Further Mathematical Methods (Linear Algebra) 2002

Solutions For Problem Sheet 4

In this Problem Sheet, we revised how to find the eigenvalues and eigenvectors of a matrix and the circumstances under which matrices are diagonalisable. We also used these skills in our study of age-specific population growth.

1. To find the characteristic polynomial, $p(\lambda)$, of the matrix A, we use

$$p(\lambda) = \det(\mathsf{A} - \lambda\mathsf{I}) = \begin{vmatrix} 1-\lambda & 0 & a \\ 0 & 1-\lambda & 0 \\ 1 & 0 & 2-\lambda \end{vmatrix}$$

This determinant can be evaluated by doing a *co-factor expansion* along the top row, and this yields

$$p(\lambda) = (1-\lambda)^2(2-\lambda) - a(1-\lambda) = (1-\lambda)\{(1-\lambda)(2-\lambda) - a\}$$

which will be useful in a moment. But, multiplying out the brackets instead of factorising, we find that the characteristic polynomial is

$$p(\lambda) = -\lambda^3 + 4\lambda^2 + (a-5)\lambda + 2 - a.$$

Notice that, as expected from the lectures, the characteristic polynomial of a 3×3 matrix is a cubic where the λ^3 term has a coefficient of $(-1)^3 = -1$. (Further, if you look at Question 11, we should expect that the constant term, i.e. $2 - a$, gives the value of $\det(\mathsf{A})$ — and you can verify that it does!)

We are now asked to show that the matrix A is diagonalisable when $a = 0$, but not when $a = -\frac{1}{4}$. To do this, recall that an $n \times n$ matrix A is diagonalisable iff it has n linearly independent eigenvectors. So, we have to establish that if $a = 0$ the matrix A has three linearly independent eigenvectors, but if $a = -\frac{1}{4}$ it doesn't.

<u>When $a = 0$:</u> From our earlier analysis of the characteristic polynomial it should be clear that in this case the matrix A has eigenvalues given by $\lambda = 1, 1, 2$. (That is the eigenvalue $\lambda = 1$ is of *multiplicity* two.) Now, if a matrix has distinct eigenvalues, then the eigenvectors corresponding to these distinct eigenvalues are linearly independent[1] and so as there must be at least one eigenvector corresponding to each eigenvalue[2] and so, we have at least two linearly independent eigenvectors already. But, to diagonalise A we require three and so, the question is actually asking whether this matrix has enough linearly independent eigenvectors corresponding to $\lambda = 1$, i.e. two.[3] To show that this is indeed the case when $a = 0$, we can find out precisely what the eigenvectors corresponding to the eigenvalue $\lambda = 1$ are.

So, to find the eigenvectors $\mathbf{x}$ corresponding to this eigenvalue, we have to solve the simultaneous equations which, in matrix form, are given by the expression

$$(\mathsf{A} - \mathsf{I})\mathbf{x} = \mathbf{0}$$

(Note, *by definition*, the trivial solution to this system of simultaneous equations, i.e. $\mathbf{x} = \mathbf{0}$, is *not* an eigenvector corresponding to $\lambda = 1$.) That is, we have

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathbf{0}$$

[1] See, for example, Question 2, or its generalisation in Question 8.

[2] That is, for each eigenvalue, we can always find an eigenspace with a dimension of at least one with this eigenvector serving as a basis for this space. Obviously, to get an eigenspace whose dimension is greater than one, we require that the corresponding eigenvalue has multiplicity (i.e. it is a *repeated* root of the characteristic polynomial).

[3] Indeed, in terms of the previous footnote, this is just asking whether the eigenspace corresponding to the multiplicitous eigenvalue has a sufficiently large dimension — in this case two — to yield an eigenbasis containing two linearly independent eigenvectors.

which gives us one equation, namely $x + z = 0$, in three variables. So, taking y and z to be the two free parameters we find that eigenvectors corresponding to $\lambda = 0$ will have the form

$$\mathbf{x} = \begin{bmatrix} -z \\ y \\ z \end{bmatrix} = z \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + y \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Consequently, two linearly independent eigenvectors corresponding to $\lambda = 1$ would be $[-1, 0, 1]^t$ and $[0, 1, 0]^t$. Thus, we can find three linearly independent eigenvectors which can be used as the column vectors of an invertible matrix P, and so the matrix A is diagonalisable if $a = 0$ (as required).

When $a = -\frac{1}{4}$: From our earlier analysis of the characteristic polynomial it should be clear that in this case, the matrix A has eigenvalues given by

$$(1 - \lambda)\{(1 - \lambda)(2 - \lambda) + \tfrac{1}{4}\} = 0$$

and solving this we find that $\lambda = 1, \frac{3}{2}, \frac{3}{2}$. (That is, the eigenvalue $\lambda = \frac{3}{2}$ is of *multiplicity* two.) As in the previous case, we know that we are going to get at least two linearly independent eigenvectors — one corresponding to $\lambda = 1$ and at least one corresponding to $\lambda = \frac{3}{2}$. So, to show that A is not diagonalisable, it is sufficient to show that, in fact, there is only one linearly independent eigenvector corresponding to $\lambda = \frac{3}{2}$.

To do this, we note that the eigenvectors $\mathbf{x}$ corresponding to this eigenvalue, can be found by solving the simultaneous equations which, in matrix form, are given by the expression

$$(\mathsf{A} - \tfrac{3}{2}\mathsf{I})\mathbf{x} = \mathbf{0}$$

(Again, note that *by definition*, the trivial solution to this system of simultaneous equations, i.e. $\mathbf{x} = \mathbf{0}$, is *not* an eigenvector corresponding to $\lambda = \frac{3}{2}$.) That is, we have

$$\begin{bmatrix} -\frac{1}{2} & 0 & -\frac{1}{4} \\ 0 & -\frac{1}{2} & 0 \\ 1 & 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathbf{0}$$

which gives us three equations, namely

$$\begin{aligned} -\tfrac{1}{2}x - \tfrac{1}{4}z &= 0 \\ -\tfrac{1}{2}y &= 0 \\ x + \tfrac{1}{2}z &= 0 \end{aligned}$$

But, the top equation is only the bottom equation multiplied by $-\frac{1}{2}$ and so, really, we just have two equations in three variables. So, taking z to be the free parameter we find that eigenvectors corresponding to $\lambda = \frac{3}{2}$ have the form

$$\mathbf{x} = \begin{bmatrix} -\frac{1}{2}z \\ 0 \\ z \end{bmatrix} = z \begin{bmatrix} -\frac{1}{2} \\ 0 \\ 1 \end{bmatrix}$$

Consequently, we can only find one linearly independent eigenvector corresponding to $\lambda = \frac{3}{2}$, say $[-\frac{1}{2}, 0, 1]^t$. Thus, as we can *not* find three linearly independent eigenvectors which could be used as the column vectors of an invertible matrix P, the matrix A is *not* diagonalisable if $a = -\frac{1}{4}$ (as required).

2. Let us consider a general[4] matrix A (which is square and at least 2×2) that has two distinct eigenvalues λ_1 and λ_2.[5] Also, $\mathbf{x}_1$ and $\mathbf{x}_2$ are eigenvectors of A corresponding to the eigenvalues λ_1 and λ_2 respectively, i.e. they must be such that

$$\mathsf{A}\mathbf{x}_1 = \lambda_1\mathbf{x}_1 \quad \text{and} \quad \mathsf{A}\mathbf{x}_2 = \lambda_2\mathbf{x}_2$$

[4]The restriction that the matrix should be 2×2 is irrelevant and possibly distracting.

[5]This is the generalisation mentioned in the question above. If A is an $n \times n$ matrix, then it may have just two distinct (and multiplicitous) eigenvalues or it may have many distinct eigenvalues (some of which may be multiplicitous) of which we are just considering two. The proof that is presented here does not depend on such details.

and we have to show that these vectors are linearly independent. To do this we consider the vector equation

$$\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 = \mathbf{0}$$

where we can establish that the set of vectors $\{\mathbf{x}_1, \mathbf{x}_2\}$ is linearly independent if this equation *only* has the trivial solution $\alpha_1 = \alpha_2 = 0$. So that we can solve this equation for α_1 and α_2 using *all* of the information that we have about $\mathbf{x}_1$ and $\mathbf{x}_2$, we now multiply this equation by the matrix A , i.e.

$$\alpha_1\mathsf{A}\mathbf{x}_1 + \alpha_2\mathsf{A}\mathbf{x}_2 = \mathsf{A}\mathbf{0} \implies \alpha_1\lambda_1\mathbf{x}_1 + \alpha_2\lambda_2\mathbf{x}_2 = \mathbf{0}$$

where we have used the eigenvalue-eigenvector relations above in the second step. Consequently, if we multiply our original vector equation by λ_1 (say, assuming [without loss of generality] that $\lambda_1 \neq 0$) and subtract it from the vector equation above we get

$$\alpha_2(\lambda_2 - \lambda_1)\mathbf{x}_2 = \mathbf{0}$$

where as $\mathbf{x}_2 \neq \mathbf{0}$ (it is an eigenvector) and $\lambda_1 \neq \lambda_2$ it must be the case that $\alpha_2 = 0$. Substituting this into our original vector equation we get

$$\alpha_1\mathbf{x}_1 = \mathbf{0}$$

and as $\mathbf{x}_1 \neq \mathbf{0}$ (it is an eigenvector too) we get $\alpha_1 = 0$. Clearly, the solution that we have found, i.e. $\alpha_1 = \alpha_2 = 0$, is the only solution that this pair of vector equations is going to yield and so the set $\{\mathbf{x}_1, \mathbf{x}_2\}$ is linearly independent (as required).

