

BMA 771 Introduction: A Simple Population Model

Alun L. Lloyd

August 16, 2017

We are interested in understanding how the number of individuals of some (imaginary) animal species changes over time. For our species, an individual can either be a juvenile or an adult: individuals are born as juveniles and can mature into adults. Only adults can give birth to new juveniles. We make no distinction between males and females (perhaps our species is asexual).

We make observations on the population at fixed time intervals, perhaps once a day or once a week. We write the number of juveniles at time t as x_t , the number of adults at that time as y_t . The initial numbers of juveniles and adults are written as x_0 and y_0 .

Our population model will relate the values of x and y at successive times t and $t+1$. This general type of model is known as a discrete time model, and the particular type of model that we will create is often called a **Leslie matrix model**.

Over the time interval we imagine that one of three things can happen to a juvenile. They can mature into an adult, they could die, or nothing could happen to them. For adults, we imagine that they can either survive the time interval or die during it. We complete the description of the lifecycle by assuming that the number of births (=new juveniles) that occur is proportional to the number of adults alive at the start of the time interval.

In our simplest model, we consider the **average numbers** of juveniles and adults. By this we mean that, if the probability of a juvenile maturing into an adult over the time interval is equal to 0.5 and we have x_t juveniles, then they will give rise to an average of $0.5x_t$ additional adults in the next generation.

To complete our discussion of the model, let's make some assumptions about the probabilities of the other changes that can occur over the time interval. We assume that the probability that a juvenile will die is 0.1. (The probability that a juvenile remains a juvenile is, therefore, equal to $1-0.5-0.1=0.4$.) For the adults, we assume that the probability of death is 0.3. (The probability of an adult surviving is, therefore, equal to 0.7.) Completing the lifecycle, we imagine that adults give

rise to an average of 2 juvenile offspring over the time interval.

If we start off with 100 juveniles and 5 adults we can now calculate how many juveniles and adults we expect to be present at the start of the next time period.

For the number of adults, we know that each of the 5 adults has a 0.7 probability of surviving, giving an average of 3.5 adults. Each of the 100 juveniles has a 0.5 chance of maturing, giving an average of 50 adults. In total, we expect to see, on average, 53.5 adults.

For the juveniles, each of the 100 juveniles has a 0.4 chance of surviving and remaining as a juvenile, giving 40 juveniles. Each of the 5 adults gives birth to an average of 2 new juveniles, giving 10 juveniles. So in total, we expect to see 50 juveniles.

We can now repeat these calculations, but now beginning with 50 juveniles and 53.5 adults, to find the number of individuals at the end of the second time interval.

In general, if we have x_t juveniles and y_t adults, we have

$$x_{t+1} = 0.4x_t + 2y_t \tag{1}$$

$$y_{t+1} = 0.5x_t + 0.7y_t. \tag{2}$$

1 Numerical Simulation of the Model Using Excel

We could continue to calculate the numbers in successive generations by hand calculation. This quickly starts to get tedious. It is easy to get a computer to perform the repeated calculations. For this sort of problem, a spreadsheet like Excel is an excellent choice. But feel free to use whatever package you feel most comfortable with. What follows is a detailed explanation of how to simulate and plot the behavior of this model using Excel. (I wrote these instructions based on the PC version of Excel; the keyboard shortcuts mentioned are a little different on the Mac.) If you are comfortable with Excel, you can probably skip most of the rest of this section.

Excel labels the cells in a spreadsheet using letters (column label) and numbers (row labels). So A1 denotes the top left entry of a spreadsheet. It's most convenient to work with columns.

Use the first column to hold the values of t running from 0 to, say, 20. A quick way to do this is to type 0 in the first cell (then press **ENTER**) and 1 in the second (then press **ENTER**). Select these two cells (click on the zero with the left mouse button and drag the pointer down into the second cell while keeping the button held down). Without holding down any mouse buttons, move the pointer towards the bottom right hand corner of cell A3 (which should still be selected). The white cross

pointer changes to a black cross. When this happens, click and hold the left mouse button and drag the selection down as far as the 21st row. Excel will then fill in the remaining cells according to the pattern it has spotted.

We will use the next two columns (B and C) to represent the numbers of juveniles and adults. Enter 100 and 5 as the initial numbers of juveniles and adults: these numbers go into cells B1 and C1. (As always, press ENTER after entering the contents of a cell.) To get the numbers in the new row, we use equations (1) and (2) from above. Click on cell B2 and enter the following

$$=0.4*B1+2*C1$$

Click on cell C2 and enter the following

$$=0.5*B1+0.7*C1$$

Excel should have calculated the numbers that we obtained earlier.

