Homework #1

Sam Fleischer

October 7, 2016

Problem 1

The whooping crane has been the target of protection efforts by the National Audubon Society and the federal governments of the US and Canada since 1941. There is one wild breeding population that winters at the Arkansas National Wildlife Refuge on the coast of Texas. The size of this population from 1941 to 1989 was studied by Dennis et al. (1991) in their Ecological Monographs paper Estimation of growth and extinction parameters for endangered species. The data set used by Dennis et al. is in the cranes.csv data file in the resources folder at SmartSite.

(a) Assuming the log population numbers can be modeled by

$$\log N(t) = \log N(0) + rt + \text{"observation error"}$$

where the "observation error" is normally distributed, use linear regression to estimate $\log N(0)$ and r. Use this model to estimate $N(2011)$.

(b) Assuming the log population numbers can be modeled by

$$\log \frac{N(t+1)}{N(t)} = r + \text{"process error"}$$

where the "process error" is normally distributed, use linear regression to estimate r. Use this model to estimate $N(2011)$.

(c) Do an online search to find an estimate for the *wild* whooping population size in August 2011. How do your model predictions compare?

(a) The following code snippet was used to load the CSV file and run a linear regression on the log population numbers.

```
1   # Load the file
2   cranes = read.csv("cranes.csv")
3
4   # Set local variables to access the data
5   Count = cranes$Number
6   Year = cranes$Year
7
8   # Plot Count vs. Year
9   plot(Year, Count, pch=4)
10
11  # Take the logarithm of the population numbers.
12  logPop = log(Count)
13
14  # Perform a linear regression
15  regression = lm(logPop~Y)
16
17  # Add the exponential (line in log-space) to the plot
18  lines(Year,exp(regression$fitted.values))
19
20  # Print the summary of the regression
21  summary(regression)
```

The output from the summary command shows $\log N(0) \approx -67.695$ and $r \approx 0.036390$, so $\log N(t) \approx -67.695 + 0.036390t$, which then implies

$$N(t) \approx N_0 \exp[0.036390t] \qquad \text{with} \qquad N_0 = \exp[-67.695].$$

Below is the plot generated from the code snippet. Using this linear regression, $N(2011) \approx 241.12$.
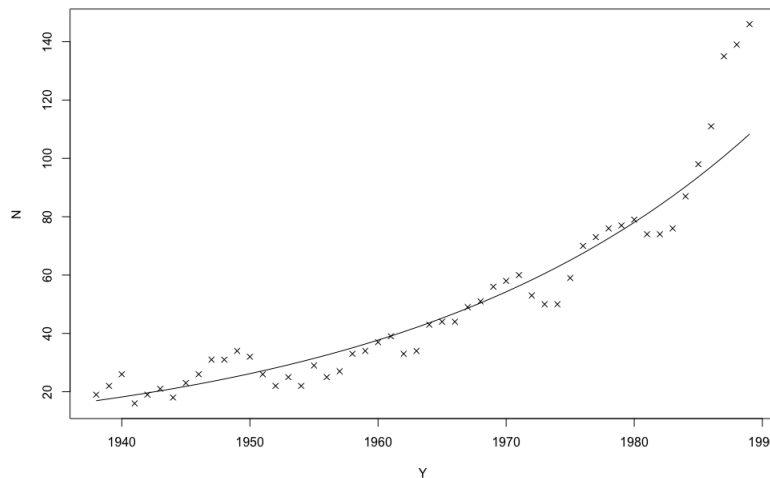
(b) The following code snippet was used to load the CSV file and run a linear regression on the ratio of successive log population numbers.

```
1  # Load the file
2  cranes = read.csv("cranes.csv")
3
4  # Use the variable ''N'' as the population
5  N = cranes$Count
6
7  # Take the log of the population
8  logN = log(N)
9
10 # Get ''future'' and ''past'' arrays
11 logNtp1 = logN[-1]
12 logNt = logN[-length(logN)]
13
14 # Plot them against each other
15 plot(logNt,logNtp1)
16
17 # Linearly regress them and plot the result
18 regression = lm(logNtp1~logNt)
19 lines(logNt,regression$fitted.values)
20
21 # Print the result of the regression
22 summary(regression)
```

The output from the summary command shows $\log N_{t+1} \approx 0.99783 \log N_t + 0.04810$, which implies

$$N_{t+1} = e^{0.04810} N_t^{0.99783}$$

Using N_{1989} from the data, ($N_{1989} = 146$), we get the approximation $N_{2011} = 325.53$. Using the initial condition $N_{1938} = 19$ and the above recursive rule, we get the approximation $N_{2011} = 318.37$.

