

covid19_modelling

April 5, 2020

1 Modelling COVID-19 community spread

This model for the spread of COVID-19 would actually apply in general to any homogenous population in a closed infection environment where transmission is internal within that population (i.e. no source of infection from arrival of external travellers). The growth rate is modelled as a proportion of the population who are infected and contagious, which is less than the population who are infected since some infected people are still within the incubation period. The infected population is divided into known and unknown cases, with only the unknown proportion propagating the disease in the community. The known proportion is assumed to be isolating and therefore not transmitting.

The model is not statistical in the sense of modelling random sampling and random events in the population. All such details are approximated by continuous differential equations as functions of time. Constant values are used for incubation time, detection time, duration of illness and death rate. In the practical example, such parameters are tuned to the Australian context where necessary, and otherwise reckoned from broad variety of sources (not yet well referenced).

In the growth rate model, I include a scaling factor on the growth rate parameter that tapers the growth as the number of resolved cases approaches a certain population threshold. In the early growth phase, this scaling factor is essentially 1, so the model behaves like exponential growth. As the proportion of resolved cases becomes a significant proportion of the population, the growth rate will taper until the total number of infections (active + resolved) levels off at the assumed herd immunity level.

Mathematically speaking, we end up with a system of differential equations where derivatives of some variables depend on other variables at negative time delays. It is this feature that I believe is important in capturing the dynamics of the growth of the infection. The philosophy applied here is to model the true number of infections over time (even though in practice this is not known) and from that, derive observable quantities such as the known number of active infections over time, which we can test against observable data.

The growth rate parameter α is a key input and is not derived from any first principles. Instead, this parameter is tuned to fit known data to date, and then either held constant or computed dynamically according to some assumed policy or behavioural model to try to forecast the resulting model variables. For example, if social distancing was imposed on a particular date, then α could be reduced from that date onwards. Or α could be made to depend on the number of known cases, modelling reaction of government to the daily situation. However the actual values of α to use in such cases should be modelled on, or fitted to, actual data, so have limited predictive purpose. It is reasonable, however, to fit α to known data up to now, then run the model forwards to project the outcomes of the growth rate remaining constant.

1.1 Formulation of the model

1.1.1 Definitions

Definitions of the symbols used in the formulation.

$P(t)$ = total population, $P(t \leq 0) = P_0$

$N(t)$ = number of actual actively infected (both known and unknown), $N(t = 0) = N_0 = 1$, $N(t < 0) = 0$

$M(t)$ = number of immune (vaccinated or recovered), $M(t \leq 0) = M_0 = 0$

$D(t)$ = number of deceased, $D(t \leq 0) = D_0 = 0$

$G(t)$ = number of infected growth rate, $G(t \leq 0) = G_0 = 0$

$K(t)$ = number of known actively infected, $K \leq N$, $K(t \leq 0) = K_0 = 0$

t = time variable

α = growth rate parameter

f = fatality rate as a fraction of all (known and unknown) resolved cases

s = fraction of actively infected that become symptomatic, $0 < s \leq 1$

t_i = mean incubation period, time from initial infection until becoming contagious

t_c = symptom response and testing confirmation period

τ = mean resolution period to recovery or mortality, after which case is no longer active

h = "herd immunity" threshold as a proportion of total population

p_{hosp} = proportion of known active cases requiring hospitalization

The "symptom response and testing confirmation period" is the delay from becoming contagious to detection by testing (if symptomatic).

1.1.2 Dynamic equations

The growth rate G in simple exponential growth represents the rate of change in N per unit time from new infections and is proportional to N itself, i.e. it must look something like $\frac{dN}{dt} = \alpha N$. However this needs to be modified to account for three effects. Firstly, the growth rate is not proportional to N at the present instant, rather proportional to the infected who are contagious, which is $N(t - t_i)$. Secondly, the growth rate needs to be tempered as the number of infected or previously infected approaches the herd immunity level. Third, because the known cases K are in self-isolation, they reduce the number of available actively infected that propagate to others, so that the growth rate is proportional to $N - K$ (assuming 100% compliance with self-isolation). The modified growth rate is thereby formulated as

$$G(t) = \alpha \left(1 - \frac{N(t) + M(t)}{P(t)} \right) (N(t - t_i) - K(t))$$

During the growth phase when $N \ll P$ and $M \ll P$, this formula behaves as $G(t) \approx \alpha (N(t - t_i) - K(t))$.

The rate of change of the number of actually infected, $N(t)$, is positively contributed by $G(t)$, but negatively contributed by cases passing the resolution period τ . The balance of the rate of change of actively infected is therefore

$$\frac{dN}{dt} = G(t) - G(t - \tau)$$

The recovery rate is the complement of the death rate applied to the rate of new cases τ ago in time:

$$\frac{dM}{dt} = (1 - f) G(t - \tau)$$

The death rate is the fraction f of the infection rate τ ago in time:

$$\frac{dD}{dt} = f.G(t - \tau)$$

The rate of change of the population is the negative of the death rate:

$$\frac{dP}{dt} = -\frac{dD}{dt}$$

The rate of change of known active cases $K(t)$, like the rate of change of true active cases $N(t)$, consists of a growth term from new cases and a depletion term from expiry of cases older than τ . The growth term is based on G at a time delay of $t_i + t_c$. The balance of these terms gives:

$$\frac{dK}{dt} = s (G(t - t_i - t_c) - G(t - \tau))$$

Note that the expanded first term here is the new case rate, which is integrated below to produce the total known cases (active and resolved),

$$\text{new case rate} = s.G(t - t_i - t_c)$$

1.1.3 Derived quantities

The number of uninfected in the population as a function of time will be

$$U(t) = P(t) - N(t) - M(t)$$

The cumulative total of all actual cases as a function of time, which includes known and unknown active infections, recovered and deceased is

$$N_{tot}(t) = N(t) + M(t) + D(t)$$

The total number of known cases as a function of time will be

$$K_{tot}(t) = \int_0^t sG(t - t_i - t_c).dt$$

The number of hospital beds needed over time will be

$$B(T) = p_{hosp}K(t)$$

1.1.4 Invariants

Identifying invariants explicitly allows us to confirm that our solution implementation respects such invariants. If the invariant is violated, the implementation is definitely wrong. The invariant here is the total initial population P_0 which must be conserved through the following expressions:

$$\begin{aligned} P(t) + D(t) &= P_0 \\ U(t) + N(t) + M(t) + D(t) &= P_0 \end{aligned}$$

1.2 Summary of dynamic equations

Collecting the dynamic equations above, we have the following system to solve with time:

$$\begin{aligned} G(t) &= \alpha \left(1 - \frac{N(t) + M(t)}{P(t)} \right) (N(t - t_i) - K(t)) \\ \frac{dN}{dt} &= G(t) - G(t - \tau) \\ \frac{dM}{dt} &= (1 - f) G(t - \tau) \\ \frac{dD}{dt} &= f \cdot G(t - \tau) \\ \frac{dP}{dt} &= -\frac{dD}{dt} \\ \frac{dK}{dt} &= s (G(t - t_i - t_c) - G(t - \tau)) \end{aligned}$$

[]: