

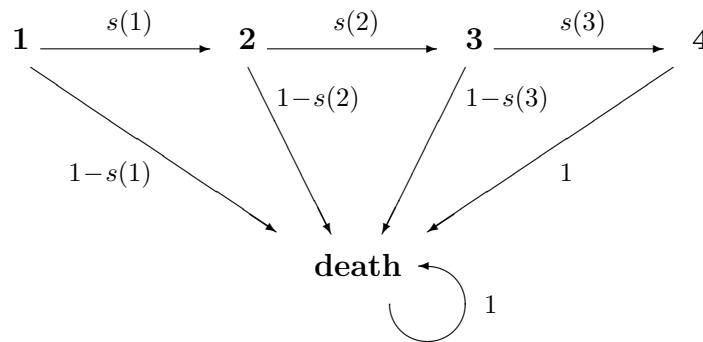
9 Demography

We have already considered population dynamics and growth using the differential reproduction model in Chapter 3. For simplicity, that model assumed that the entire population “turns over” each period. In other words, it is a *discrete generations* model. In contrast, demographers are concerned with *age-structured* populations.

9.1 Survival

From a demographic perspective, life is essentially an absorbing Markov chain. Each individual begins life in the 1st age class (e.g., age 0), from which she may transition to the 2nd age class (e.g., age 1) or else death (the sole absorbing state). More generally, from any age class i , the individual either transitions (“survives”) to age class $i + 1$ with probability $s(i)$ or else dies with probability $1 - s(i)$. If no one ever survives beyond age class n , so that $s(n) = 0$, we thus obtain an absorbing chain with $n + 1$ states.

To illustrate, we’ll adopt a coarser partition of the population into 4 age classes (although our results extend easily to the general case with n age classes). For concreteness, consider a society where no one lives beyond age 79, so that these classes can be interpreted as ages 0-19, 20-39, 40-59, and 60-79. Relatedly, each “period” should be interpreted a 20-year interval. The transition diagram is shown below.



Alternatively, specifying the probability transition matrix in canonical form (with the absorbing state listed first), we can write

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1-s(1) & 0 & s(1) & 0 & 0 \\ 1-s(2) & 0 & 0 & s(2) & 0 \\ 1-s(3) & 0 & 0 & 0 & s(3) \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

For our present purposes, we could employ this transition matrix directly. However, following the convention in demography, we will use the closely related *survival* matrix. We can get from the transition matrix to the survival matrix in two steps. First, as we saw in Chapter 4, much can be learned about an absorbing chain with transition matrix P simply from its lower right submatrix Q characterizing transitions from non-absorbing states to non-absorbing states. For the P matrix above,

$$Q = \begin{bmatrix} 0 & s(1) & 0 & 0 \\ 0 & 0 & s(2) & 0 \\ 0 & 0 & 0 & s(3) \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Second, when demographers specify transition matrices, they adopt the convention that element (i, j) denotes the probability of transition *to* age class i *from* age class j . Consequently, demographers specify the survival matrix as

$$S = \begin{bmatrix} 0 & 0 & 0 & 0 \\ s(1) & 0 & 0 & 0 \\ 0 & s(2) & 0 & 0 \\ 0 & 0 & s(3) & 0 \end{bmatrix}$$

which is the *transpose* of the Q matrix. While it is tempting to say that demographers specify transition matrices “backwards,” we have already seen one rationale for this convention in Chapter 3. Recall that Theorem 2 involved the eigenvalues and eigenvectors of the *transpose* of the RP matrix. Implicitly adopting the demography convention, we restated the population dynamics as $\mathbf{x}_t = A\mathbf{x}_0$ where $A = (RP)'$ and \mathbf{x}_t is a column vector. Had we adopted the demography convention from the outset, there would have been no need to transpose A before computing its eigenvalues and eigenvectors. So perhaps it is our original convention that is “backwards.”¹

We saw in Chapter 4 that $Q^t(i, j)$ can be interpreted as the probability that a process initially in state i will occupy state j after t periods. Taking care to note the transposition of i and j , we can thus interpret $S^t(j, i)$ as the probability that an individual initially in age class i will survive to age class j after t periods. Given the S matrix above, iterated multiplication yields

$$S^2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ s(1)s(2) & 0 & 0 & 0 \\ 0 & s(2)s(3) & 0 & 0 \end{bmatrix},$$

¹On the other hand, this convention is well-entrenched in literatures on Markov chains (e.g., Kemeny and Snell 1960) and social network analysis (e.g., Wasserman and Faust 1994). Given the existence of dual conventions, we’ll continue to alternate between them.

$$S^3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ s(1)s(2)s(3) & 0 & 0 & 0 \end{bmatrix},$$

and it is easy to see that every element of S^t is equal to 0 for all $t \geq 4$. Intuitively, after 2 periods, an individual initially in age class 1 survives to age class 3 with probability $s(1)s(2)$, while an individual initially in class 2 survives to class 4 with probability $s(2)s(3)$. Similarly, after 3 periods, an individual initially in age class 1 survives to class 4 with probability $s(1)s(2)s(3)$. All elements of S^t are 0 for any $t \geq 4$ because there is no chance of surviving beyond the 4th period.

Further recall from Chapter 4 that the fundamental matrix

$$N = I + Q + Q^2 + \dots = (I - Q)^{-1}$$

can be used to determine the expected number of periods until absorption from each non-absorbing state. Relatedly, the matrix

$$N' = I + S + S^2 + \dots = (I - S)^{-1}$$

is useful for survival analysis because $N'(j, i)$ can be interpreted as the probability of survival from age class i to age class j . Given 4 age classes, this matrix can be written generically as

$$N' = \begin{bmatrix} 1 & 0 & 0 & 0 \\ s(1) & 1 & 0 & 0 \\ s(1)s(2) & s(2) & 1 & 0 \\ s(1)s(2)s(3) & s(2)s(3) & s(3) & 1 \end{bmatrix}.$$

The first column of this matrix is of special interest because it reports survival probabilities from birth (i.e., computed for someone who initially occupies the first age class). For our present demographic application, the sole absorbing state is death, and hence the “expected number of periods to absorption” becomes simply “life expectancy.” The sum of the i th *column* of the N' matrix thus reports the expected number of periods remaining for an individual currently in age class i .

Using the first column and the column sums of the N' matrix, we can construct the *life table* shown below.

<i>age class</i>	<i>survival probabilities</i> (from age class 1)	<i>expected periods remaining</i> (from current age class)
1	1	$1 + s(1) + s(1)s(2) + s(1)s(2)s(3)$
2	$s(1)$	$1 + s(2) + s(2)s(3)$
3	$s(1)s(2)$	$1 + s(3)$
4	$s(1)s(2)s(3)$	1

If we had adopted a finer partition of the age distribution into one-year intervals – so that “age class” becomes simply “age” – then the third column of the life table would become simply “life expectancy.” Due to our coarser specification of age classes, we need to multiply each element in the final column by period length (20 years) in order to obtain expected number of *years* remaining.²

To illustrate, let’s assume some hypothetical survival rates, setting $s(1) = 0.9$, $s(2) = 0.8$, and $s(3) = 0.7$. The resulting life table (along with intermediate computations) is shown below.

```
>> S = [0 0 0 0; .9 0 0 0; 0 .8 0 0; 0 0 .7 0] % survival matrix

S =
     0         0         0         0
    0.9000         0         0         0
     0     0.8000         0         0
     0         0     0.7000         0

>> S^2 % survival after 2 periods

ans =
     0         0         0         0
     0         0         0         0
    0.7200         0         0         0
     0     0.5600         0         0

>> S^3 % survival after 3 periods

ans =
     0         0         0         0
     0         0         0         0
     0         0         0         0
    0.5040         0         0         0

>> N = eye(4) + S + S^2 + S^3 % (transpose of) fundamental matrix

N =
    1.0000         0         0         0
    0.9000     1.0000         0         0
    0.7200     0.8000     1.0000         0
    0.5040     0.5600     0.7000     1.0000

>> lifetable = [(1:4)' N(:,1) sum(N)'] % life table

lifetable =
    1.0000     1.0000     3.1240
```

²This computation implicitly assumes that everyone who enters an age class lives through the entire interval. That is, death occurs only at the end of each period. Of course, in reality, deaths occur continuously throughout each period. While further discussion is beyond our present scope, demographers would account for this by making “corrections” to life tables.