(c) The website https://www.learner.org/jnorth/tm/crane/PopulTotals.html gives the wild Whooping Crane Population in the Western Flock in 2011 as 278. It looks like the linear regression from (a) underestimated the most (relative error $= -13.3\%$), while the results from (b) are 325.53 and 318.37 (average of 321.95), which is an overapproximation of the actual data (relative error $= 15.8\%$).

<div style="border:1px solid">

Problem 2

Monad's nightmare: Starting with a single cell of *E. coli*, how long would it take to cover the Earth 1 foot deep with *E. coli*? Assume that *E. coli* are rectangular prisms of volume 1.1250×10^{-18} meters^3 and that they reproduce every 20 minutes.

</div>

The radius of the Earth is approximately 20903520 feet, and thus the volume of the 1 foot film of *E. coli* around the Earth is

$$V_{\text{film}} = \frac{4}{3}\pi 20903521^3 - \frac{4}{3}\pi 2090352103^3 = 5490965741568000 \text{ feet}^3 \approx 155486834545908.5625 \text{ meters}^3$$

This means it would take approximately $\frac{155486834545908.5625}{1.1250 \times 10^{-18}} \approx 1.3821051959636316 \times 10^{32}$ *E. coli* cells to cover the earth with a 1 foot film. If we discretize time into twenty minute intervals and assume that the *E. coli* cells never die, then we just need to solve $1.3821051959636316 \times 10^{32} = 2^t$ for t, which has the solution

$$t = 106.7685664639302$$

Since this is the number of twenty minute intervals, this corresponds to about 35.6 hours.

Problem 4

In this problem, you will play around with the Bellows model using some of his data. To get things started you need to download the data file bellows1d.txt and the bifurcation file bellowsBif.R.

(a) For the data file bellows1d.txt, you are going to estimate the parameters a, b, and c for the survivorship function

$$S(N) = \frac{aN}{1 + (bN)^c}$$

using nonlinear regression. To this end, load the data and rename it as follows

```
>temp=read.table("bellows1d.txt",header=FALSE)
>N=temp[,1]
>S=temp[,2]
```

To fit the model $S(N)$ to the data, use the nonlinear regression command

```
>reg=nls(S~a*N/(1+(b*N)^c),start=c(a=1,b=0.01,c=2))
```

where the second argument provides initial guesses for the parameter values. To see the parameter estimates, type

```
>summary(reg)
```

To see what the fit looks like, type

```
>plot(N,S,col="red")
>lines(N,fitted.values(reg),col="blue")
```

(b) With your fitted parameters in hand, simulate the Bellows model

$$N_{t+1} = FS(N_t)$$

with $N_0 = 1$ for the following fecundity values: $F = 1, 3, 4, 6, 7, 10$. Discuss what happens for each of these parameter values.

(c) Determine analytically the minimum value F for which the population persists.

(d) To see how the behaviors change as you vary F, open the R-file bellowsBif.R and change the values of a, b, and c in this file to what you hound in (a) and for the range of F values used in (b). Run the file. R will produce a *bifurcation diagram* that for each parameter value plots the abundances observed in the long-term in the vertical direction e.g. if there is only one point above a given F value then the population tends toward that equilibrium value; if there are two values plotted either the population oscillates between the two values or there are two stable equilibria. Using this plot, do the following:

  (i) Estimate for what range of fecundities the population goes to a positive equilibrium.

  (ii) Estimate for what range of fecundities the population goes to a period two orbit.

  (iii) Estimate for what range of fecundities the population goes to a period four orbit.

  (iv) Discuss whether fecundity is always stabilizing or destabilizing.
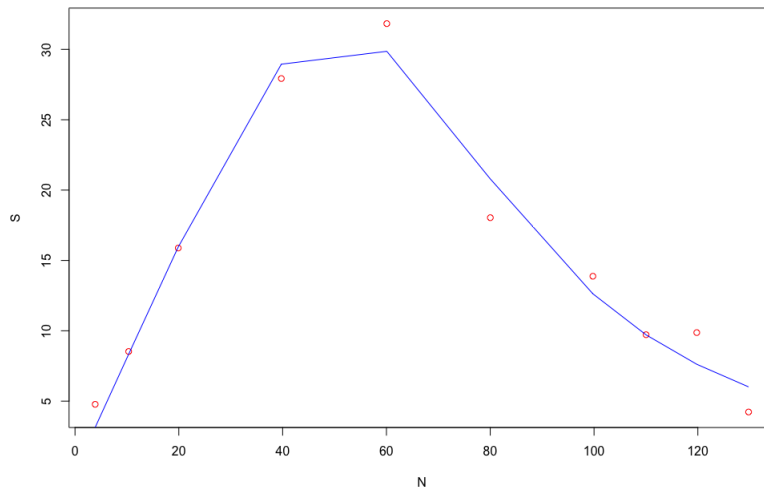
(a) Running the given code gives the following parameter values:

