Homework #4

ES_APPM 346-0

Due Mar. 6, 2021

Introduction

This project is based on a paper by Mandal and Banerjee that illustrates how stochastic fluctuation affects can change the outcome of a model, in this case a population model concerning tilapia fish and pelicans.

In this model there are tilapia fish that are subject to an infectious disease, and the pelicans that feed on the fish. Infected fish are rendered sterile by the disease, so only uninfected (also called susceptible) tilapia are able to reproduce, but all tilapia compete for the same food resources. Transmission of the disease is by contact between fish. Let $S(t)$, $I(t)$, and $P(t)$ be the populations of susceptible tilapia, infected tilapia, and pelicans respectively. The proposed model for this eco-system is given by the differential equations:

$$\frac{dS}{dt} = rS\left(1 - \frac{S+I}{k}\right) - \lambda SI$$
$$\frac{dI}{dt} = \lambda SI - \mu I - \frac{mIP}{I+a}$$
$$\frac{dP}{dt} = \frac{\theta IP}{I+a} - \delta P$$

Let's take a look at the different terms in this system of equations. The equation for susceptible fish starts with a term that comes from the logistic equation for population that you may have seen in EA4. In that context, the equation is written as

$$y' = \rho y(K - y)$$

where ρK is the reproduction rate, and K is called the *carrying capacity*. The carrying capacity is a phenomenological construction that makes it so that the steady state population in the model is given by K. This type of model is meant to reflect that the food resources in the environment is not unlimited, but is capable of of supporting only a finite population. It's a way of encapsulating that complex subsystem without having to also model the food supply. In our tilapia model, r is playing the role of ρK, and k is playing the role of K. The total number of fish is given by $S + I$ because both infected and susceptible will consume the local food supply.

The second term in the susceptible equation is the mechanism for disease transmission. It requires contact between an infected fish and a susceptible fish. The rate of infection per contact is contained in the parameter λ. When the infection occurs, the newly infected fish are moved from the susceptible to the infected. That's why there is a negative sign for the λSI term in the first equation, and a positive sign in the second equation.

Infected fish are at a higher risk for death due to the disease. This process is included in the model for the infected fish as a $-\mu I$ term where μ represents the higher death rate for infected fish.

The model assumes that healthy fish are able to maintain their buoyancy well, and hence are able to stay in deep enough water to avoid falling prey to pelicans. However, infected fish may start to stray toward the surface due to their illness. Let's look at the term $I/(I + a)$. When there are relatively few infected fish, then the number that surface will be some fraction of them given by $\approx I/a$. If there are lots of infected fish, and they flood the surface of the water, there is a limit to the rate at which pelicans can gorge themselves so, when $I >> a$, then $I/(I + a) \approx 1$ and we can interpret m as a maximum rate of fish consumption by

pelicans. Of course, the pelicans have to be present to actually consume the fish, and hence we get the pelican consumption rate of $-mIP/(I + a)$.

Of course, it requires more than one fish to create a new pelican, so we have to give the rate of pelican reproduction as proportional to the rate of fish consumption, but at a lower rate. Thus, the pelican population increases by the rate $\theta IP/(I + a)$. Finally, pelicans will die according to a given rate, hence the final term $-\delta P$. Since in the case of pelicans we are also accounting for the food source, we don't employ a logistic term as we did for the fish population.

When we look at a system like this, an important question to ask is where are the critical points. Two critical points are rather easy to spot based solely on the physical interpretation of the system. One is at the origin, where obviously if there are no fish, and no pelicans, the system will remain as it is. Another obvious one is when the fish population is uninfected and there are no pelicans. In that case, the tilapia population will be at carrying capacity k and remain there. There are two additional critical points, you will be asked to compute one of them, the other is given by

$$S = k - \frac{\delta a(k\lambda + r)}{(\theta - \delta)r}, \quad I = \frac{\delta a}{\theta - \delta}, \quad P = \frac{a\theta(r(\theta - \delta)(k\lambda - \mu) - \delta a\lambda(k\lambda + r))}{rm(\theta - \delta)^2}. \tag{1}$$