2.0000	0.9000	2.3600
3.0000	0.7200	1.7000
4.0000	0.5040	1.0000

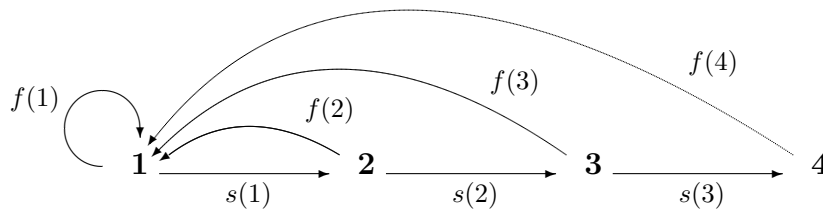
Thus, given our hypothetical survival rates, an individual in age class 1 has a 90% chance of reaching age class 2, a 72% chance of reaching age class 3, and a 50.4% chance of reaching age class 4. At birth (i.e., upon entry into age class 1), she is expected to live 3.124 periods (= 62.48 years). If she survives to age class 2 (reaching age 20), she is expected to live an additional 2.36 periods (= 47.2 years). If she survives to age class 3 (reaching age 40), she is expected to live an additional 1.7 periods (= 34 years). Of course, given our assumption that no one lives beyond age class 4, those who enter this class (at age 60) will live exactly one more period (= 20 years).³

9.2 Fertility

Having addressed survival, we can now expand our focus to include *fertility* and thereby obtain a specification of population dynamics. Letting $f(i)$ denote the average number of children born to individuals in age class i , and $\mathbf{x}_t(i)$ denote the number of individuals in age class i in period t , the age distribution in period $t + 1$ is determined by the following system of equations,

$$\begin{aligned} \mathbf{x}_{t+1}(1) &= f(1)\mathbf{x}_t(1) + f(2)\mathbf{x}_t(2) + f(3)\mathbf{x}_t(3) + f(4)\mathbf{x}_t(4) \\ \mathbf{x}_{t+1}(2) &= s(1)\mathbf{x}_t(1) \\ \mathbf{x}_{t+1}(3) &= s(2)\mathbf{x}_t(2) \\ \mathbf{x}_{t+1}(4) &= s(3)\mathbf{x}_t(3) \end{aligned}$$

which can be illustrated by the diagram below.⁴



Intuitively, individuals currently in age class i contribute to next period's population in two ways: by surviving to age class $i + 1$, and by "sending back" children who

³Again, demographers would correct life tables to account for deaths occurring within periods. But these computations help connect our present analysis to our analysis of absorbing Markov chains in Chapter 4.

⁴Technically, this diagram is not a transition diagram because the model is not a Markov chain process. Nevertheless, the diagram remains useful for visualizing the system of equations.

begin life in age class 1. To specify population dynamics in more elegant matrix notation, we can construct the matrix

$$L = \begin{bmatrix} f(1) & f(2) & f(3) & f(4) \\ s(1) & 0 & 0 & 0 \\ 0 & s(2) & 0 & 0 \\ 0 & 0 & s(3) & 0 \end{bmatrix},$$

which demographers call the *Leslie* matrix. The system of equation given above can now be written as

$$\mathbf{x}_{t+1} = L\mathbf{x}_t$$

where \mathbf{x}_t is a column vector. Given the initial age distribution \mathbf{x}_0 , we obtain

$$\begin{aligned} \mathbf{x}_1 &= L\mathbf{x}_0 \\ \mathbf{x}_2 &= L\mathbf{x}_1 = L^2\mathbf{x}_0 \\ \mathbf{x}_3 &= L\mathbf{x}_2 = L^3\mathbf{x}_0 \end{aligned}$$

and hence by induction

$$\mathbf{x}_t = L^t\mathbf{x}_0.$$

By this point in the course, the form of this equation should be familiar. However, let me emphasize (one last time) that we now following the demography convention, postmultiplying L^t by a column vector rather than premultiplying by a row vector.

Given that this model provides the standard demographic framework for studying population dynamics, it may be surprising to observe that it is a one-sex model. That is, we are continuing to maintain the fiction that each individual has one parent. Of course, demographers are well aware that children have both mothers and fathers. But perhaps due to the complexity of two-sex models (which we will address in Chapter xx), the linear model based on the Leslie matrix remains the standard framework. Given its one-sex formulation, the model is best interpreted as a model of the *female* population, so that \mathbf{x}_t is the distribution of the *female* population across age classes.⁵ Having clarified this point, we might further recognize that fertility rates will be zero for women who are either too young or too old to have children. But this fact is easily captured within the present framework by setting $f(i) = 0$ when i is very low or very high.