To get the successive numbers is now easy. Select cells B2 and C2 (left-click on B2, hold the button and drag the pointer to C2 before releasing the button). Then copy the cells (either using the copy option on the 'Edit' menu or by pressing the CTRL and C keys simultaneously). Finally, paste into B3 and C3: click on B3 and use the paste option on the 'Edit' menu or press CTRL and V simultaneously.

You can copy and paste repeatedly to fill in the rest of the rows. Alternatively, you can do it all at once by selecting the entire region from B4 to C21 and then paste.

Notice that you don't have to edit the references (e.g. to B1 and C1) each time. This is because the references B1 and C1 are **relative references**. If you copy and paste a cell containing the reference B1 to the cell immediately below, Excel replaces the B1 by B2. If you click on cell B3, you will see that Excel has done this, and has replaced C1 by C2. (If you copy and paste to the cell to the right, Excel would replace B1 by C1 and C1 by D1.) Looking down the B column, clicking on each cell reveals that Excel has repeatedly increased the row labels by 1.

An alternative way of referencing a cell is by an **absolute reference**. If we enter a reference as **\$A\$1**, then Excel will not change the row or column number when we copy and paste the cell elsewhere. So, for instance, if we wanted another column (say column E) to contain some multiple of the entries in the second column, we could enter the multiplier into some cell (for instance D1) and enter **= \$D\$1 * B1** into the first cell of column E. We could then copy and paste this cell into the rest of column E. Going down the column, Excel would alter the reference B1 to point to the appropriate entry of column B, but would not alter the reference to D1.

You can mix absolute and relative references, such as **\$A1** or **A\$1**. Only one of the two labels would then be altered as the cell is copied and pasted elsewhere. You might sometimes want to simply copy a column from one place to another. But if the cells contain expressions that involve relative references there will be a problem. If you simply want to copy the values that are present at that point in time, you can use the ‘paste special’ option from the ‘Edit’ menu, and choose ‘paste values’.

In column D, get Excel to calculate the total number of individuals. In the first entry, D1, enter the following

= B1 + C1

Fill out the rest of the column by copying the contents of D1 into the remaining cells.

One advantage of using Excel is that it is easy to generate graphs. To plot the numbers of juveniles and adults against time, first select the region from A1 to C21. From the ‘Insert’ menu, choose ‘Chart...’, bringing up the Chart Wizard, and then select ‘XY (scatter)’. Choose any appropriate chart type (I prefer the straight lines with symbols) and click on ‘Next’. On the second screen of the Chart Wizard, you can click on the ‘Series’ tab, from where you can label Series 1 as juveniles and Series 2 as adults. Clicking on ‘Next’ brings up the third screen of the Chart Wizard, where you can insert some labeling (e.g. axes and titles) and formatting, if you wish. The final step of the Chart Wizard asks where you want to place the new figure. I usually choose to place it as an object in the current sheet.

If you wanted to plot the total number of individuals against time, you would have to select column A, then add column D to the selection. You can do this by selecting column A, releasing the mouse button, then hold down the **CTRL** key and select column D. (A shortcut to selecting an entire column is to click on the column label. So you could select columns A and D by clicking on the ‘A’ on the spreadsheet and then click on the ‘D’ while holding down **CTRL**.) Once these two columns are selected, plot the graph as above.

2 Model Behavior

Both the number of juveniles and the number of adults appear to grow exponentially over time. Similarly for the total number of individuals. We can confirm this by replotting the total number of individuals on a logarithmic scale. An easy way to do this in Excel is to click on the y-axis of the graph : the ‘Format Axis’ window should appear. Click on the ‘Scale’ tab and then check the ‘Logarithmic scale’ box.

An alternative way of getting a logarithmic plot is to set up a new column containing the logarithm

of the total number of individuals. We will put this into column E. In cell E1 enter `= LN(D1)`. Copy and paste this into the cells E2 to E21. Plot a graph as usual.

The straight line behavior of the logged totals against time shows that the population grows exponentially in the long term. Excel's linear regression functions can give the slope and intercept of the line and give. (You may need to have loaded the Analysis or Statistics addin toolpaks to get these functions to work.) To get the slope, enter

`=SLOPE(E1:E21,A1:A21)`

into some empty cell, and

`=INTERCEPT(E1:E21,A1:A21)`

into another. The arguments of the functions give the ranges of cells that contain the y and x values for the regression line. You should see that the slope is approximately equal to 0.4422, which means that the total population size grows as $e^{0.4422t}$, which can be written as 1.56^t .

It is instructive to look at the ratio of the number of adults to the number of juveniles. Calculate this ratio in column F (in cell F1 enter `= C1/B1`, copy and paste as usual). Plotting this ratio against time shows that the ratio of adults to juveniles approaches a constant over time. The numbers of juveniles and adults approach what is known as a **stable age distribution**: this happens for many models.