3. To verify that $\mathbf{v}_1$ is an eigenvector of the Leslie matrix L, you should have shown that $\mathsf{L}\mathbf{v}_1 = \lambda_1\mathbf{v}_1$. This is easily accomplished as

$$\mathsf{L}\mathbf{v}_1 = \begin{bmatrix} a_1 & a_2 & a_3 & \cdots & a_{n-1} & a_n \\ b_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & b_2 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & b_{n-1} & 0 \end{bmatrix} \begin{bmatrix} 1 \\ b_1/\lambda_1 \\ b_1b_2/\lambda_1^2 \\ \vdots \\ b_1b_2\ldots b_{n-1}/\lambda_1^{n-1} \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 \\ b_1 \\ b_1b_2/\lambda_1 \\ \vdots \\ b_1b_2\ldots b_{n-1}/\lambda_1^{n-2} \end{bmatrix} = \lambda_1 \begin{bmatrix} 1 \\ b_1/\lambda_1 \\ b_1b_2/\lambda_1^2 \\ \vdots \\ b_1b_2\ldots b_{n-1}/\lambda_1^{n-1} \end{bmatrix} = \lambda_1\mathbf{v}_1.$$

Notice that we have used two facts from the handout for Lecture 8, namely that

$$a_1 + a_2b_1/\lambda_1 + a_3b_1b_2/\lambda_1^2 + \cdots + a_nb_1b_2\ldots b_{n-1}/\lambda_1^{n-1} = \lambda_1 q(\lambda_1)$$

and $q(\lambda_1) = 1$.

4. Clearly, given the information in the question, the required Leslie matrix is

$$\mathsf{L} = \begin{bmatrix} 1 & 4 \\ 1/2 & 0 \end{bmatrix}$$

and the initial population distribution vector is $\mathbf{x}^{(0)} = [1000, 1000]^t$. To find an *exact* formula for $\mathbf{x}^{(k)}$, the population distribution vector after k time periods (in this case, after $5k$ years), we need to evaluate

$$\mathbf{x}^{(k)} = \mathsf{L}^k\mathbf{x}^{(0)}.$$

Obviously, we do this by diagonalising the matrix L, and so finding the eigenvalues and eigenvectors we obtain

$$\mathsf{P} = \begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix}, \quad \mathsf{P}^{-1} = \frac{1}{6}\begin{bmatrix} 1 & 2 \\ -1 & 4 \end{bmatrix} \quad \text{and} \quad \mathsf{D} = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$$

where $\mathsf{P}^{-1}\mathsf{LP} = \mathsf{D}$. Thus, $\mathsf{L} = \mathsf{PDP}^{-1}$ and so, L^k is given by

$$\mathsf{L}^k = \underbrace{(\mathsf{PDP}^{-1})(\mathsf{PDP}^{-1})\ldots(\mathsf{PDP}^{-1})}_{k \text{ times}} = \mathsf{PD}^k\mathsf{P}^{-1}$$

as $\mathsf{P}^{-1}\mathsf{P} = \mathsf{I}$. This means $\mathbf{x}^{(k)} = \mathsf{PD}^k\mathsf{P}^{-1}\mathbf{x}^{(0)}$, and so

$$\begin{aligned}
\mathbf{x}^{(k)} &= \begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2^k & 0 \\ 0 & (-1)^k \end{bmatrix} \begin{bmatrix} 1/6 & 1/3 \\ -1/6 & 2/3 \end{bmatrix} \begin{bmatrix} 1000 \\ 1000 \end{bmatrix} \\
&= \begin{bmatrix} 4 \cdot 2^k & -2(-1)^k \\ 2^k & (-1)^k \end{bmatrix} \begin{bmatrix} 500 \\ 500 \end{bmatrix} \\
\therefore\ \mathbf{x}^{(k)} &= 500 \begin{bmatrix} 4 \cdot 2^k - 2(-1)^k \\ 2^k + (-1)^k \end{bmatrix}
\end{aligned}$$

is the required expression.[6]

We are now asked to *check* that

$$\lim_{k\to\infty} \frac{\mathbf{x}^{(k)}}{\lambda_1^k} = c\mathbf{v}_1$$

for some constant c, where in this case $\lambda_1 = 2$ is the unique positive real eigenvalue and $\mathbf{v}_1 = [4, 1]^t$ is an eigenvector of the Leslie matrix corresponding to this eigenvalue. Thus, dividing both sides of our exact expression for $\mathbf{x}^{(k)}$ by 2^k we obtain

$$\frac{\mathbf{x}^{(k)}}{2^k} = 500 \begin{bmatrix} 4 - 2(-1/2)^k \\ 1 + (-1/2)^k \end{bmatrix}$$

and so in the limit as $k \to \infty$, $(-1/2)^k \to 0$, which means that

$$\lim_{k\to\infty} \frac{\mathbf{x}^{(k)}}{2^k} = 500 \begin{bmatrix} 4 \\ 1 \end{bmatrix}$$

as required.[7]

5. Clearly, given the information in the question, the required Leslie matrix is

$$\mathsf{L} = \begin{bmatrix} 0 & 1/4 & 1/2 \\ 1/2 & 0 & 0 \\ 0 & 1/4 & 0 \end{bmatrix}.$$

Calculating the characteristic polynomial, you should find that the eigenvalues of L are the roots of the cubic equation:

$$\lambda^3 - \frac{1}{8}\lambda - \frac{1}{16} = 0.$$

To solve this, you have to use 'trial and error', and this should lead you to the conclusion that $\lambda = 1/2$ is a root. Further, as the theory in the handout for Lecture 8 guarantees that the Leslie matrix will have only one positive real eigenvalue, this must be it.[8]

We now notice that the Leslie matrix has two successive fertile classes, and so the theory in the handout for Lecture 8 tells us that the eigenvalue which we have calculated is *dominant*.[9] This means

[6] Incidentally, notice that the second line of this calculation tells us that the first entry of $\mathsf{P}^{-1}\mathbf{x}^{(0)}$ is 500.