Demographers have developed two measures to summarize the level of fertility within a population. The first measure, called the *gross reproduction rate*, is defined as

$$\text{GRR} = f(1) + f(2) + f(3) + f(4).$$

⁵Recognizing that men and women have different survival probabilities, demographers would conduct a “side” analysis of the age structure of the male population. But because only women have children, they constitute the “engine” of population dynamics.

Mathematically, the GRR is simply the sum of the first row of the Leslie matrix (reporting the average fertility level for each age class). Conceptually, the GRR represents the expected number of children that a women would bear if she survives through the final age class. The second measure is the *net reproduction rate*, defined as

$$\begin{aligned} \text{NRR} &= [f(1) \quad f(2) \quad f(3) \quad f(4)] \begin{bmatrix} 1 \\ s(1) \\ s(1)s(2) \\ s(1)s(2)s(3) \end{bmatrix} \\ &= f(1) + s(1)f(2) + s(1)s(2)f(3) + s(1)s(2)s(3)f(4). \end{aligned}$$

Mathematically, the NRR is the inner product of the first row of the Leslie matrix and the first column of the life table (reporting survival probabilities from age class 1). Conceptually, this measure adjusts downward the GRR to account for the fact that some women die before reaching the end of their child-bearing years. Intuitively, we might expect that the population size is rising given $\text{NRR} > 1$, stable given $\text{NRR} = 1$, and falling given $\text{NRR} < 1$. This result will be established formally in Section xx, where we also examine the relationship between the NRR and the dominant eigenvalue of the Leslie matrix.

To illustrate, consider the numerical example below.

```
>> L = S; L(1,:) = [0 1.2 .5 0] % Leslie matrix
```

```
L =
      0      1.2000      0.5000      0
  0.9000      0      0      0
      0      0.8000      0      0
      0      0      0.7000      0
```

```
>> GRR = sum(L(1,:)) % gross reproduction rate
```

```
GRR =
  1.7000
```

```
>> NRR = L(1,:) * N(:,1) % net reproduction rate
```

```
NRR =
  1.4400
```

Further assuming an initial age distribution, it is straightforward to compute the age distribution for the next 20 periods (representing 400 years). Demographers would refer to these sorts of computations as population “projections.”

```
>> x0 = [1 .9 .72 .504]'; % initial age distribution (period 0)
```

```
>> popfreq = [ ]; for t = 0:20; popfreq = [popfreq; (L^t * x0)']; end; popfreq
```

```
popfreq =
    1.0000    0.9000    0.7200    0.5040
    1.4400    0.9000    0.7200    0.5040
    1.4400    1.2960    0.7200    0.5040
    1.9152    1.2960    1.0368    0.5040
    2.0736    1.7237    1.0368    0.7258
    2.5868    1.8662    1.3789    0.7258
    2.9290    2.3281    1.4930    0.9653
    3.5403    2.6361    1.8625    1.0451
    4.0945    3.1862    2.1089    1.3038
    4.8779    3.6851    2.5490    1.4762
    5.6966    4.3901    2.9481    1.7843
    6.7422    5.1269    3.5121    2.0636
    7.9084    6.0680    4.1015    2.4585
    9.3323    7.1175    4.8544    2.8711
   10.9682    8.3991    5.6940    3.3981
   12.9259    9.8714    6.7193    3.9858
   15.2053   11.6333    7.8971    4.7035
   17.9085   13.6848    9.3067    5.5280
   21.0750   16.1177   10.9478    6.5147
   24.8151   18.9675   12.8941    7.6635
   29.2081   22.3336   15.1740    9.0259
```