$$a \approx 0.8068723 \qquad b \approx 0.0148920 \qquad c \approx 4.2470522$$
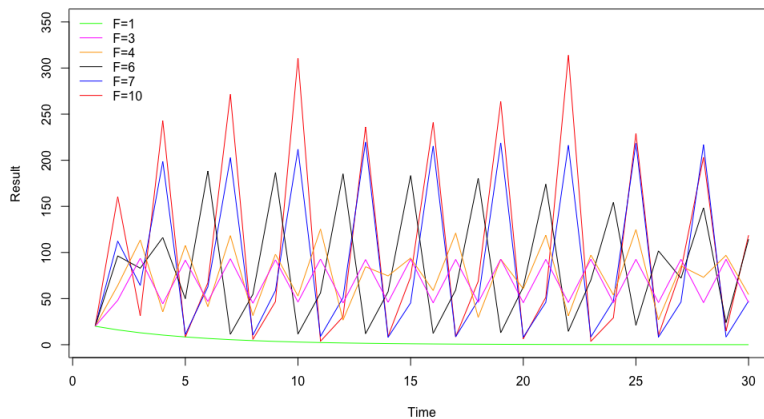
so the model is

$$S(N) \approx \frac{0.8068723N}{1 + (0.0148920N)^{4.2470522}}$$

and here is the plot with the data points in red dots and the model approximation as a blue line:

(b) Here is the result of the model for the six different values of F:

We see that different fecundity values produce chaotic, cyclic, and equilibrium behavior, and in particular, low fecundity results in extinction.

(c) For the model

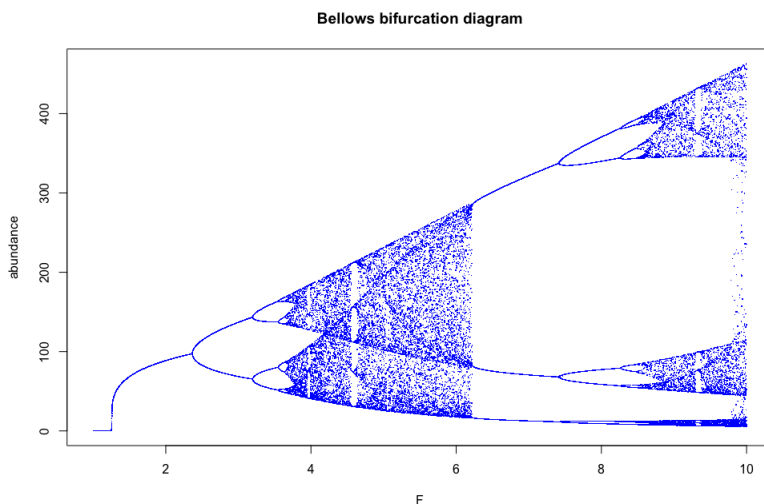$$N_{t+1} = R(N_t) \qquad \text{where} \qquad R(N) = F\frac{aN}{1 + (bN)^c},$$

we know 0 is an equiilbrium since $R(0) = 0$. To find the stability of this equilibrium, notice

$$R'(0) = Fa$$

and thus 0 is stable if $|Fa| < 1$, i.e. $F < \frac{1}{a}$ and unstable if $|Fa| > 1$, i.e. $F > \frac{1}{a}$. Since $a \approx 0.8068723$, this critical value of F is

$$F^* = \frac{1}{a} \approx 1.23935349.$$

(d) Here is the resulting plot from `BellowsBif.R`:

Bellows bifurcation diagram

(i) We see the population reaches equilibrium for $F \lesssim 2.35$, and specifically a positive equilibrium for $F \gtrsim 1.25$.

(ii) We see period two cycles for $2.35 \lesssim F \lesssim 3.19$.

(iii) We see period four cycles for $3.19 \lesssim F \lesssim 3.55$.

