Lab 5

Population Growth Models

Purpose

To test the logistic model of population growth, using the United States population census figures, and to learn how to *differentiate* experimental data.

5.1 Malthusian Law of Population Growth

You have learned that population growth can be modelled by a differential equation. However, the population at any given time is an integer. So, how can one talk about taking derivatives, which necessitates an interval of real numbers? For very large populations, any change of the population by a single person is a very small change compared to the total population. Therefore, we can assume that the population changes continuously and can differentiated.

Let $P(t)$ denote the population at any given time. The population will increase due to births and decrease as a result of deaths. The rate of change of the population per person is equal to the difference in the birth and death rates.

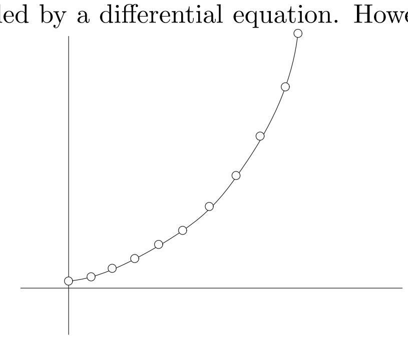

Figure 5.1: Exponential growth.

The simplest model is obtained by assuming that this difference is a constant $k = b - d$. This makes sense since for larger populations, we would expect more births and deaths. If the initial population is given as P_0 at $t = 0$, then we have an *initial value problem* of the form

$$\frac{dP}{dt} = kP, \quad P(0) = P_0. \tag{5.1}$$

We can verify that

$$P(t) = P_0 e^{kt} \tag{5.2}$$

is a solution to equation (5.1).

5.2 The Logistic Law of Population Growth

In the Malthusian model of population growth it is assumed that populations obey an exponential law of the form

$$P(t) = P_0 e^{kt} \tag{5.3}$$

Let's apply this to the global population of the earth. During the period 1960-1970, the world population was increasing at a rate of 2% per year. On January 1, 1965 the earth's population was estimated to be 3.34 billion people. Therefore, by taking $k = .02$ and $P_0 = 3.34 \times 10^9$, the population at any time is predicted by

$$P(t) = (3.34 \times 10^9)e^{.02t}, \tag{5.4}$$

where t is the number of years since 1965. In particular, this model predicts that the global population will be 1,800,000 billion in the year 2625. However, the total surface area of the Earth is about 1,860,000 billion square feet. So, this data implies that there will only be 1 square foot of land per person by 2625. It is evident that the Malthusian model will have to be modified.

The exponential model is satisfactory as long as the population does not get too large. In fact for the US population data given below it is easy to see that the Malthusian model fits well only for the first 6 or 7 data points. When the population gets large, there is competition for the food supply and the space needed to accommodate the people. This competition can be modelled by adding a term, which is proportional to P^2 :

$$\frac{dP}{dt} = kP - bP^2. \tag{5.5}$$

This equation is known as the *logistic equation* and was first introduced in 1837 by the Dutch biologist Verhulst. The constant b is generally very small compared to k. If P is small, then bP^2 will be negligible compared to kP. However, as the population gets larger, bP^2 will begin to have the effect of slowing down the population increase, which is the effect that we need.

The logistic equation can be solved by the method of separation of variables. Separate the P and t dependent parts of equation (5.5) and integrate both sides:

$$\int dt = \int \frac{dP}{kP - bP^2}. \tag{5.6}$$

In order to integrate the righthand side of this equation, we use the method of partial fractions to rewrite the integrand as

$$\frac{1}{kP - bP^2} = \frac{1}{k}(\frac{1}{P} + \frac{b}{k - bP}). \tag{5.7}$$

Equation (5.6) can then be written as

$$k\int dt = \int [\frac{1}{P} + \frac{b}{k - bP}]\, dP. \tag{5.8}$$

Integrating and solving for P, we obtain

$$P(t) = \frac{kAe^{kt}}{1 + bAe^{kt}}. \tag{5.9}$$

The integration constant A can be determined using the initial value of the population, $P(0) = P_0$. Inserting this value into (5.9), gives the solution to the logistic equation:

$$P(t) = \frac{kP_0}{(k - bP_0)e^{-kt} + bP_0}. \tag{5.10}$$

A typical solution looks like an S-shaped curve (sigmoid) as shown to the right. Such curves are found when the initial population lies between the two equilibrium solutions, $P = 0$ and $P = k/b$. For these values of population, one has $\frac{dP}{dt} > 0$. So, the curve has positive slope throughout the region. Since solution curves do not cross, we see that the population is limited to stay below $P = k/b$. In fact, by computing the limit as t gets large, we find that the population predicted by this model is

