Mathematics 4MB3/6MB3 Mathematical Biology
2019 ASSIGNMENT 1

This assignment is due in class on **Monday 21 January 2019 at 11:30am**.

General Notes

(i) You must show your work and justify all your assertions carefully in order to obtain credit for your work.

(ii) Your solutions must be typeset using LaTeX. Carefully read all the technical comments below before beginning to work on your solutions document.

(iii) Any documents you hand in must state your group name and the names of all members of the group.

(iv) Always state the units of quantities you use or state as results. The reader shouldn't have to guess whether you chose your time unit to be years or days or whatever.

(v) In addition to a hardcopy, your group must submit a final `pdf` file by e-mail to the instructor together with all source files (*e.g.,* `.tex` for LaTeX, `.R` or `.Rnw` for R, and `.ode` for `XPPAUT`). E-mail a compressed folder (zipped or tarballed archive) containing all the source files (and no unnecessary files!), and *separately* include the final `pdf` as well. The instructor must be able to reproduce your `pdf` file from the source files you submit. Make sure each member of your group has successfully reproduced the final `pdf` given the compressed folder before you submit it by e-mail to the instructor.

(vi) The subject line of the e-mail in which you submit your assignment must be

`Math 4MB3 Assignment 1: Group Name`

Please copy and paste this line into the subject line of your e-mail and then replace "`Group Name`" with your group name. This will reduce the chance for typos that make it harder for me to find your e-mail by subject. Remember to attach both the compressed source file folder and the compiled `pdf` file to your e-mail message. Be sure to copy all members of your group on this e-mail message so the instructor can easily reply to the whole group.

Technical Comments (read carefully!)

The comments below apply to all work in this course. If you are not creating any graphs with R for this assignment then some of the comments won't seem relevant now, but they will be very relevant when you start making graphs in R.

1. Change the default font size from 10 point to 12 point. The default font size is set in the first line of your LaTeX document: `\documentclass[12pt]{article}`.

2. The LaTeX *preamble* used for this question sheet can be downloaded from the Assignments page on the course web site. You should `\input{4mbapreamble.tex}` in the preamble of your solutions document (*i.e.,* between `\documentclass[12pt]{article}` and `\begin{document}`). This file addresses the following issues and many others:

   - You will need to keep referring to $\mathcal{R}_0$ in your solutions. To make your life easier, define a new `\R` command like so:

     ```
     \newcommand{\R}{{\mathcal R}}
     ```

     Then if you type `$\R_0$` in your LaTeX source file you will get $\mathcal{R}_0$ in your `pdf` output.

   - Please use the logo to refer to the `R` language. Do this by defining this as an `\Rlogo` macro in your LaTeX preamble like so:

     ```
     \usepackage{xspace}
     \newcommand{\Rlogo}{\protect\includegraphics[height=2ex,keepaspectratio]
       {images/Rlogo.pdf}\xspace}
     ```

     Note that you will also need the image file `Rlogo.pdf`, which can be downloaded from the Assignments page on the course web site. Place `Rlogo.pdf` in an `images` subfolder of the folder where your script lives.

   - Define a comment macro as follows:

     ```
     \usepackage{color}
     \newcommand{\de}[1]{{\color{red}{\bfseries DE:} #1}}
     ```

     This macro allows me to add comments easily in your LaTeX document. For example, the LaTeX code `\de{What a great idea!}` yields **DE:** What a great idea!

3. Good notation is important for making your documents easily comprehensible. Given the LaTeX definitions of `\R` and `\Rlogo` above, it is easy to distinguish the removed class (R), the reproduction number ($\mathcal{R}$) and the programming language (). Always do this! Not doing so is sloppy and confusing to readers of your work. Also pay close attention to any other potentially confusing notational issues.

4. Run your source file(s) through a spell checker. Don't submit work that has typos or spelling errors. There is information about spell-checkers that work with LaTeX on the Software page of the course web site.

5. File and folder names: Please avoid spaces and non-alphanumeric characters in file names and folder names (*e.g.,* do not use `\`, `&`, `#`, `!`, `^`, `%`, `$`, `*` or brackets, though the underscore `_` is fine). For example, instead of naming an script

   ```
   My Group's 4MB3 Assignment #1 Question 2(b).R
   ```

   choose the file name to be something like

   ```
   MyGroup_4MB3_A1_2b.R
   ```

A sensible filename for a LaTeX document that contains your group's solutions to Assignment 1 would be

`MyGroup_4MB3_A1.tex`

and a sensible name for the folder that contains all files for the assignment would be

`MyGroup_4MB3_A1`

6. Do not include any absolute file paths in your code. For example, you should *not* refer to `C:\MyName\Documents\MyFavoriteCourse\datafile.csv` in your code. Keep files you need to read in a subfolder of the folder where you are executing your scripts. Make sure everyone in your group can produce the final `pdf` file without altering any of the files in any way.