(iv) While there is a general trend of de-stabilization as fecundity increases, the bifurcation diagram clearly shows a bifurcation at $F \approx 6.22$ where the dynamics move from chaos to period 3 cycles. As F increases, those cycles then bifurcate to period 6, then period 12, and so on, back to chaos, but there are certain intervals of F which produce finite cycles bounded by regions which produce chaos. There are many of these regions, as we can see from the bifurcation plot for $1 \leq F \leq 30$:

Bellows bifurcation diagram
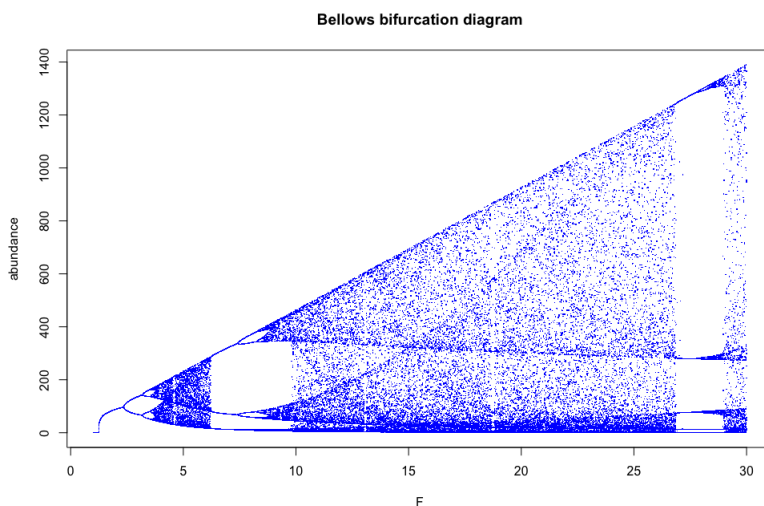
Problem 5

You can find five more population data sets in the resources folder at Smartsite in the files data1.txt, data2.txt, data3.txt, data4.txt, data5.txt. For each of these data sets use Bayesian Information Criterion (BIC) to select one of the following three models where the "process error" is assumed to be normally distributed:

Gompertz model: is of the form

$$\log \frac{N(t+1)}{N(t)} = a + b \log N(t) + \text{"process error"}$$

Ricker model: is of the form

$$\log \frac{N(t+1)}{N(t)} = a + bN(t) + \text{"process error"}$$

Random walk model: is of the form

$$\log \frac{N(t+1)}{N(t)} = a + \text{"process error"}$$

Here is the code used to generate the parameter values for each model, as well as the Log-Likelihood, AIC, and BIC values:

```
1   # Define the Ricker, Gompertz, and Random Walk models
2   RICKER = function(x, a, b) x + a + b*exp(x)
3   GOMPERTZ = function(x, a, b) a + (b+1)*x
4   RANDOM_WALK = function(x, a) x + a
5
6   # Define a list of files to load
7   data_list <- c("data1.txt", "data2.txt", "data3.txt", "data4.txt", "data5.txt")
8
9   for (FILE in data_list) {
10    # Load the file
11    data = read.csv(FILE)
12    N = data$N
13
14    # Take the logarithm of the data
15    y = log(N)
16    k = length(y)
17    y.future = y[-1]
18    y.past = y[-k]
19
20    # Perform the regressions
21    MODEL_GOMPERTZ = nls(y.future~GOMPERTZ(y.past, a, b), start=list(a=0, b=0))
22    MODEL_RICKER = nls(y.future~RICKER(y.past, a, b), start=list(a=0, b=0))
23    MODEL_RANDOM_WALK = nls(y.future~RANDOM_WALK(y.past, a), start=list(a=0))
24
25    # Print the name of the file
26    print(FILE)
27    # Print the Coefficients
28    print("COEFFICIENTS")
29    print(coef(MODEL_GOMPERTZ))
30    print(coef(MODEL_RICKER))
31    print(coef(MODEL_RANDOM_WALK))
32    # Print the Log-Likelihood values
33    print("LOGLIK")
34    print(logLik(MODEL_GOMPERTZ))
35    print(logLik(MODEL_RICKER))
```

```
36    print(logLik(MODEL_RANDOM_WALK))
37    # Print the AIC and BIC values
38    print("AIC")
39    print(AIC(MODEL_GOMPERTZ))
40    print(AIC(MODEL_RICKER))
41    print(AIC(MODEL_RANDOM_WALK))
42    print("BIC")
43    print(BIC(MODEL_GOMPERTZ))
44    print(BIC(MODEL_RICKER))
45    print(BIC(MODEL_RANDOM_WALK))
46  }
```