It is also interesting to look at the ratios of the numbers of juveniles (and adults) in successive time periods. Calculate these ratios in columns G and H (in cell G2 enter `= B2/B1` and in cell H2 enter `= C2/C1`, copy and paste as usual). We see that these ratios settle down to the same growth factor that we saw for the total number of individuals, namely 1.56. We call this number λ , and we will come back to look more closely at its value later.

Putting the results of the last three paragraphs together, we see that the numbers of both juveniles and adults increase exponentially, both with the same rate constant.

3 Mathematical Analysis of the Model

Remember that we have the following equations for the model

$$x_{t+1} = 0.4x_t + 2y_t \quad (3)$$

$$y_{t+1} = 0.5x_t + 0.7y_t. \quad (4)$$

These equations can be recast in the following **matrix form**

$$\begin{pmatrix} x_{t+1} \\ y_{t+1} \end{pmatrix} = \begin{pmatrix} 0.4 & 2 \\ 0.5 & 0.7 \end{pmatrix} \begin{pmatrix} x_t \\ y_t \end{pmatrix}, \quad (5)$$

or in the even more compact notation

$$\mathbf{v}_{t+1} = M\mathbf{v}_t. \quad (6)$$

Here $\mathbf{v}_t$ is a vector whose entries give the numbers of juveniles and adults at time t . M is a 2×2 matrix: it has two rows and two columns.

4 Long-Term Model Behavior: Eigenvalues and Eigenvectors

We saw that in the long term, the numbers of juveniles and adults both grow exponentially. Going from one time period to the next, the numbers of juveniles and adults are both multiplied by the same factor, which we called λ . In terms of the vector notation, if the number of juveniles at some time equals a and the number of adults equals b , this means that

$$\begin{aligned} \mathbf{v}_{t+1} &= \begin{pmatrix} \lambda a \\ \lambda b \end{pmatrix} \\ &= \lambda \begin{pmatrix} a \\ b \end{pmatrix}. \end{aligned} \quad (7)$$

Comparing equations (6) and (7) we see that the numerically observed long-term behavior of the model implies that

$$\lambda \begin{pmatrix} a \\ b \end{pmatrix} = M \begin{pmatrix} a \\ b \end{pmatrix}. \quad (8)$$

Writing this out in full, we see that we have a pair of simultaneous linear equations. These can be solved using techniques learned in high-school algebra. We would then obtain the values of λ , a and b (or at least the ratio b/a).

We can write equation (8) as

$$\lambda \mathbf{v} = M\mathbf{v}, \quad (9)$$

for some non-zero vector $\mathbf{v}$. (If both entries of $\mathbf{v}$ were zero then there would be no individuals in the population!) This means that when we multiply the vector $\mathbf{v}$ by the matrix M , we get back out a vector that is just some multiple of the original vector. We say that $\mathbf{v}$ is an **eigenvector** of the matrix M with eigenvalue λ .

Notice what this says in biological terms: the eigenvector gives the stable age distribution and the eigenvalue gives the growth factor that is observed for a population that is in the stable age distribution.

4.1 How Do We Find the Eigenvalue and Eigenvector?

We just saw that the eigenvalue and eigenvector can be found by writing equation (8) out in full and solving the resulting simultaneous equations. We shall now see how techniques from linear algebra can be used to solve the equations. (These techniques will be useful when it comes to looking at more complex problems.)

Rearranging equation (9) gives

$$\begin{aligned}\mathbf{0} &= M\mathbf{v} - \lambda\mathbf{v} \\ &= (M - \lambda I)\mathbf{v}.\end{aligned}\tag{10}$$

Here, I is the identity matrix: $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, and $\mathbf{0}$ is the zero vector: $\mathbf{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. For our problem, the matrix $M - \lambda I$ is given by $\begin{pmatrix} 0.4 - \lambda & 2 \\ 0.5 & 0.7 - \lambda \end{pmatrix}$.

Equation (10) might look odd. If this were ordinary multiplication of numbers, then this equation could only hold if one of the terms on the right hand side equalled zero. However, this is matrix multiplication and this is one of the situations in which the rules of matrix multiplication differ from those of ordinary multiplication of numbers. In matrix multiplication, you can have the product $M\mathbf{v}$ equalling zero without either term being zero.

Standard theory from linear algebra tells us that equation (10) can have a non-zero solution if something called the **determinant** of the matrix $M - \lambda I$ equals zero. (This is pretty much the only result in the whole of these notes that you need to take on trust if you haven't come across it before.) There is a standard formula for the determinant of a matrix: the determinant of the general 2×2 matrix $\begin{pmatrix} x & y \\ w & z \end{pmatrix}$ is given by $xz - wy$. For our problem, we have

$$(0.4 - \lambda)(0.7 - \lambda) - 0.5 \times 2 = 0.\tag{11}$$

If we multiply this out, we get the quadratic

$$\lambda^2 - 1.1\lambda - 0.72 = 0,\tag{12}$$

which we can solve using the formula for the general solution of the quadratic equation. This gives us two answers: 1.5611..., which is the growth factor we saw in the numerical simulations, but also -0.4611.... We will come back to this second solution later.