[7] With reference to the previous footnote, observe that the constant c is 500. This is what we expect from the theory given in the handout for Lecture 8.

[8] Incidentally, dividing the cubic by the factor $\lambda - 1/2$ gives us a quadratic equation, namely $\lambda^2 + \frac{1}{2}\lambda + \frac{1}{8} = 0$, to solve for the other two eigenvalues. Doing this we find that they are $\lambda = \frac{1}{4}(-1 \pm i)$.

[9] Recall that an eigenvalue, λ_1 of a matrix L is *dominant* if $\lambda_1 > |\lambda_i|$ where the λ_i are the other eigenvalues of L. In this case, the moduli of the other eigenvalues are given by $\frac{1}{4}|-1 \pm i| = \frac{1}{4}\sqrt{1^2 + 1^2} = \frac{\sqrt{2}}{4} = \frac{1}{2\sqrt{2}} < \frac{1}{2}$. Thus the eigenvalue $\lambda = 1/2$ is dominant, as expected.

that we can describe the long-term behaviour of the population distribution of this species using the results in the handout for Lecture 8, i.e. for large k,

$$\mathbf{x}^{(k)} \simeq c\lambda_1^k \mathbf{v}_1 \quad \text{and} \quad \mathbf{x}^{(k)} \simeq \lambda_1 \mathbf{x}^{(k-1)}.$$

But, to do this, we need an eigenvector, $\mathbf{v}_1$, corresponding to our unique positive real eigenvalue, so using the result of Question 3, we get

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ b_1/\lambda_1 \\ b_1 b_2/\lambda_1^2 \end{bmatrix} = \begin{bmatrix} 1 \\ (1/2)/(1/2) \\ (1/2)(1/4)/(1/2)^2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1/2 \end{bmatrix}$$

as here, $\lambda_1 = 1/2$. Thus, the long-term behaviour of the population distribution of this species is given by

$$\mathbf{x}^{(k)} \simeq c(\tfrac{1}{2})^k \begin{bmatrix} 1 \\ 1 \\ 1/2 \end{bmatrix} \quad \text{and} \quad \mathbf{x}^{(k)} \simeq \tfrac{1}{2}\mathbf{x}^{(k-1)}.$$

Consequently, we can see that, in the long run, the proportion of the population in each age class becomes steady in the ratio $1 : 1 : \frac{1}{2}$ and that the population in each age class decreases by 50% every time period (i.e. every ten years). Further, as the population is decreasing in this way, the species in question will ultimately become extinct.

Other Problems

The Other Problems on this sheet were intended to give you some further insight into the consequences of our model of age-specific population growth.

6. We are given the Leslie matrix

$$\mathsf{L} = \begin{bmatrix} 0 & 0 & 6 \\ 1/2 & 0 & 0 \\ 0 & 1/3 & 0 \end{bmatrix}.$$

This question is about what happens when you do not have a dominant eigenvalue, and clearly, we don't have one here because we do not have two successive fertile classes. Calculating the eigenvalues as asked, you should have found that the characteristic polynomial is $1 - \lambda^3$. Thus, the eigenvalues are the roots of the equation

$$\lambda^3 = 1$$

This has one real root, namely $\lambda = 1$, and two complex roots (which are complex conjugates because the equation has real coefficients) given by

$$\lambda = e^{\pm 2\pi i/3} = -\frac{1}{2} \pm \frac{\sqrt{3}}{2} i.$$

(These are found by noting that $1 = e^{2\pi n i}$ and taking $n = -1, 0, 1$.) Now, the modulus of these complex roots is given by

$$|\lambda|^2 = \left(\frac{1}{2}\right)^2 + \left(\pm\frac{\sqrt{3}}{2}\right)^2 = \frac{1}{4} + \frac{3}{4} = 1 \implies |\lambda| = 1$$

which is equal to the value of the unique positive real eigenvalue given by $\lambda = 1$, thus this eigenvalue is not dominant (as expected).

