Introduction

This coursework is worth 100% and is due Thursday, 6 January 2022, 16:00 (but please always check what Canvas and Sussex Direct says!). Submission is electronic, via Canvas. Submission guidelines are identical to the formative coursework, namely: (a) Analytical work can be hand-written or typeset; (b) simulation code + analysis can be presented separately or integrated into a Jupyter / Colab notebook (whatever you prefer!). Unlike the previous coursework, there is no page limit **but** conciseness of exposition is always best!

In this coursework, you are going to engage in some proper modelling work. You are given a scenario along with a couple of ordinary differential equations that describe the operation of this system (in principle I could have asked you to derive them from the scenario description but I don't). The questions take you through a number of steps by which you will gain a detailed understanding of the system / model. Although all questions are related, it is possible to get results even if you find yourself unable to complete some of the steps.

Statement of problem

We consider a system of individuals (the nature of which is irrelevant) that can take two states A and B. At any point in time, an individual can only be in one state. There are no other states possible. We will denote by $[A]$ the number of individuals in state A and by $[B]$ the number of individuals in state B. There are N individuals in total. All individuals can be assumed to be in contact with all other individuals. At time $t = 0$, there are B_0 individuals in state B. All other individuals are therefore in state A. Individuals in state A turn into state B at a rate $\beta\frac{[B]}{N}$, i.e., proportional to the number of their neighbours that are in state B. Individuals in state B turn into state A at a fixed rate γ.

Now, in an ideal world, you should be able to write the two ordinary differential equations (ODE) that describe this system. To make sure that everyone can proceed with the coursework, I am providing the ODEs in the last page of this document. Before consulting that section, you are strongly encouraged to try and come up with your own formulation.

Questions

We start with some analytical work.

Analytical work

1. Using the fact that $[A] + [B] = N$ at all times, write $[\dot{B}]$ as a function of $[B]$, i.e., the expression should no longer involve $[A]$. This is your (so-called) *mean-field equation*.

   (2 marks)

2. Find the equilibria of the system and determine their stability. From now on, we will refer to the non-zero equilibrium as B^*. You may find it useful to write your results in terms of the following quantity $R_0 = \frac{\beta}{\gamma}$. Plot the phase portrait, i.e., $[B]$ vs $[A]$, identifying the equilibria and their stability (following convention described in the 2nd synchronous lecture of Unit 5).

   (8 marks)

3. Produce the bifurcation plot for this system, that is, plot the value of the equilibria as a function of R_0, with R_0 taking values from 0.1 to 5.0. For this question, the value of N is irrelevant (provided it's strictly positive) so use 1000 for example. This should be done using Python.

   (4 marks)

4. In this question, you are going to integrate $[\dot{B}]$ *analytically* to obtain an expression for $[B](t)$, i.e., an expression that gives the number of individuals in state B in time. It is rarely the case that this can be done but with this system, it is possible. You will do this in four steps:

   - Starting from the mean-field equation, factorise the right hand side by $[B]^2$, then write an expression for $\frac{1}{[B]^2}[\dot{B}]$.

     (4 marks)

   - Consider the following variable substitution: $y = \frac{1}{[B]}$. Using the chain rule, express $\dot{y}$ in terms of $[B]$, then derive a simple expression for $\dot{y}$, i.e., this expression should only involve y terms and parameters of the system. There shouldn't be any $[B]$ or $[A]$. However, it will be helpful to use B^* (calculated in the 2nd question) to simplify the expression.

     (4 marks)

- Integrate this equation. You should be able to do this without any help, but if help is needed, you should note that this expression looks very much like the equation we solved during a synchronous lecture in Unit 4, replacing λ and I by appropriate quantities. You can then use the result to derive an equation for $y(t)$. Please see short document summarising the derivation from the lecture.

  (4 marks)

- You can now produce a fully worked out expression for $[B](t)$ by remembering that $[B] = B_0$ at time $t = 0$.

  (4 marks)

  NB: In going through this question, do make sure to consider all scenarios possible regarding the value of R_0.

5. Using different values of B_0 (between 1 and N – briefly discuss the case $B_0 = 0$), plot solutions of $[B](t)$ for various values of R_0 between 0.1 and 5.0 (with $\gamma = 0.5$ for example). Confirm your expression for $[B](t)$ is correct by (a) verifying that it converges to B^* for large times t and (b) visually confirming agreement when integrating the mean-field equation using Euler (use Python). What happens when $R_0 = 1$? Speculate as to what this means. We will get back to this. For a given value of R_0, what happens when the value of γ changes? Provide a brief explanation.

