special vectors, however, the matrix does not rotate the vector, but only stretches or shrinks it. These special vectors are the eigenvectors of **M**. Now, imagine what would happen if we used the right eigenvectors as the axes along which we tracked the dynamics of a model, instead of the original axes (e.g., instead of axes describing the population size of each species). In this new coordinate system, the dynamics would appear simple: any point is just stretched or shrunk along the ith axis by a factor λ_i in each generation (Figure 9.1). But how do we accomplish this change of coordinate system? The answer involves the matrices **A** and $\mathbf{A}^{-1}$.

We can think of the matrices **A** and $\mathbf{A}^{-1}$ as *transformation* matrices. In general, transformation matrices can be used to change from one set of coordinate axes to another. Specifically, the matrix $\mathbf{A}^{-1}$ transforms a point $\vec{n}(t)$ in the original coordinate system to a point $\vec{y}(t)$ in the new coordinate system defined by the right eigenvectors; that is, $\vec{y}(t) = \mathbf{A}^{-1}\,\vec{n}(t)$. Similarly, the reverse transformation from the new to the original coordinate system can be accomplished by multiplying by **A**; that is, $\vec{n}(t) = \mathbf{A}\vec{y}(t)$. We now apply these transformations to a system of recursion equations.

Box 9.1: Long-Term Dynamics and the Role of the Leading Eigenvalue

In this box, we describe an approximation to the dynamics of a linear discrete-time model in multiple variables. Consider a transition matrix **M** whose eigenvalues are known and placed along the diagonal of a matrix **D**, and whose eigenvectors are known and placed in the columns of a matrix **A**. We are free to place the eigenvalues in **D** in any order (as long as we place their associated eigenvectors in the same order in **A**), so let us place the leading eigenvalue (the eigenvalue with the largest magnitude, λ_1) in the first row and first column of **D**. To simplify the presentation further, we adjust the length of the eigenvector associated with the leading eigenvalue $\vec{u}_1$, so that its elements sum to one. As discussed in Primer 2, eigenvectors point in a particular direction but can be of any length, and therefore we are free to choose whatever length is convenient.

We start by factoring out λ_1^t from $\mathbf{D}^t$:

$$\mathbf{D}^t = \lambda_1^t \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \left(\frac{\lambda_2}{\lambda_1}\right)^t & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \left(\frac{\lambda_d}{\lambda_1}\right)^t \end{pmatrix}. \qquad (9.1.1)$$

As long as λ_1 is larger in magnitude than all other eigenvalues, (λ_i/λ_1) will be less than one in magnitude and $(\lambda_i/\lambda_1)^t$ will approach zero over time. Consequently, over time, $\mathbf{D}^t$ becomes more and more similar to

$$\widetilde{\mathbf{D}}^t = \lambda_1^t \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}. \tag{9.1.2}$$

Whether or not $\widetilde{\mathbf{D}}^t$ provides a sufficiently accurate approximation depends on how much time has passed and on the magnitude of the other eigenvalues relative to the leading eigenvalue. If, for example, the eigenvalues of a 4 × 4 matrix are 3/2, −3/2, 1/3, and −1/2, two eigenvalues are equally large in magnitude ($\lambda = 3/2$ and −3/2), and the approximation (9.1.2) should be avoided in favor of the exact general solution (9.12) (see Appendix 3 for a description of which matrices can and cannot have more than one eigenvalue equal to the leading eigenvalue). On the other hand, if the eigenvalues are 5/2, −3/2, 1/3, and −1/2, then after only five time steps (9.1.1) becomes

$$\mathbf{D}^3 = \left(\frac{5}{2}\right)^5 \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \left(-\frac{3}{5}\right)^5 & 0 & 0 \\ 0 & 0 & \left(\frac{2}{15}\right)^5 & 0 \\ 0 & 0 & 0 & \left(-\frac{1}{5}\right)^5 \end{pmatrix} = \left(\frac{5}{2}\right)^5 \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -0.078 & 0 & 0 \\ 0 & 0 & 0.000042 & 0 \\ 0 & 0 & 0 & 0.00032 \end{pmatrix},$$

which is very close to (9.1.2). In this case, it would be reasonable to approximate $\mathbf{D}^t$ with $\widetilde{\mathbf{D}}^t$ unless you needed short-term or extremely accurate predictions.