Here are the results from the various regressions:

data1.txt	a	b	Log-Likelihood	AIC	BIC
Gompertz	2.549111	−0.538343	−68.97711	143.9542	149.4401
Ricker	0.794930970	−0.003908944	−63.17832	132.3566	137.8426
Random Walk	−0.02756187	N/A	−76.10542	156.2108	159.8681

data2.txt	a	b	Log-Likelihood	AIC	BIC
Gompertz	2.2840263	−0.6554932	−11.72863	29.45726	34.74086
Ricker	0.57313072	−0.01663231	−12.86428	31.72857	37.01217
Random Walk	−0.0035849	N/A	−20.00474	44.00948	47.53188

data3.txt	a	b	Log-Likelihood	AIC	BIC
Gompertz	1.2411181	−0.2311494	−2.360294	10.72059	14.12707
Ricker	0.736252393	−0.003664602	−3.534877	13.06975	16.47624
Random Walk	0.1956439	N/A	−12.05806	28.11611	30.3871

data4.txt	a	b	Log-Likelihood	AIC	BIC
Gompertz	0.6757822	−0.1522380	24.10618	−42.21236	−38.6782
Ricker	0.162914901	−0.001894426	24.00982	−42.01963	−38.48547
Random Walk	0.008361279	N/A	22.87781	−41.75561	−39.39951

data5.txt	a	b	Log-Likelihood	AIC	BIC
Gompertz	1.1102957	−0.3326013	−113.7893	233.5786	241.078
Ricker	0.332103758	−0.007314065	−118.2062	242.4124	249.9119
Random Walk	−0.002411484	N/A	−121.9505	247.901	252.9006

We choose the Ricker Model for the first data set and the Gompertz Model for the second, third, fourth, and fifth data sets since these models produce the lowest BIC scores for the data.

Problem 6

Consider a population who in the absence of predation exhibits exponential growth i.e. $G(N) = rN$ without predation. If this population is subject to a predator with a saturating function response, the predation would be of the form

$$\frac{aNP}{H + aN}$$

where P is the predator abundance (assumed to be constant), a is the searching efficiency of the predator and H is the predator's "half-saturation constant". In the presence of predation, the population growth rate is given by

$$G(N) = rN - \frac{aN}{H + aN}.$$

Let $R(N) = G(N)/N$ be the per-capita growth rate term.

(a) Find the maximum per-capita growth rate of the population. Discuss under what conditions the population is doomed to extincton for all initial conditions.

(b) Find the minimum per-capita growth rate of the population. Discuss under what conditions the population persists for all positive initial conditions. Is the population regulated?

(c) Assume that the maximum per-capita growth rate of the population is positive and the minimum per-capita growth rate is negative. Find the positive equilibrium for the model. What happens for populations whose initial abundance lies below this equilibrium? What happens if the initial abundance lies above this equilibrium? Discuss how the resilience of this population can be improved (by decreasing the threshold value).

(a) We see $R(N)$ is a strictly increasing function of N, and thus the maximum per-capita growth rate occurs as $N \to \infty$, i.e. $\lim_{N \to \infty} R(N) = r$. If $r < 0$, the population is doomed to extinction for all initial conditions.

(b) Since R is an increasing function of N, R attains its minimum on the boundary $N = 0$. This minimum is $R(0) = r - \frac{a}{H}$. If $r > \frac{a}{H}$ then the per-capita growth rate, at its worst, is positive, and thus the population persists for all persitive initial conditions. The per-capita growth rate at large densities is positive and bounded from zero, so no, the population is not regulated.

(c) Assuming $0 < r < \frac{a}{H}$, then $R(0) < 0$ and $R(\infty) > \infty$. The positive equilibrium is th value of N for which

$$r = \frac{a}{H + aN^*} \qquad \text{which simplifies to} \qquad N^* = \frac{a - Hr}{ar}.$$

Since $R(N)$ is an increasing function, $R(N) < 0$ for $N < N^*$ and $R(N) > 0$ for $N > N^*$. This means N^* is a threshold value below which the population must not fall if it intends on surviving. If a population density is initially below this threshold, it will approach extinction asymptotically. If a population density is initially above this threshold, it will grow without bound. Note that N^* is a decreasing function of H, and N^* it is not monotonic in a or r. This means we can surely decrease the threshold value by increasing H. Decreasing the threshold value allows the population to undergo larger perturbations without increasing the risk of extinction. This means the population is more resilient to these large perturbations.