Calculating L^3 as asked, you should have found that

$$\mathsf{L}^3 = \begin{bmatrix} 0 & 0 & 6 \\ 1/2 & 0 & 0 \\ 0 & 1/3 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 6 \\ 1/2 & 0 & 0 \\ 0 & 1/3 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 6 \\ 1/2 & 0 & 0 \\ 0 & 1/3 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 6 \\ 1/2 & 0 & 0 \\ 0 & 1/3 & 0 \end{bmatrix} \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 3 \\ 1/6 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

This tells us that the population distribution vector $\mathbf{x}^{(k)}$, which is given by $\mathbf{x}^{(k)} = \mathsf{L}^k\mathbf{x}^{(0)}$, will evolve as follows:

$$\mathbf{x}^{(1)} = \mathsf{L}\mathbf{x}^{(0)}, \quad \mathbf{x}^{(2)} = \mathsf{L}^2\mathbf{x}^{(0)}, \quad \mathbf{x}^{(3)} = \mathsf{L}^3\mathbf{x}^{(0)}$$

But, $\mathsf{L}^3 = \mathsf{I}$, and so $\mathbf{x}^{(3)} = \mathbf{x}^{(0)}$! Thus,

$$\mathbf{x}^{(4)} = \mathsf{L}^4\mathbf{x}^{(0)} = \mathsf{L}\mathsf{L}^3\mathbf{x}^{(0)} = \mathsf{L}\mathbf{x}^{(0)} = \mathbf{x}^{(1)}, \quad \mathbf{x}^{(5)} = \mathsf{L}^5\mathbf{x}^{(0)} = \mathsf{L}^2\mathbf{x}^{(0)} = \mathbf{x}^{(2)}, \quad \mathbf{x}^{(6)} = \mathsf{L}^6\mathbf{x}^{(0)} = \mathbf{x}^{(0)},$$

and, repeating this argument for $k \geq 7$ we conclude that:

$$\mathbf{x}^{(k)} = \begin{cases} \mathbf{x}^{(0)} & \text{for } k = 0, 3, 6, \ldots \\ \mathbf{x}^{(1)} & \text{for } k = 1, 4, 7, \ldots \\ \mathbf{x}^{(2)} & \text{for } k = 2, 5, 8, \ldots \end{cases}$$

Thus, the population distribution exhibits a *cyclical* behaviour which repeats itself every three time periods.

This *seems* to contradict the key result of from the handout for Lecture 8 where we established that in the infinite time limit there will be a fixed proportion of the population in each age class (where the relevant proportion is given by the ratio of the elements in the eigenvector $\mathbf{v}_1$ corresponding to the unique real positive eigenvalue of the Leslie matrix). But, of course, the contradiction is only apparent as the analysis in the handout for Lecture 8 presupposes that the unique real positive eigenvalue is *dominant*, and as we have seen, this is not the case here.

7. We are asked to show that the *net reproduction rate*, given by

$$R = a_1 + a_2b_1 + a_3b_1b_2 + \cdots + a_nb_1b_2\ldots b_{n-1},$$

is the average number of daughters born to a female during her expected lifetime. To do this, we consider how many daughters are 'produced' in the lifetime of a certain *generation*. Here, a generation is the class of females who are born in a certain time period (i.e. a period of L/n years in our model, where L is the maximum age of a female) and, let us assume that N females are born in the time period that we are considering.[10] Now, a_1 is the average number of daughters born to a female when she is in the first age class, and so this generation would be expected to produce a_1N daughters in this initial time period of L/n years. However, only a fraction (i.e. b_1) of these females will live for more than L/n years and hence make it into the second age class. Thus, only b_1N of the females in this generation will pass into the second age class. But, on average, those that do survive will have a_2 daughters, and so during this second time period this generation would be expected to produce a_2b_1N daughters. Thus, after two time periods, i.e. $2L/n$ years, we would expect this generation to have produced $a_1N + a_2b_1N$ daughters. Repeating this argument we would expect that this generation would ultimately produce

$$a_1N + a_2b_1N + a_3b_1b_2N + \cdots + a_nb_1b_2\ldots b_{n-1}N$$

daughters.[11] Now, this is the total number of daughters born to the N females in this generation during its lifetime. Thus, the net reproduction rate, R, which is defined as the average number of daughters born to *a* female during her expected lifetime, is just this quantity divided by N (as required).

Now, assuming that two consecutive a_i are non-zero, we are asked to show that the population is eventually increasing iff its net reproduction rate is greater than 1. To 'cash this out', we notice that the two consecutive fertile classes guarantee the existence of a *dominant* real positive eigenvalue (our λ_1) and so we know that the long term behaviour is given by

$$\mathbf{x}^{(k)} \simeq c\lambda_1^k\mathbf{v}_1.$$

[10] So, obviously, these N females are all in the first age class during this time period.

[11] Notice that the series terminates because there is absolutely no chance of a female living for more than L years (by assumption) and so the expected number of daughters from a female in the ith age class, where $i > n$, is zero.

So, the population is *eventually* increasing if $\lambda_1 > 1$. This means that we have to show that $\lambda_1 > 1$ iff $R > 1$. To do this, we use the hint and recall that $q(\lambda_1) = 1$, i.e.

$$q(\lambda_1) = \frac{a_1}{\lambda_1} + \frac{a_2 b_1}{\lambda_1^2} + \frac{a_3 b_1 b_2}{\lambda_1^3} + \cdots + \frac{a_n b_1 b_2 \ldots b_{n-1}}{\lambda_1^n} = 1$$

and comparing this with our expression for R, i.e.

$$R = a_1 + a_2 b_1 + a_3 b_1 b_2 + \cdots + a_n b_1 b_2 \ldots b_{n-1}.$$

we can construct a [quick!] argument for the required result, namely:

- **RTL:** When $R > 1$, $q(\lambda_1) = 1$ can only be satisfied if $\lambda > 1$, as required.
- **LTR:** When $R < 1$, $q(\lambda_1) = 1$ can only be satisfied if $\lambda < 1$. Thus, [by the contrapositive of this,] we have: if $\lambda \geq 1$, then $R \geq 1$ and hence the required result.

