of states

p = probability of each state. Vector of the same length as s

```

1 invtransform<-function(p,s){
2     #generate a random number in the states s,
3     #with discrete distribution p with the method of
4     #the inverse transform
5
6     #s=vector of states
7     #p=probability of each state. Vector of the same length as s
8     n<-length(p)
9     cs<-cumsum(p) # vector of the cumulative sums
10    u<-runif(1,min = 0,max = 1) # generate a random number

```

```

11     d<-cs-u # difference between all the values of the vector and the random value
12     i<-min(which(d>0)) # algorithm to change the state based on the given
        probability
13     s[i] # returns the right states, picked from the vector containing the possible
        states
14 }
15
16 ###n. of simulated steps###
17 nsteps<-10000

```

- Step 2: Matrix definition

Generate a matrix with the variables created for interaction using Markov Chains modelling

```

1 ### initial distribution###
2 p0<-c(1,0,0)
3
4 # Transition Matrix
5 P<-matrix(c(0.91,0,0.09,0.09,0.73,0,0,0.27,0.91),nrow = 3,ncol=3)
6
7 x<-matrix(0,nsteps,1)
8
9 ###Epidemic Spread###
10 states<-c("S","I","R")
11
12 #starting state
13 x[1]<-invtransform(p0,states)
14
15 #define mc as a markov chain
16 mc<-new("markovchain",state=states,transitionMatrix=P)
17 plot(mc)

```

- Step 3: Markov Chain simulation

Run the simulation using the given parameters

```

1 for (j in c(2:nsteps)){ # runs the changing state algorithm for each step
2     if (x[j-1]=="S") # checks if the state is equal to S
3         nrow<-1
4     else {
5         if (x[j-1]=="I") { # checks if the state is equal to I
6             nrow<-2
7         } else { # if the state is not S or I then it's R
8             nrow<-3
9         }
10    }
11    x[j]<-invtransform(P[nrow,],states) # runs the algorithm for the given
        probability (row vector of the transition matrix)
12 }
13
14 x
15 table(x) # n? of people
16 table(x)/nsteps # percentage of people

```

After the simulation, we obtain the row vector π for the stationary distribution of the Markov Chain

$$\begin{array}{ccc} I & R & S \\ (1322 & 4150 & 4528) \end{array} \quad (7)$$

We can finally confirm its nature by checking that it fits properties of stationary distributions such as

- the π_k are nonnegative real numbers such that $\sum_{k=0}^{n-1} \pi_k = 1$.
- $\pi P = \pi$.

In conclusion, we can state that in the long run, the epidemic spread will asses itself on a stationary distribution having the 13,9% of the population Infected, 43,5% of the population Removed and 42,4% of the population Susceptible again. With no further information update in the model, this distribution will remain stable over time.

3.1 Appendix

For the sake of Reproducibility, here I attach the full commented R code used to obtain all Figures, Tables, equations and results.

```
1 library(stats)
2 library(markovchain)
3 library(igraph)
4
5 invtransform<-function(p,s){
6   #generate a random number in the states s,
7   #with discrete distribution p with the method of
8   #the inverse transform
9
10  #s=vector of states
11  #p=probability of each state. Vector of the same length as s
12  n<-length(p)
13  cs<-cumsum(p) # vector of the cumulative sums
14  u<-runif(1,min = 0,max = 1) # generate a random number
15  d<-cs-u # difference between all the values of the vector and the random value
16  i<-min(which(d>0)) # algorithm to change the state based on the given probability
17  s[i] # returns the right states, picked from the vector containing the possible states
18 }
19
20 ### n. of simulated steps###
21 nsteps<-10000
22 ### initial distribution###
23 p0<-c(1,0,0)
24
25 # Transition Matrix
26 P<-matrix(c(0.91,0,0.09,0.09,0.73,0,0,0.27,0.91),nrow = 3,ncol=3)
27
28 x<-matrix(0,nsteps,1)
29
30 ###Epidemic Spread###
31 states<-c("S","I","R")
32
33 #starting state
34 x[1]<-invtransform(p0,states)
35
36 #define mc as a markov chain
37 mc<-new("markovchain",state=states,transitionMatrix=P)
38 plot(mc)
39
40 #representation of the markov chain with igraph
41 mcgraph<-as(mc, "igraph")
42 V(mcgraph)["S"]$color<-"dark blue"
43 V(mcgraph)["I"]$color<-"dark red"
44 V(mcgraph)["R"]$color<-"dark green"
45 plot(mcgraph, vertex.color=V(mcgraph)$color, vertex.size=20, vertex.label.cex=1.5, vertex.
   label.dist=3, edge.arrow.size=0.2, edge.label=E(mcgraph)$prob, edge.curve=0.2)
46
47
48 for (j in c(2:nsteps)){ # runs the changing state algorithm for each step
49   if (x[j-1]=="S") # checks if the state is equal to S
50     nrow<-1
51   else {
52     if (x[j-1]=="I") { # checks if the state is equal to I
53       nrow<-2
54     } else { # if the state is not S or I then it's R
55       nrow<-3
56     }
57   }
58   x[j]<-invtransform(P[nrow,],states) # runs the algorithm for the given probability (row
   vector of the transition matrix)
59 }
60
61 x
```

```
62 table(x) # n? of people
63 table(x)/nsteps # percentage of people
```