Reality is not so kind as to be completely smooth. Many of the processes we've discussed are actually subject to noise. The model with noise is given below:

$$dS = \left[rS\left(1 - \frac{S + I}{k}\right) - \lambda SI\right] dt + \sqrt{rS}\, dW_1 - \sqrt{\frac{r}{k}S(S + I)}\, dW_2 - \sqrt{\lambda SI}\, dW_3$$

$$dI = \left[\lambda SI - \mu I - \frac{mIP}{I + a}\right] dt + \sqrt{\lambda SI}\, dW_3 - \sqrt{\frac{mPI}{I + a} + \mu I}\, dW_4$$

$$dP = \left[\frac{\theta IP}{I + a} - \delta P\right] dt + \sqrt{\frac{\theta IP}{I + a}}\, dW_5 - \sqrt{\delta P}\, dW_6$$

Essentially, all the terms present in the deterministic model have multiplicative noise added to them. This requires some extra calculations involving checking that the covariance of the noise terms are handled properly. In this stochastic model, note that each of the dW_j are independent Wiener processes. The noise process dW_3 appears twice, and that is meant to be taken as the same random sample.

In this project, you will explore both the deterministic system, and the stochastic system to see how adding noise affects the system.

Written Assignment

1. Let's play modeler. For each of these processes, explain how you might modify the deterministic system to account for the given process. In other words, write down the new system of equations that you would use to model the modified system. Treat these as separate modifications, don't combine them.

   (a) Pelicans have an additional constant food supply that is available beside the tilapia.

   (b) Infected tilapia are able to reproduce at the same rate as healthy tilapia and all the offspring are also infected at birth. (Hint: don't forget to think about the carrying capacity)

   (c) Same as 1b, except half the offspring are infected, and half are healthy.

2. Determine the stability of the critical points $S = I = P = 0$ and $S = k$, $I = P = 0$. Remember, a critical point is stable if all the eigenvalues of the Jacobian have negative real part. If *any* of the eigenvalues have positive real part, the critical point is unstable.

3. Find the critical point where the number of susceptible fish is zero, but the number of infected fish is not zero. Explain why this critical point is not in the feasible set (i.e. the critical point is not physically possible).

4. Use Matlab to determine the location and the stability of the critical point (1) for the cases of $k = 400$, $\lambda = 0.06$, $\mu = 3.4$, $m = 15.5$, $a = 15$, $\theta = 10$, $\delta = 8.3$, and

   (a) $r = 7$,

   (b) $r = 24$

Programming Assignment

1. I have provided Runge-Kutta 4 method to solve the deterministic system of equations. The code is available on canvas as two files, an rk_sip.m file and a F_SIP.m file. You will have to update this file with the solution to the critical point in problem 4 above. Validate the method by doing a convergence check using the time interval $0 \le t \le 50$ and $N = 1000$, 2000, and 4000 steps. Plot the solution for the following two cases for the time interval $0 \le t \le 50$ and using initial conditions $S_0 = 100$, $I_0 = 80$, and $P_0 = 20$. If your results are stored in the variables S, I, and P, the critical point (1) is stored in Sc, Ic, Pc as computed in Problem 4 above, and time is stored in t, then the data should be plotted using the following code (included in the file):

```
subplot(2,1,1);
plot(t,S,'r-',t,I,'g-',t,P,'b-',[t(1),t(end)], ...
        [Sc,Sc],'r:',[t(1),t(end)],[Ic,Ic],'g:',[t(1),t(end)],[Pc,Pc],'b:');
legend('S','I','P');
subplot(2,1,2)
plot3(S,I,P,'k-',Sc,Ic,Pc,'bo');
xlabel('S'); ylabel('I'); zlabel('P');
```

   The first plot is a simple graph of S, I, and P over time, while the second plot is a phase plot in (S, I, P)–space. Do this for each of the following list of parameters:

   (a) $r = 7$, $k = 400$, $\lambda = 0.06$, $\mu = 3.4$, $m = 15.5$, $a = 15$, $\theta = 10$, $\delta = 8.3$.