(Too quick?) But, this result is obvious anyway, because demographically, each female must 'produce' [on average] *at least* one daughter in her lifetime for the population to increase!

Harder Problems

The Harder Problems on this sheet gave you the opportunity to prove some other theorems concerning the eigenvalues and eigenvectors of a matrix.

8. In this question, we generalise the result of Question 2 and consider a general[12] matrix A (which is square and at least $m \times m$) that has $n \leq m$ distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. We are then required to show that: If $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ are the eigenvectors corresponding to these eigenvalues, then the set of vectors $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ is linearly independent. This can be done by [finite] induction[13] on k where we consider the linear independence of the set of vectors $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ with $2 \leq k \leq n$, i.e.

- **Induction Hypothesis:** Let A be an [at least] $m \times m$ matrix that has $n \leq m$ distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ with corresponding eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$. If $2 \leq k \leq n$, then the set of vectors $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ is linearly independent.
- **Basis:** In the case where $k = 2$, we must show that if A is a square matrix (which is at least 2×2) that has two distinct eigenvalues λ_1 and λ_2 with corresponding eigenvectors $\mathbf{x}_1$ and $\mathbf{x}_2$, then the set of vectors $\{\mathbf{x}_1, \mathbf{x}_2\}$ is linearly independent. But, we have already done this in Question 2.
- **Induction Step:** Suppose that $2 \leq k < n$ and that the Induction Hypothesis is true for k, we want to show that if $2 < k+1 \leq n$, then the set of vectors $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k, \mathbf{x}_{k+1}\}$ is linearly independent. To do this, consider the vector equation

  $$\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \cdots + \alpha_k \mathbf{x}_k + \alpha_{k+1} \mathbf{x}_{k+1} = \mathbf{0}$$

  and as in Question 2, we multiply through by the matrix A to get

  $$\alpha_1 \mathsf{A}\mathbf{x}_1 + \alpha_2 \mathsf{A}\mathbf{x}_2 + \cdots + \alpha_k \mathsf{A}\mathbf{x}_k + \alpha_{k+1} \mathsf{A}\mathbf{x}_{k+1} = \mathsf{A}\mathbf{0}$$

  But, we have $k+1$ eigenvalue-eigenvector relations, i.e.

  $$\mathsf{A}\mathbf{x}_1 = \lambda_1 \mathbf{x}_1, \quad \mathsf{A}\mathbf{x}_2 = \lambda_2 \mathbf{x}_2, \quad \ldots, \quad \mathsf{A}\mathbf{x}_k = \lambda_k \mathbf{x}_k \quad \text{and} \quad \mathsf{A}\mathbf{x}_{k+1} = \lambda_{k+1} \mathbf{x}_{k+1}$$

  which therefore gives us a second vector equation

  $$\alpha_1 \lambda_1 \mathbf{x}_1 + \alpha_2 \lambda_2 \mathbf{x}_2 + \cdots + \alpha_n \lambda_k \mathbf{x}_k + \alpha_{k+1} \lambda_{k+1} \mathbf{x}_{k+1} = \mathbf{0}$$

[12] Again, the restriction that the matrix should be $n \times n$ is irrelevant (but in this case, it is not particularly distracting).

[13] Alternatively, you could follow Anton and Rorres and use a proof by contradiction.

Consequently, if we multiply our original vector equation by λ_{k+1} and subtract it from this new vector equation we get

$$\alpha_1(\lambda_1 - \lambda_{k+1})\mathbf{x}_1 + \alpha_2(\lambda_2 - \lambda_{k+1})\mathbf{x}_2 + \cdots + \alpha_k(\lambda_k - \lambda_{k+1})\mathbf{x}_k = \mathbf{0}$$

Now, by the Induction Hypothesis, the set of vectors $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ is linearly independent which means that the coefficients in this vector equation must be zero, i.e.

$$\alpha_1(\lambda_1 - \lambda_{k+1}) = \alpha_2(\lambda_2 - \lambda_{k+1}) = \cdots = \alpha_k(\lambda_k - \lambda_{k+1}) = 0$$

and as the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct this implies that $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$. So, substituting this into the original vector equation, we get

$$\alpha_{k+1}\mathbf{x}_{k+1} = \mathbf{0}$$

and as $\mathbf{x}_{k+1} \neq \mathbf{0}$ (it is an eigenvector), we have $\alpha_{k+1} = 0$ too. Consequently, our original vector equation has the trivial solution as its only solution and so the set of vectors $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k, \mathbf{x}_{k+1}\}$ is linearly independent.