Turning to the eigenvector that corresponds to $\lambda = 1.5611\dots$, equation (10) tells us that $\mathbf{v}$ satisfies

$$\begin{aligned}\mathbf{0} &= (M - \lambda I) \mathbf{v} \\ &= \begin{pmatrix} 0.4 - \lambda & 2 \\ 0.5 & 0.7 - \lambda \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} \\ &= \begin{pmatrix} -1.1611 & 2 \\ 0.5 & -0.8611 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}.\end{aligned}\tag{13}$$

Multiplying out, this gives us two simultaneous linear equations:

$$-1.1611a + 2b = 0\tag{14}$$

$$0.5a - 0.8611b = 0.\tag{15}$$

Actually, these two equations are essentially the same. If we multiply both sides of the first equation by $(-0.5/1.1611) = -0.3103$, we obtain the second equation. Rearranging either equation shows that $b/a \approx 0.58$, as we saw in the numerical simulations. These equations do not fix the overall sizes of a and b . This is a general result: eigenvectors are only determined up to an arbitrary multiplier. If $\mathbf{v}$ is an eigenvector with eigenvalue λ , then so is $\alpha\mathbf{v}$, for any multiple α .

What about the other eigenvalue that we found, $\lambda = -0.4611\dots$? For this eigenvalue, we can show that the corresponding eigenvector has $b/a \approx -0.431$. Later, we will see the significance of this eigenvalue and eigenvector.

To summarize: the eigenvector and eigenvalue can tell us about the stable age distribution of the population and how the population grows when it has reached such a state. The eigenvector and eigenvalue can be found using standard theory from linear algebra.

5 General Solution of the Model

Imagine that the initial numbers of juveniles and adults are equal to $\mathbf{v}$, where $\mathbf{v}$ is an eigenvector of M . We could then find the numbers after the first time period by the product $M\mathbf{v}$, which we know equals $\lambda\mathbf{v}$. We could then find the numbers after the second time period: $M(\lambda\mathbf{v}) = \lambda M\mathbf{v} = \lambda^2\mathbf{v}$. Repeating, it is clear that $\mathbf{v}_t = \lambda^t\mathbf{v}$.

What we have found using the eigenvalue and eigenvector approach, therefore, are two independent solutions of our model:

$$\mathbf{v}_t = \lambda_1^t \mathbf{v}^{(1)}\tag{16}$$

and

$$\mathbf{v}_t = \lambda_2^t \mathbf{v}^{(2)},\tag{17}$$

where $\lambda_1 \approx 1.56$ and $\lambda_2 \approx -0.46$. $\mathbf{v}^{(1)}$ and $\mathbf{v}^{(2)}$ are the corresponding eigenvectors.

It turns out that the **general** solution of our model is given by adding together multiples of these two solutions:

$$\mathbf{v}_t = c_1 \lambda_1^t \mathbf{v}^{(1)} + c_2 \lambda_2^t \mathbf{v}^{(2)}. \quad (18)$$

The constants c_1 and c_2 are determined by the **initial conditions** of the problem, i.e. the numbers of juveniles and adults at the start. If we set $t = 0$, we get a pair of simultaneous equations that relate c_1 and c_2 to the initial numbers of juveniles and adults. We could, in theory, solve these to find c_1 and c_2 .

The ability to add solutions together to give another solution is a characteristic property of **linear systems**. This will be an area that we will study in more detail in BMA 771.

Notice that the absolute value of the second eigenvalue is smaller than one. As t increases, λ_2^t will tend to zero. The long term behavior is exponential growth, as described by the first eigenvalue and eigenvector. The second term is only important in the short-term, and describes the fluctuations seen about the exponentially growing solution. For this reason, we sometimes call λ_1 the **dominant eigenvalue** of M .

For the parameter values chosen in this problem, the absolute value of the second eigenvalue was less than one. This need not always be the case, and so it could be that both terms in (18) are growing exponentially. Typically, one of the two terms will grow faster than the other; its behavior will dominate after a short time.

6 Concluding Comments

In these notes we have introduced a simple population model, seen one way of numerically simulating its behavior and looked at some mathematical analyses of the behavior. We have seen the importance of eigenvalues and eigenvectors and have touched upon how we find their values. Many of the points that came up will be taken up and looked at in more detail during BMA 771.

Further reading: Chapter 9 of Neuhauser, *Calculus for Biology and Medicine, 2nd edition*, Prentice-Hall, in particular: sections 9.2.5 and 9.3. There is a copy of this in the Hill library.