7. LaTeX needs you to use single opening and closing quotes. A double quote (`"`) is always interpretted as a closing double quote (”). Thus, if you say `"quoted words"` you get ”quoted words”, whereas ``` ``quoted words'' ``` yields “quoted words”.

8. Always use math mode to typeset math. For example: `f(x)` yields f(x) whereas `$f(x)$` yields $f(x)$.

9. Use `typewriter type` when referring to filenames. For example, `{\tt filename.R}` yields `filename.R`. (An alternative is `\texttt{filename.R}`.)

10. When including images in a LaTeX document, it is best to save the images from R as `pdf`. If you save as `png` or `jpg` then LaTeX will still be happy to display them, but the quality of the image is reduced unnecessarily.

11. The LaTeX command for the “much less than” symbol is `\ll`. Don't use `<<` for this. For example: `$a<<b$` yields $a << b$ whereas `$a\ll b$` yields $a \ll b$. In general, if you typeset some math and it doesn't look like what you would expect to see in a professionally typeset math book then you can be certain you're not using the intended LaTeX syntax.

12. Always use LaTeX's built-in function names, *e.g.,* `$\log(t)$` correctly yields $\log(t)$ whereas `$log(t)$` yields $log(t)$.

13. Avoid explicit spacing commands in LaTeX if possible. For example, if you want to have space between each paragraph of your document, then don't include an extra line break at the end of each paragraph (which could be done via `\\`). A better approach is to set the value of the paragraph skip in the preamble (via `\parskip=10pt`, for example). Then you can change the spacing easily in the entire document, and if you later want to use a different format then there won't be explicit spacing commands lurking around to wreck your output. Even setting the `\parskip` explicitly is considered undesirable, because it will override directives in a LaTeX style file, but if you can keep formatting changes to the preamble, it will make your life simpler.

14. Every R script should begin with an opening comment explaining what the script does. What is the purpose of the script? What output will be produced when it is run?

15. Take advantage of R's vector syntax wherever convenient. For example, if setting line styles for a sequence of lines in a plot or legend, rather than `lty=c(1, 2, 3, 4, 5, 6)` say `lty=1:6`.

16. Wherever appropriate, use R's assignment operator (`<-`) rather than equals (`=`).

17. Your LaTeX code must compile without any errors. Producing a pdf file is not adequate. Others must be able to reproduce the pdf without getting any R or LaTeX errors.

18. To make your R code readable, it is very important that you indent appropriately. If you are using `Emacs` then tab will indent the current line of code according to standard convention.

19. Make sure figures appear where you want them. The `figure` environment has options that allow you to control placement in the document.

1 Analysis of the SI model

The SI model can be written

$$\frac{dI}{dt} = \beta I(N - I)\,, \tag{1}$$

where I denotes prevalence and $N = S + I$ is the total population size.

(a) Prove that the endemic equiliibrium (EE) is a globally asymptotically stable (GAS) equilibrium by finding an appropriate Lyapunov function. Note that "global" here refers to all biologically relevant initial conditions except the (unstable) disease free equilibrium (DFE).
Hint: Lyapunov functions often look paraboloidal.
Note: Notions of stability and Lyapunov functions were discussed in Math 3F03 Lecture 27 in 2013 (http://www.math.mcmaster.ca/earn/3F03).

(b) In class we proved only stability of the EE, not asymptotic stability. Prove GAS "directly" in two distinct ways:

   (i) find the exact solution of the model and take the limit as $t \to \infty$, and conclude that every solution that starts in the interval $(0, N)$ converges to the EE (this approach works only in situations where you can find the exact solution);

   (ii) given $\epsilon > 0$, prove that for any $I(0) \in (0, N)$ $\exists t < \infty$ such that $I(t) \in [N-\epsilon, N)$ and use this to establish GAS. (Do not use your exact solution in this part; the point is to use an approach that also works for models that cannot be solved exactly.)

2 Analysis of the basic SIR model

The basic SIR model is specified by the following system of differential equations.

$$\frac{dS}{dt} = -\mathcal{R}_0 SI \tag{2a}$$

$$\frac{dI}{dt} = \mathcal{R}_0 SI - I \tag{2b}$$

$$\frac{dR}{dt} = I \tag{2c}$$

The state variables S, I and R are the proportions of the population that are susceptible, infectious and removed, respectively. The parameter $\mathcal{R}_0$ is the basic reproduction number. The time unit has been chosen to be the mean infectious period for convenience.

(a) A quantity of some practical importance is the ***peak prevalence*** of disease in the population, *i.e.,* the maximum proportion of the population that is simultaneously infected. Find an exact expression for the peak prevalence, given initial conditions (S_0, I_0). Why might a public health official want to know this quantity?

(b) It would be helpful to have an analytical expression for the solution of the model. Most valuable would be a formula for $I(t)$, which is most closely related to time series data. You probably will not find a formula for $I(t)$ (extra credit if you do!!) but it is definitely possible to find an exact expression that relates R (proportion removed) and t (time).