Thus, by the Principle of Induction, if A is an $m \times m$ matrix that has $n \leq m$ distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ with corresponding eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ and $2 \leq k \leq n$, then the set of vectors $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ is linearly independent, as required.

9. For an $n \times n$ matrix A, we are asked to prove that: A is diagonalisable iff A has n linearly independent eigenvectors. As this is an 'if and only if' claim, we have to prove it both ways:

LTR: We assume that the matrix A is diagonalisable, and so there is an invertible matrix, say P , with column vectors denoted by $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$, i.e.

$$\mathsf{P} = \begin{bmatrix} | & | & & | \\ \mathbf{p}_1 & \mathbf{p}_2 & \cdots & \mathbf{p}_n \\ | & | & & | \end{bmatrix}$$

such that $\mathsf{P}^{-1}\mathsf{A}\mathsf{P}$ gives a diagonal matrix, say D, where

$$\mathsf{D} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

It follows from the way that we have chosen P and D that $\mathsf{AP} = \mathsf{PD}$, which means that as

$$\mathsf{AP} = \begin{bmatrix} | & | & & | \\ \mathsf{A}\mathbf{p}_1 & \mathsf{A}\mathbf{p}_2 & \cdots & \mathsf{A}\mathbf{p}_n \\ | & | & & | \end{bmatrix}$$

and

$$\mathsf{PD} = \begin{bmatrix} | & | & & | \\ \mathbf{p}_1 & \mathbf{p}_2 & \cdots & \mathbf{p}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \lambda_1\mathbf{p}_1 & \lambda_2\mathbf{p}_2 & \cdots & \lambda_n\mathbf{p}_n \\ | & | & & | \end{bmatrix}$$

it must be the case that the successive columns of AP and PD are equal, i.e.

$$\mathsf{A}\mathbf{p}_1 = \lambda_1\mathbf{p}_1, \quad \mathsf{A}\mathbf{p}_2 = \lambda_2\mathbf{p}_2, \quad \ldots, \quad \mathsf{A}\mathbf{p}_{n-1} = \lambda_{n-1}\mathbf{p}_{n-1} \quad \text{and} \quad \mathsf{A}\mathbf{p}_n = \lambda_n\mathbf{p}_n$$

Now, as P is invertible, we can deduce that:

- all of the column vectors of P must be *non-zero* and so the vectors $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$ are eigenvectors of A corresponding to the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$.

- $\det(\mathsf{P})$ must be non-zero and so the set of vectors $\{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n\}$ is linearly independent.

Thus, the matrix A has n linearly independent eigenvectors (as required).

RTL: Assume that the matrix A has n linearly independent eigenvectors $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$ corresponding to the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and let the matrix P be such that

$$\mathsf{P} = \begin{bmatrix} | & | & & | \\ \mathbf{p}_1 & \mathbf{p}_2 & \cdots & \mathbf{p}_n \\ | & | & & | \end{bmatrix}$$

Multiplying the two matrices A and P together, we get

$$\mathsf{AP} = \begin{bmatrix} | & | & & | \\ \mathsf{A}\mathbf{p}_1 & \mathsf{A}\mathbf{p}_2 & \cdots & \mathsf{A}\mathbf{p}_n \\ | & | & & | \end{bmatrix}$$

and we know that we have n eigenvalue-eigenvector relations, i.e.

$$\mathsf{A}\mathbf{p}_1 = \lambda_1\mathbf{p}_1, \quad \mathsf{A}\mathbf{p}_2 = \lambda_2\mathbf{p}_2, \quad \ldots, \quad \mathsf{A}\mathbf{p}_{n-1} = \lambda_{n-1}\mathbf{p}_{n-1} \quad \text{and} \quad \mathsf{A}\mathbf{p}_n = \lambda_n\mathbf{p}_n$$

which on substituting gives

$$\mathsf{AP} = \begin{bmatrix} | & | & & | \\ \lambda_1\mathbf{p}_1 & \lambda_2\mathbf{p}_2 & \cdots & \lambda_n\mathbf{p}_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \mathbf{p}_1 & \mathbf{p}_2 & \cdots & \mathbf{p}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = \mathsf{PD}$$

where D is the diagonal matrix with the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ along its 'main' diagonal. Further, as the column vectors of P are linearly independent, $\det(\mathsf{P})$ is non-zero, and so the matrix P is invertible. Consequently, we can re-write the above expression as $\mathsf{P}^{-1}\mathsf{AP} = \mathsf{D}$, which means that the matrix A is diagonalisable (as required).