While these projections might seem similar in form to those obtained from the differential reproduction model in Chapter 4, recall that we are now considering an *age structured* population. It is no longer true that the entire population “turns over” each period. Rather, we can follow the progress of those individuals born in some period – the *birth cohort* for that period – as they progress through their life course. To illustrate, consider the fourth row of the `popfreq` table (which corresponds to period 3 since we started indexing from period 0). In period 3, there were 1.9152 individuals born into age class 1. In period 4, the number of survivors in this cohort is $(1.9152)(0.9) = 1.7237$, which appears as the 2nd element in the 5th row of the table. Similarly, after 2 periods, the number of survivors in this cohort is $(1.7237)(0.8) = 1.3789$, which appears as the 3rd element in the 6th row of the table. Finally, after 3 periods, the number of survivors in this birth cohort has dwindled to $(1.3789)(0.7) = 0.9653$, which appears as the 4th element in the 7th row of the table. Schematically, we might keep track of birth cohorts using the following diagram, which is conceptually similar to “Lexis” diagrams used by demographers.


```

0.4040    0.2525    0.2020    0.1414
0.3636    0.3273    0.1818    0.1273
0.4030    0.2727    0.2182    0.1061
0.3730    0.3100    0.1865    0.1305
0.3945    0.2846    0.2103    0.1107
0.3796    0.3018    0.1935    0.1251
0.3897    0.2902    0.2050    0.1150
0.3829    0.2980    0.1972    0.1219
0.3875    0.2927    0.2025    0.1173
0.3844    0.2962    0.1989    0.1204
0.3865    0.2939    0.2013    0.1183
0.3851    0.2955    0.1997    0.1197
0.3860    0.2944    0.2008    0.1188
0.3854    0.2951    0.2001    0.1194
0.3858    0.2946    0.2006    0.1190
0.3855    0.2950    0.2002    0.1193
0.3857    0.2948    0.2005    0.1191
0.3856    0.2949    0.2003    0.1192
0.3857    0.2948    0.2004    0.1191
0.3856    0.2949    0.2003    0.1192

```

Thus, we find that the population is indeed converging to a stable growth equilibrium, with a long-run growth rate of 17.72% and a limiting distribution of the population with approximately 38.56% in age class 1, 29.49% in age class 2, 20.03% in age class 3, and 11.92% in age class 4.

Based on our analysis in Chapter 3, you might further anticipate (again correctly) that the long-run growth rate and limiting distribution could have been found directly by computing the eigenvalues and eigenvectors of the Leslie matrix.

```

>> [eigvec, eigval] = eig(L)

eigvec =
    0    0.7161   -0.4627   -0.0975
    0    0.5475    0.5270    0.2268
    0    0.3721   -0.5336   -0.4689
    1.0000  0.2212    0.4727    0.8481

eigval =
    0         0         0         0
    0    1.1772         0         0
    0         0   -0.7902         0
    0         0         0   -0.3870

>> eigvec(:,2)/sum(eigvec(:,2)); ans'

ans =
    0.3857    0.2948    0.2004    0.1191

```

In this case, Matlab has listed the dominant eigenvalue and the associated eigenvector in the *second* column of these matrices. However, the eigenvalue and eigenvector

reported in the first column also warrant discussion. If you return to Chapter 3 and inspect closely the relevant computations, you will find that only the dominant eigenvalue is associated with a non-negative eigenvector (i.e., an eigenvector with all non-negative elements). (—Indeed, we made use of this result in our discussion of Theorem 2.—) In contrast, for the Leslie matrix in the current example, we see that both the eigenvalue in the first column ($= 0$) and the dominant eigenvalue reported in the second column ($= 1.1772$) are associated with non-negative eigenvalues.

On one hand, there is a simple demographic interpretation of the results in the first column. Suppose we start from an initial age distribution in which everyone belongs to the 4th age class. After one period, those original members of the population will have died, leaving behind no children. Thus, the size of the population falls immediately to zero and remains at this level forever. On the other hand, the existence of multiple non-negative eigenvectors raises an important question about the long-run dynamics of age-structured populations. Will the population generally converge to the stable growth equilibrium determined by the dominant eigenvalue and associated eigenvector? Or can other long-run outcomes arise? To answer this question, we need to look more closely at the Leslie matrix.