   (b) $r = 24$, $k = 400$, $\lambda = 0.06$, $\mu = 3.4$, $m = 15.5$, $a = 15$, $\theta = 10$, $\delta = 8.3$.

2. Use the Euler-Maruyama method to solve the stochastic system of equations and using the same parameters values as in Problem 1. (Note: I would ordinarily recommend using a higher order method, but for systems of equations, it requires multiple integrals of noise terms that are beyond what we can

cover in this class, so we'll stick with the easy method.) Plot a typical sample path. Every time you solve this system, you will get a different answer, so choose a sample path that is most representative of the *many* trials you take. In terms of the ecosystem, what happens in the stochastic case, that doesn't seem to happen in the deterministic case? Does anything non-physical or non-sensical happen? Explain what you think is happening and why it didn't happen in the deterministic case.

3. In theory, it is possible that ΔW can be huge and negative, though with very low probability. If the model is supposed to maintain non-negative values, then what do we do? One solution, of course, is to truncate the path so that if, e.g. the number of susceptible tilapia goes negative, we just cut it off at zero, but that would violate the mass balance of the system of equations. Alternatively, one strategy is to decide that too large of a time-step must have been taken, and cut the time step down temporarily. Thus, your main update loop should look like this, where dt0 is the base time step size, and t is the variable tracking the time steps:

```
while  t(end) < T
        dt = min([T−t(end), dt0]);
        dW = sqrt(dt) * randn(6,1);

        % Compute  S_{n+1},  I_{n+1},  P_{n+1}  using  step  size  dt

        while  min([S(n+1),I(n+1),P(n+1)]) < 0
                dt = dt/2;
                dW = dW/sqrt(2);

                % Recompute  S_{n+1},  I_{n+1},  P_{n+1}  using  new  step  size  dt

        end
        t = [t,t(end)+dt];
        n = n+1;
end
```

Doing this will keep all your creatures alive. This solver with the patch for correcting against negative populations is the one you will turn in as code. Call it

```
function [t,x] = sde(K, T, dt0, x0)
```

where K is a vector of parameters in the order: r, k, λ, μ, m, a, θ, and δ. The vector x, and the initial data x0 should be the vector of species in the order: S, I, and P. The parameter T will be the terminal time, and the parameter dt0 will be the default time step size, used whenever possible assuming no species populations go negative. As usual, you will submit this code via email to me. For your write-up also submit three representative plots of sample paths (all three species on one plot, one plot for each sample path). Again, use the same parameters as for Problem 1 except only go to time $T = 10$ (for a total of 6 plots, 3 for each of the parameter sets), and use dt0= 10^{-4}.

4. Using your function sde(), run at least 100 sample paths for the case $r = 7$ and record the populations of each of the species at time $T = 10$. Plot a histogram of your data and compute the expected value, i.e. the mean, and the variance of the populations at that time. You may find the Matlab functions hist, mean, and var helpful in this effort. For the hist function in Matlab, use 20 bins.

5. (**Extra Credit**) For this problem use the parameter values in Problem 1b, but varying the value of a in the range $1 \leq a \leq 30$. For what value of a will the pelicans have the best expected population at time $T = 10$? To complete this task, you must compute the expected value of the random variable P_{10} as a function of a. This means that for each value of $a = 1, 2, 3, \ldots, 30$, you will need to take the mean value of P_{10} over many sample paths. To get a decent approximation, you will need to take a minimum

of 400 sample paths for each value of a to estimate the expected value $E(P_{10}(a))$. In other words, you will need to run your simulation a total of $30 \times 400 = 12,000$ times. Plot $E(P_{10}(a))$ as a function of a, and then identify the value of a that results in the highest expected pelican population. This may be a *very* expensive calculation, so only do this if you are really curious. This illustrates the true challenge of solving stochastic differential equations. It's not that computing any single sample path is any kind of challenge, it's the number of sample paths required in order to get a reliable, reproducible result that is the challenge.