10. We are asked to prove that: $\lambda = 0$ is an eigenvalue of a matrix A iff A is not invertible. To do this we just note that if $\lambda = 0$ is an eigenvalue of the matrix A , it is the case that

$$\det(\mathsf{A} - 0\mathsf{I}) = 0$$

which is equivalent to asserting that $\det(\mathsf{A}) = 0$, i.e. the matrix A is not invertible and *vice versa* (as required).[14]

11. We are asked to prove that: If A is an $n \times n$ matrix, then

$$\det(\mathsf{A} - \lambda\mathsf{I}) = (-1)^n\lambda^n + c_1\lambda^{n-1} + \cdots + c_{n-1}\lambda + c_n$$

where the c_i ($1 \leq i \leq n$) are constants. To do this, we note that the eigenvalues of an $n \times n$ matrix A are such that

$$\mathsf{A}\mathbf{x} = \lambda\mathbf{x}$$

for some vector $\mathbf{x} = [x_1, x_2, \ldots, x_n]^t \neq \mathbf{0}$. But, re-writing this as

$$(\mathsf{A} - \lambda\mathsf{I})\mathbf{x} = \mathbf{0}$$

and denoting the column vectors of the matrix $\mathsf{A} - \lambda\mathsf{I}$ by $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ we find that this matrix equation is equivalent to the vector equation

$$x_1\mathbf{u}_1 + x_2\mathbf{u}_2 + \cdots + x_n\mathbf{u}_n = \mathbf{0}$$

[14] Notice that this result should be obvious as the eigenvalues are the numbers which make the columns of the matrix $\mathsf{A} - \lambda\mathsf{I}$ linearly dependent. Consequently, if $\lambda = 0$ is an eigenvalue, the columns of the matrix A are linearly dependent, and so this matrix is not invertible (as $\det(\mathsf{A})$, say, will then be zero).

Now, as $\mathbf{x} \neq \mathbf{0}$, we have non-trivial solutions to this vector equation and so the eigenvalues must be such that they make the column vectors of the matrix $\mathsf{A} - \lambda\mathsf{I}$ linearly dependent. That is, if λ is an eigenvalue of the matrix A, then the matrix $\mathsf{A} - \lambda\mathsf{I}$ has linearly dependent columns and this means that $\det(\mathsf{A} - \lambda\mathsf{I}) = 0$ (i.e. the matrix $\mathsf{A} - \lambda\mathsf{I}$ is not invertible).

Next, we note that the matrix $\mathsf{A} - \lambda\mathsf{I}$, which has column vectors $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ say, can be written as

$$\mathsf{A} - \lambda\mathsf{I} = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{21} & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{n1} & u_{n2} & \cdots & u_{nn} \end{bmatrix} = \begin{bmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{bmatrix}$$

and will have a determinant given by[15]

$$\det(\mathsf{A} - \lambda\mathsf{I}) = \sum_{\pi} \tilde{\pi} u_{1j_1} u_{2j_2} \cdots u_{nj_n}$$

where the summation is over all $n!$ permutations π of the ordered n-tuple $(j_1, j_2, \ldots, j_n)$ with $1 \leq j_i \leq n$ for all i such that $1 \leq i \leq n$ and $\tilde{\pi}$ is $+1$ or -1 if the permutation π is even or odd respectively. In particular, it should be clear that *if* we chose to write this summation out, we would get a polynomial in λ of degree n as the the term in the summation corresponding to the permutation where the ordered n-tuple is $(1, 2, \ldots, n)$ is

$$(a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda)$$

and as this permutation is even, $\tilde{\pi} = +1$.[16] Consequently, we can see that as the eigenvalues satisfy the equation $\det(\mathsf{A} - \lambda\mathsf{I}) = 0$, they must also satisfy the equation $p(\lambda) = 0$ where $p(\lambda)$ is a polynomial in λ of degree n. Also notice that the λ^n term in this summation will have a coefficient of $(-1)^n$. As such, [including multiplicity and the possibility that some eigenvalues are complex,] there must be n solutions[17] $\lambda_1, \lambda_2, \ldots, \lambda_n$ to these equations, and as such they will be roots of the polynomial $p(\lambda)$, i.e. we have

$$p(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \cdots (\lambda_n - \lambda)$$

which on expanding would give us a polynomial of the desired degree.

But, we are not quite there yet. So far, we have shown that the eigenvalues can be seen as solutions of

$$\det(\mathsf{A} - \lambda\mathsf{I}) = 0 \quad \text{or} \quad p(\lambda) = 0$$

However, this only guarantees that

$$\det(\mathsf{A} - \lambda\mathsf{I}) = kp(\lambda)$$

for some constant k. Although, we have also found that the λ^n term in the expansion of $\det(\mathsf{A} - \lambda\mathsf{I})$ must have a coefficient of $(-1)^n$ and so it should be clear that the same term in the polynomial $p(\lambda)$ will have this coefficient too. Thus, for the equality above to hold, it must be the case that $k = 1$. Consequently, it should be obvious that if we expand our expression for $p(\lambda)$, we will get

$$\det(\mathsf{A} - \lambda\mathsf{I}) = p(\lambda) = (-1)^n \lambda^n + c_1 \lambda^{n-1} + \cdots + c_{n-1}\lambda + c_n$$

for some constants c_i where $1 \leq i \leq n$ (as required).

Further, if we let $\lambda = 0$ in this expression it reduces to

$$\det(\mathsf{A}) = c_n$$

(as required).

[15] For example, see Anton and Rorres, pp. 81-5.

[16] You should convince yourself that every *other* possible permutation leads to a term in the summation which has *at most* $n - 1$ [linear] factors involving λ.

[17] That is, eigenvalues.