9.3 Reducibility of the Leslie matrix

For human populations, women in higher age classes have zero fertility levels.⁶ We may thus write the set of age classes as $\{1, \dots, k, \dots, n\}$ where k is the highest age class with positive fertility (so that $f(k) > 0$ and $f(i) = 0$ for all $i > k$). Recalling the structure of transition diagram for the Leslie matrix from the previous section, the reduced transition diagram is

$$\{1, \dots, k\} \rightarrow \{k + 1\} \rightarrow \{k + 2\} \cdots \rightarrow \{n\}$$

and hence the reduced influence diagram is

$$\{1, \dots, k\} \leftarrow \{k + 1\} \leftarrow \{k + 2\} \cdots \leftarrow \{n\}$$

It is apparent that the Leslie matrix is reducible (and hence not primitive). Thus, Theorem 2 does not guarantee that the population will converge to the stable growth equilibrium. However, the reducibility of this simultaneous-equation system allows us to solve sequentially for the long-run outcome. Letting L_1 denote the first $k \times k$ submatrix of the Leslie matrix (corresponding to the first communication class), we

⁶This is not always true for non-human populations. See, e.g., the Leslie matrix for sandhill cranes in Mooney and Swift (A Course in Mathematical Modeling, 1999).

can begin by solving the k -equation system

$$\lambda \begin{bmatrix} \mathbf{x}(1) \\ \vdots \\ \mathbf{x}(k) \end{bmatrix} = L_1 \begin{bmatrix} \mathbf{x}(1) \\ \vdots \\ \mathbf{x}(k) \end{bmatrix}.$$

If L_1 is primitive, then we obtain a unique population growth factor λ (given by the dominant eigenvalue of L_1) and a limiting distribution $\{\mathbf{x}(1), \dots, \mathbf{x}(k)\}$ that is unique up to a scaling factor. We can then solve for the next element of the limiting distribution using the equation

$$\lambda \mathbf{x}(k+1) = s(k) \mathbf{x}(k).$$

More generally, having obtained $\mathbf{x}(k+t)$, we can then solve for the next element of the limiting distribution using the equation

$$\lambda \mathbf{x}(k+t+1) = s(k+t) \mathbf{x}(k+t).$$

In this way, we ultimately obtain the entire limiting distribution $\{\mathbf{x}(1), \dots, \mathbf{x}(n)\}$.

To illustrate, consider again our numerical example from the preceding section. Given $k = 3$ and $n = 4$, we already know that the two communication classes will be $\{1, 2, 3\}$ and $\{4\}$. However, this result could also have been established using our “recipe” for finding communication classes from Chapter 7.

```
>> L % Leslie matrix
```

```
L =
```

```
      0      1.2000      0.5000      0
    0.9000      0          0          0
      0      0.8000      0          0
      0          0      0.7000      0
```

```
>> R = (eye(4) + L)^3 > 0 % reachability matrix
```

```
R =
```

```
      1      1      1      0
      1      1      1      0
      1      1      1      0
      1      1      1      1
```

```
>> C = R & R' % can-reach-and-be-reached-by matrix
```

```
C =
```

```
      1      1      1      0
      1      1      1      0
      1      1      1      0
      0      0      0      1
```

```

>> U = unique(C, 'rows') % communication classes

U =
    0    0    0    1
    1    1    1    0

>> M = U*L*U' > 0

>> M = M & ~eye(2) % image matrix for reduced influence graph

M =
    0    1
    0    0

```

Focusing now on the submatrix for the first communication class, we can first check for primitivity, and then solve for the first $k = 3$ elements of the limiting distribution.

```

>> L1 = L(1:3,1:3) % submatrix for class {1,2,3}

L1 =
    0    1.2000    0.5000
   0.9000    0    0
    0    0.8000    0

>> L1^5 % to demonstrate primitivity

ans =
   0.7776    1.5437    0.5832
   1.0498    0.7776    0.1620
   0.2592    0.9331    0.3888

>> [eigvec, eigval] = eig(L1)

eigvec =
   0.7343   -0.5251    0.1841
   0.5614    0.5981   -0.4280
   0.3815   -0.6055    0.8848

eigval =
   1.1772    0    0
    0   -0.7902    0
    0    0   -0.3870

>> lambda = eigval(1,1)

lambda =
    1.1772

>> x = eigvec(:,1); x'

```

```
ans =
    0.7343    0.5614    0.3815
```

Note that the dominant eigenvalue for the L_1 submatrix ($= 1.1772$) is the same as the dominant eigenvalue for the entire Leslie matrix (reported in the preceding section). Essentially, the first communication class is the “engine” of population growth, while the remaining classes are merely “pulled along.”