   (i) Find such an expression. *Hint:* Combine the equations for dS/dt and dR/dt into one equation that can be solved for S as a function of R. Then recall that $S+I+R = 1$ and use the dR/dt equation again. *Note:* You will end up with an expression for t as a function of R, not R as a function t.

   (ii) Use your expression for $t(R)$ to find an expression for the time at which peak prevalence will occur. Why might this be useful?

   (iii) How could your expressions be used to compare with the time series for pneumonia and influenza in Philadelphia in 1918? (Don't actually do it; just clearly explain your thinking including any assumptions you are making.) Would you advise your assistant who just graduated with a degree in math and biology to do this (to help you prepare your report for the public health agency)? Why or why not?

   (iv) Is it possible to find an exact analytical expression for t as a function S?

(c) Prove that all solutions of the basic SIR model approach $I = 0$ asymptotically, and explain why this makes biological sense. *Hint:* Is the function $L(S, I) = I$ a Lyapunov function? Read the Notes on Lyapunov functions below.

(d) Find and classify the stability of all equilibria of the basic SIR model.

Notes on Lyapunov functions

Consider Lyapunov's Stability Theorem as stated in Math 3F03 Lecture 28 in 2013:

Theorem 1 (Lyapunov's Direct Method)**.** *Consider an equilibrium X_* of $X' = F(X)$ and an open set $\mathcal{O}$ containing X_*. If $\exists$ a differentiable function $L : \mathcal{O} \to \mathbb{R}$ such that*

(a) $L(X_*) = 0 \quad \text{and} \quad L(X) > 0 \quad \forall X \in \mathcal{O} \setminus \{X_*\}$ *(L positive definite on $\mathcal{O}$)*

(b) $\dot{L}(X) \le 0 \quad \forall X \in \mathcal{O} \setminus \{X_*\}$ *($\dot{L}$ negative semi-definite on $\mathcal{O}$)*

then X_* is stable and L is called a ***Lyapunov function****. If, in addition,*

(c) $\dot{L}(X) < 0 \quad \forall X \in \mathcal{O} \setminus \{X_*\}$ *($\dot{L}$ negative definite on $\mathcal{O}$)*

then X_* is asymptotically stable and L is called a ***strict Lyapunov function****.*

Theorem 1 can be generalized for analysis of stability of sets more complicated than isolated equilibria, such as periodic orbits or line segments. If you think through the proof of the theorem above (*e.g.,* [1, §9.2, theorem stated on p. 193 and proved on p. 196]), you should be able to convince yourself that the proof still works if the equilibrium X_* is replaced by any *closed forward-invariant set* (often simply called a *closed invariant set*). This observation allows us to state the following more general theorem.

Theorem 2 (Lyapunov's Direct Method for Closed Invariant Sets)**.** *Consider a closed invariant set $\mathcal{C}$ of $X' = F(X)$ and an open set $\mathcal{O}$ containing $\mathcal{C}$. If $\exists$ a differentiable function $L : \mathcal{O} \to \mathbb{R}$ such that*

(a) $L(X) = 0 \ \forall X \in \mathcal{C} \quad \text{and} \quad L(X) > 0 \quad \forall X \in \mathcal{O} \setminus \mathcal{C}$ *(L positive definite on $\mathcal{O}$)*

(b) $\dot{L}(X) \le 0 \quad \forall X \in \mathcal{O} \setminus \mathcal{C}$ *($\dot{L}$ negative semi-definite on $\mathcal{O}$)*

then $\mathcal{C}$ is stable and L is called a ***Lyapunov function****. If, in addition,*

(c) $\dot{L}(X) < 0 \quad \forall X \in \mathcal{O} \setminus \mathcal{C}$ *($\dot{L}$ negative definite on $\mathcal{O}$)*

then $\mathcal{C}$ is asymptotically stable and L is called a ***strict Lyapunov function****.*

Note in the above theorems that open sets are defined relative to the subset of interest; in our case this subset is $\Delta = \{(S, I) : S \ge 0,\ I \ge 0,\ S + I \le 1\}$, not all of $\mathbb{R}^2$. An open set of Δ is a set of the form $U \cap \Delta$ where U is an open set of $\mathbb{R}^2$. (These sets are said to be open in the ***relative topology*** on Δ.) In particular, note that Δ is *open* as a subset of itself, in spite of the fact that it is *not open* as a subset of $\mathbb{R}^2$, whereas Δ is closed in both the relative topology on Δ and the usual topology on $\mathbb{R}^2$.

References

[1] Hirsch MW, Smale S, Devaney RL. Differential equations, dynamical systems, and an introduction to chaos. 3rd ed. Waltham, MA: Academic Press; 2013.

— END OF ASSIGNMENT —

Compile time for this document: December 26, 2018 @ 20:59