   (11 marks)

Simulation work

Some more could be done analytically but let's now turn to simulations. For this part of the work, we will employ the Gillespie algorithm, an algorithm often used to simulate complex systems. There is something very important to understand here. The mean-field equation provided to you is *derived* from considering the interactions of a very large number of individuals. The Gillespie algorithm, instead, provides discrete simulations of the system with few individuals by explicitly simulating every transitions. In other words, a single run of the Gillespie algorithm represents one sample trajectory of all possible trajectories for this system. In principle, the average of a (sufficiently) large number of Gillespie runs should converge to the mean-field. This is what you will test here. **NB**: Gillespie provides a mathematically rigorous and computationally efficient alternative to agent-based modelling. You are not asked to compare the two approaches even if that would be an interesting thing to do (NB: but certainly **not** within this assessment!).

To make sure everybody can work, I provide a Python implementation of the Gillespie algorithm for this problem. Use the code although feel free to write your own if you feel

so inclined – you would need to read up on the Gillespie algorithm obviously. Function `gillespie_ABA` takes 5 arguments: N, B_0, β, γ and $Tmax$ (max simulation time) and produces three outputs: T, A and B which contain the time of each transition (when one, and only one, individual changed state, whatever that state was), the numbers of individuals in state A and in state B after that transition, respectively. **NB**: The Gillespie algorithm uses exponentially-distributed times to next event so each run will have a different timeline T of events. This won't be apparent to you if plotting with lines but will matter in what comes below.

1. Explore the behaviour of the system when considering suitably chosen scenarios, i.e., focus on the limit cases (e.g., small R_0, large R_0 and $R_0 = 1$; small N, large N; small B_0, large B_0). For each scenario, use the code provided to generate many realisations of the stochastic process. Plot all realisations on a single plot. Make relevant **qualitative** observations.

   (10 marks)

2. For each scenario, calculate the average (and standard deviation) of the realisations. Here, you are going to face a problem linked with the *nota bene* from the introductory paragraph. You will need to think of a solution and implement it. Superimpose the average (and error bars) to the realisations. Use a larger line width for visibility.

   (11 marks)

3. Finally, superimpose the mean-field solution. Again, use a larger line width and different colour for visibility. Describe and interpret agreement between average of stochastic realisations and mean-field in relation to the choice of parameters. In this question, using $B_0 = 1$ (i.e., only one individual in state B at $t = 0$) can help exacerbate the differences and help you think about what is happening. You may want to refer to your bifurcation plot.

   (8 marks)

4. *[Slightly challenging question]*: Consider 100 replications for $N = 1000$, $\beta = 0.51$, $\gamma = 0.5$ and 100 replications for $N = 1000$, $\beta = 0.95$, $\gamma = 0.5$. You should notice a substantial difference in agreement between the mean-field and the average of the stochastic realisations depending on which scenario is considered. How could you improve agreement for the scenario with the poorest agreement. Please note: **The difference in B^* is not the quantity of interest here**. Rather you should think about why agreement is so poor. This does not actually involve analytical work. An excellent answer would see you implement your proposed solution and provide evidence of improved agreement.

   (12 marks)

A bit of critical thinking

In this last part of the coursework, you are invited to think a bit more about what you have just done.

1. So far the brief has provided no information whatsoever regarding what the states A and B are and what the individuals are. Thinking about what is happening in this system, provide at least one example of real-world scenario to which this model could apply. Bonus points will be given for any answer that provides two examples, one in which the equilibria are of interest and the other in which the critical regime (when $R_0 = 1$) is of interest. Either way, what is the benefit of being able to model the system?

   (8 marks)

2. *[Very tough question]*: The model provided implicitly assumes that all individuals are potentially in contact with each other. What would be a more likely scenario? What changes would have to be made to the **code** of the Gillespie algorithm in order to include such a scenario? If you are able to do this, do it. Then, speculate as to what could affect the results observed in the previous questions. If you feel so inclined, demonstrate it experimentally. **NB**: Only 10 marks have been given to this question. However, anyone managing it successfully would receive an extra 10 marks for the assignment (with the total mark capped to 100 obviously).

   (10 marks)

Total marks: 100.

The ODEs

The system is fully described by the following two ordinary differential equations:

$$[\dot{B}] = \beta \frac{[B]}{N}[A] - \gamma[B]$$

$$[\dot{A}] = -\beta \frac{[B]}{N}[A] + \gamma[B]$$

where β and γ are strictly positive real numbers. The first equation provides the rate of change of the number of individuals in state B and says that it is the sum of the number of individuals in state A that become B with a rate $\beta \frac{[B]}{N}$ (proportional to the number of other individuals in state B) minus the number of individuals in state B that turn to state A with rate γ. I leave it to you to interpret the second equation. Summing both equations lead to 0 which is consistent with the fact that the total population $[A] + [B] = N$ at all times. To complete the description we need the initial conditions. We will call B_0 the number of individuals in state B at $t = 0$. The number of individuals in state A is therefore $N - B_0$.