Having solved the 3-equation system corresponding to class $\{1, 2, 3\}$, we can now proceed to communication class $\{4\}$.

```
>> x(4) = x(3)*L(4,3)/lambda; x'
```

```
ans =
    0.7343    0.5614    0.3815    0.2269
```

```
>> x/sum(x); ans' % normalizing to obtain probability vector
```

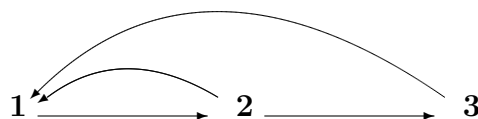
```
ans =
    0.3857    0.2948    0.2004    0.1191
```

Normalizing, we thus obtain the same limiting distribution reported in the preceding section. If our example included higher age classes (beyond 4) with zero fertility, we could continue in the same sequential fashion to derive equilibrium frequencies for those classes as well.

9.4 An alternative test for primitivity

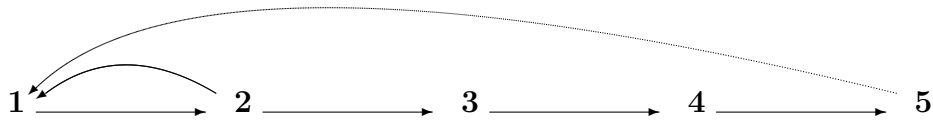
The answer to the question posed at the end of section 9.2 is now apparent. If (i) some members of the initial population have positive fertility levels (i.e., $\mathbf{x}_0(i) > 0$ for some $i \leq k$) and (ii) the first $k \times k$ submatrix of the Leslie matrix (i.e., L_1) is primitive, then the population will eventually reach the stable growth equilibrium. But while condition (ii) holds for our numerical example, should we expect it to hold more generally? The answer is yes. To see why, it will be useful to introduce an alternative test for primitivity.

To this point, we have assessed the primitivity of matrix A by checking whether there is some t such that all elements of A^t are positive. However, the primitivity of a matrix can also be assessed using the transition diagram of its zero pattern. In particular, an irreducible matrix is primitive if and only if the transition diagram of its zero pattern has at least two cycles whose lengths l_1 and l_2 are relatively prime. That is, the greatest common divisor (GCD) of l_1 and l_2 is 1. To illustrate, consider the transition diagram for the zero pattern of the L_1 matrix in our numerical example.



Inspection of this diagram reveals a cycle of length 2 (from state 1 to 2 to 1) and another cycle of length 3 (from state 1 to 2 to 3 to 1).⁷ Because $\text{GCD}(2, 3) = 1$, the L_1 matrix is primitive.

We can now see why L_1 matrix will primitive more generally (for any empirical Leslie matrix). Given two consecutive age classes j and $j + 1$ with positive fertility, the transition diagram will contain cycles of length j and $j + 1$. Hence $\text{GCD}(j, j + 1) = 1$ and the submatrix is primitive. However, even this (rather weak) condition is not necessary. Consider, for example, the transition diagram below.



Given that this diagram contains cycles of length 2 and 5, we obtain $\text{GCD}(2, 5) = 1$, and thus any L_1 matrix that generates this diagram is primitive.

In the present context, the transition diagram would contain a loop only if women in age class 1 have positive fertility levels. Thus, if the age partition is not too coarse, loops are not relevant for our present purposes. Nevertheless, for other (non-demographic) applications, it is useful to note that any loops would immediately indicate primitivity. More precisely, given the transition diagram for an irreducible matrix A , any loop implies that A is primitive. Loops constitute a cycle of length 1, and obviously $\text{GCD}(j, 1) = 1$ for any j .

9.5 Further reading

Bradley and Meek (*Matrices and Society*, 1986, Chapter 10) provide a simple introduction to matrix methods in demography. Allman and Rhodes (*Mathematical Models in Biology*, Chapter 2) and Mooney and Swift (*A Course in Mathematical Modeling*, 1999, Chapter 3) provide other helpful introductions which focus on non-human populations. For a more advanced treatment of matrix methods in demography, see Smith (*Formal Demography*, 1992, Chapter 8). Other useful references include Keyfitz (*Applied Mathematical Demography*, 1985) and Preston, Heuveline, and Guillot (*Demography: Measuring and Modeling Population Processes*, 2001). Farina and Rinaldi (*Positive Linear Systems*, Chapter 4) discuss alternative tests for primitivity.

⁷Recalling our analysis of reachability in Chapter 7, we could also find cycles using matrix methods: cycles of length t are indicated by positive elements on the main diagonal of the Z^t matrix.