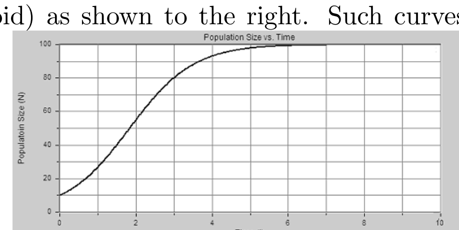

Figure 5.2: A solution of the logistic equation.

$$\lim_{t\to\infty} P(t) = \frac{k}{b}. \tag{5.11}$$

This quantity is often referred to as the *carrying capacity* for the model.

Defining $K = \frac{k}{b}$, the problem can be rewritten as

$$\frac{dP}{dt} = kP(1 - \frac{P}{K}), \tag{5.12}$$

with solution

$$P(t) = \frac{KP_0}{(K - P_0)e^{-kt} + P_0}. \tag{5.13}$$

5.3 Differentiating Data

The population of the United States is given in the following table for ten year intervals from 1790 to 1950. You will use this data to study the validity of the logistic law of population growth [1].

Year	Population	Year	Population
1790	3,929,214	1880	50,189,209
1800	5,308,483	1890	62,979,766
1810	7,239,881	1900	76,212,168
1820	9,638,453	1910	92,228,496
1830	12,866,020	1920	106,021,537
1840	17,069,453	1930	123,202,624
1850	23,191,876	1940	132,164,569
1860	31,443,321	1950	151,325,798
1870	38,558,371		

Fitting this data to the general solution of the logistic equation (5.10) can be done by applying a simple trick to straighten out the curve. First notice that

$$\frac{1}{P} = \frac{1 + bAe^{kt}}{kAe^{kt}} = \frac{1}{kA}e^{-kt} + \frac{b}{k} \tag{5.14}$$

So, by defining $z = 1/P$, we find that:

$$\begin{aligned} \frac{dz}{dt} &= -k[\frac{1}{kA}e^{-kt}] \\ &= -k(z - \frac{b}{k}) \\ \frac{dz}{dt} &= -kz + b \end{aligned} \tag{5.15}$$

Thus, the plot of z' versus z, should be a line with slope $-k$ and intercept b. Since we only have the data for P, we need to find a way to get the corresponding data for z'. You can first compute $z = 1/P$ for each data point P. How do we get the data for z'? We have to go back to our knowledge of differential calculus and recall how to *differentiate data.*

Assume that we have a table with data for t and z in the form $(t_1, z_1), (t_2, z_2), \ldots (t_i, z_i), \ldots (t_N, z_N)$. We would like to "differentiate this data". This means that we approximate the derivative by using the difference quotient

$$\frac{dz}{dt} \approx \frac{\Delta z_i}{\Delta t_i} = \frac{z_{i+1} - z_i}{t_{i+1} - t_i}. \tag{5.16}$$

[1] This problem was adapted from *Differential Equations and Their Applications,* 3rd Ed, M. Braun, Springer-Verlag, p26-32 (1983).

Since the data is given over intervals of 10 years, we could use $\Delta t_i = 10$. A much better approximation would be to use the average of two consecutive quotients, giving the approximation

$$\frac{dz}{dt} \approx \frac{z_{i+1} - z_{i-1}}{t_{i+1} - t_{i-1}}. \tag{5.17}$$

This will be the approximation that you will use. [This is called the discrete central difference approximation.] Note that in this case $\Delta t_i = 20$.

Instructions

Exponential Model

We begin by first fitting an exponential function to the population data.

- You will be provided with an MS Excel spreadsheet with the population data **P** entered into column **B** using units of a million and keeping only one decimal place. (For example, the first population entry would be 3.9.)

