Problem Set 6

Assigned Nov 21 2023, Due Dec 7 2023

Problem 1 (20 points): The goal of this problem is to take some informal descriptions of model inference problems and pose them as optimization or sampling problems. For each informal description, you should formalize the problem by specifying a set of variables encoding the model and defining a likelihood or objective function you could use to fit parameters to the system. Define whatever additional constants you need to provide a full specification of the system. Note that there may be many right answers to any given problem.

a. Suppose we are trying to learn a model for how the expression level of a gene G_* is affected by the expressions of a set of other genes $G_1, \ldots, G_n$ presumed to be upstream of it in a regulatory network. We first assume that G_* is directly regulated by $G_1, \ldots, G_n$ and that we believe a linear function will relate their activities. We have a set of training data consisting of steady-state expression levels of all $n+1$ genes in a set of m conditions.

b. Suppose we again want to learn a model of how G_*'s expression depends on those of $G_1, \ldots, G_n$ but instead that we believe this to be a dynamic system described by a set of ordinary differential equations. Developing a model requires learning the rate constants of the ODEs. Our training data consists of sets of time points at which we know the expression levels of $G_*, G_1, \ldots, G_n$.

Problem 2 (20 points): Consider the following loss function on vectors w in R^4:

$$L(w) = w_1^2 + 2w_2^2 + w_3^2 - 2w_3w_4 + w_4^2 + 2w_1 - 4w_2 + 4 \tag{1}$$

(a) What is $\nabla L(w)$?
(b) Suppose we use gradient descent to minimize this function, and that the current estimate is w = (0, 0, 0, 0). If the step size is η, what is the next estimate?
(c) What is the minimum value of L(w)?
(d) Is there is a unique solution w at which this minimum is realized?

Problem 3 (40 points): In class we discussed how to build simulations of the Birth-death Moran process on networks to determine the probability of fixation of a new mutant with selective coefficient s appearing in a random node on the network (same setup as last homework, mutant has (1+s) higher relative probability of birth than wild-type population). Simulate the death-birth process we also mentioned in class: now, at every time step, we first randomly select a node for death, a neighbor node to reproduce, proportional to fitness, followed by the replacement of the death node by the offspring of the birth node. Discuss if and how the probability of mutant fixation is different than the one we saw for the Birth-death process discussed in class. If you observe different fixation probability (for the same parameters), can you provide intuition and discuss why that may be?

***Problem 4 (20 points): The SZR model on networks**:

The models discussed so far have ignored the fact that that an individual comes into contact only with its network-based neighbors in the pertinent contact network. We have assumed homogenous mixing, which means that an infected individual can infect any other individual. To accurately predict the dynamics of an epidemic, we might need to consider the precise role the contact network plays in epidemic phenomena. For this problem, based on the work of Munz et al. [1], we will model Zombie attacks on generalized random networks (the paper is below). There are three states: S, susceptible, Z, zombie, and, R, removed. For the random mixing model, the differential equations are

$$\frac{dS}{dt} = \theta - \beta SZ - \gamma S \tag{2}$$

$$\frac{dZ}{dt} = \beta SZ + \xi R - \alpha SZ \tag{3}$$

$$\frac{dR}{dt} = \gamma S + \alpha SZ - \xi R, \tag{4}$$

where
• θ is the birth rate of new susceptibles;
• β is the rate at which susceptibles who bump into zombies become zombies;
• γ is the background, non-zombie related death rate for susceptibles;
• ξ is the rate at which the dead (removed) are resurrected as zombies;
• and α is the rate at which susceptibles defeat zombies (through traditional methods shown in movies).
For our purposes, consider a random network containing completely susceptible individuals and discrete time updates. We'll now think of the parameters above as probabilities, and ignore birth processes (θ). Assume that in each time step, all edges convey interactions, meaning each individual interacts with each of their neighbors. Simulate this model for different types of networks and see if you can determine for which network types and spreading parameters will local zombification take off (i.e., grow exponentially, at least in the short term), given one randomly chosen individual becomes the first zombie? (The long term dynamics will likely be complicated so we will focus on the initial dynamics.) See http://www.wired.com/wiredscience/2009/08/zombies/ for more information/enjoyment.

P. Munz, I. Hudea, J. Imad, and R. J. Smith? When zombies attack!: Mathematical modelling of an outbreak of zombie infection. In J. M. Tchuenche and C. Chiyaka, editors, Infectious Disease Modelling Research Progress, pages 133–150. Nova Science Publishers, Inc., 2009.

*Recall that the starred problems are only for the 02-712 students.