- Create a new column array **t** in column **A**, which will give the time in years after 1790. So, $t = 0$ corresponds to the year 1790.

- Graph the **P** versus **t** using an XY-scatterplot.

- You should see a plot of the population that looks somewhat exponential. An exponential function can be fit to this data. **Right mouse click** on a data point. Select **Add Trendline** and choose an exponential type. Before closing the window, select **Display equation on chart** and **Display R-squared value on chart**. Exiting the dialog box, you will get the equation of the exponential fit.

- Excel displays a quantity called the R^2 value under the equation. The closer this value is to 1 the better the data fits the model. Record this number and the equation for the fit. How well does this model fit?

Logistic Model

Now we will fit the logistic model to our data. We will continue using the spreadsheet from before.

- In column **C** define an array of data for $z = 1/P$. To do this let's say that the first data entry in column **B** is in cell **B3**. Then you put the cursor in cell **C3** and type **=1/B3**. The equal sign means you means that the entry in cell **C3** is computed from cell **B3** by means of a formula. To copy the formula to the rest of the appropriate cells you highlight the entire column of relevant cells starting from **C3** and type **Ctrl-D**. You should now see the column filled with different numbers, the reciprocals of the population values.

- Next, you will compute the derivative of this data, as described above, storing the values in column **D**. The starting formula in cell **D4** for the method of central differences would be:

  =(C5-C3)/(A5-A3).

  Copy this down, **being careful not to include the last entry**. (Why?)

- A plot of column **D** (dz/dt) versus column **C** (z) should result in a straight line. Right-mouse-click on a data point on the graph and choose **Add Trendline**. Make sure **Linear** is chosen this time and click on the **Options** tab. Select **Display equation on chart** and **Display R-squared value on chart**. Exiting the dialog box, you will get the equation of the linear fit. Now you can obtain the growth and competition constants from the slope and intercept of this line.
- Record the R^2 value and the equation.
- In order to see how well the logistic model fits the data, you should plot the solution (5.10) with the data. So, you need one more column, **P(t)**. In cells **F3-F5** enter the expressions **k, b, P0** as labels. In corresponding cells of column **G** type in your values for these parameters. We can use these values to develop the solution formula. In cell **E3** you can now type

  =G$3*G$5/((G$3-G$4*G$5)*EXP(-G$3*A3)+G$4*G$5).

  Copy this formula to the rest of the appropriate cells in column **E**. Now select the three columns **A**, **B**, and **E** while holding down the **Ctrl** key. Plot these columns. You should get two sets of data.

Exercises

▷ **Exercise 5.1** How well does the model fit the data? How does this model compare with the Malthusian model as shown by the instructor?

▷ **Exercise 5.2** What is the limiting value for your logistic model? What does this value mean?

▷ **Exercise 5.3** Some additional data is given below. What populations does your model predict? How well does your law fit to this data?

Year	Population	Year	Population	Year	Population
1960	179,323,175	1970	203,211,926	1980	224,454,272

Further Explorations (optional)

Here you can explore population models for other countries. The population data was taken from from the United Nations site http://esa.un.org/unpp/ with predicted populations up to the year 2050 using available data starting in 1950. Several sets of data from this site are stored at the lab web site. However, if get different data from the UN site, then you will have to edit the population data by deleting the blanks used for the "thousand's place holder.

▷ **Exercise 5.4** Load the file "CountryPopModels.xls." The file contains population data for Mexico

1. Use the sliders to adjust the logistic model to get a best fit to the data.
2. Predict the population in the year 3000 if the trend continues.

▷ **Exercise 5.5** Select a different country from the web site. Copy and paste the data for exactly the SAME time period into the excel spreadsheet.

1. Type the name of the country in the text field.
2. Readjust the parameters to get a best fit and predict the population of that country for the year 3000.

Note: For convenience, we have included the data for the following countries: Uganda (High, Medium and low Variance), Liberia, Nigeria, Rwanda, Sierra Leona, Somalia, and Sudan.

Other Links of interest can be found at

- MathDL Module http://mathdl.maa.org/mathDL/4/?pa=content& sa=viewDocument& nodeId=484& pf=1 (Remove spaces in this link.)
- Mathworld. http://mathworld.wolfram.com/LogisticEquation.html