Using (9.1.2), we can approximate the general solution (9.12) by

$$\vec{n}(t) \approx \mathbf{A}\,\widetilde{\mathbf{D}}^t\,\mathbf{A}^{-1}\,\vec{n}(0). \tag{9.1.3}$$

The first two terms in this equation can be multiplied together to give

$$\mathbf{A}\,\widetilde{\mathbf{D}}^t = \lambda_1^t \begin{pmatrix} u_{11} & 0 & \cdots & 0 \\ u_{21} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ u_{d1} & 0 & \cdots & 0 \end{pmatrix}. \tag{9.1.4}$$

Let $\widetilde{v}_{ij}$ stand for the element in the i_{th} row and j_{th} column of $\mathbf{A}^{-1}$. Multiplying $\mathbf{A}\,\widetilde{\mathbf{D}}^t$ by $\mathbf{A}^{-1}$ on the right then gives

$$\mathbf{A}\,\widetilde{\mathbf{D}}^t\,\mathbf{A}^{-1} = \lambda_1^t \begin{pmatrix} u_{11}\widetilde{v}_{11} & u_{11}\widetilde{v}_{12} & \cdots & u_{11}\widetilde{v}_{1d} \\ u_{21}\widetilde{v}_{11} & u_{21}\widetilde{v}_{12} & \cdots & u_{21}\widetilde{v}_{1d} \\ \vdots & \vdots & \vdots & \vdots \\ u_{d1}\widetilde{v}_{11} & u_{d1}\widetilde{v}_{12} & \cdots & u_{d1}\widetilde{v}_{1d} \end{pmatrix}. \qquad (9.1.5)$$

Finally, we can multiply this matrix by the vector describing the initial state of the population, $\vec{n}(0)$, to get an approximation for $\vec{n}(t)$:

$$\vec{n}(t) \approx \mathbf{A}\,\widetilde{\mathbf{D}}^t\,\mathbf{A}^{-1}\,\vec{n}(0) = \lambda_1^t \begin{pmatrix} u_{11}(\widetilde{v}_{11}\,n_1(0) + \widetilde{v}_{12}\,n_2(0) + \cdots \widetilde{v}_{1d}\,n_d(0)) \\ u_{21}(\widetilde{v}_{11}\,n_1(0) + \widetilde{v}_{12}\,n_2(0) + \cdots \widetilde{v}_{1d}\,n_d(0)) \\ \vdots \\ u_{d1}(\widetilde{v}_{11}\,n_1(0) + \widetilde{v}_{12}\,n_2(0) + \cdots \widetilde{v}_{1d}\,n_d(0)) \end{pmatrix}. \qquad (9.1.6)$$

This approximation can be rewritten as

$$\vec{n}(t) \approx \lambda_1^t\, c\, \vec{u}_1, \qquad (9.1.7)$$

where $c = \widetilde{v}_{11}\,n_1(0) + \widetilde{v}_{12}\,n_2(0) + \ldots + \widetilde{v}_{1d}\,n_d(0) = \vec{\widetilde{v}}_1\,\vec{n}(0)$. The constant c can be thought of as the initial size of the system, adjusted by the left eigenvector $\vec{\widetilde{v}}$ (which is found in the first row of $\mathbf{A}^{-1}$; see Chapter 10).

The approximation (9.1.7) indicates that the system will eventually grow at a rate equal to λ_1^t. Furthermore, because we have adjusted $\vec{u}_1$ so that its elements sum to one, the proportion of the system that is of type i is given by the ith element of $\vec{u}_1$ (i.e., by the ith element of the right eigenvector of $\mathbf{M}$ associated with the eigenvalue λ_1). This result helps us to understand why the leading eigenvalue determines the stability of an equilibrium: eventually it will come to dominate the recursions. It also forms the basis for important results in demography, as described in Chapter 10.

Box 9.2: General Solution of a Discrete-Time Linear Model with Complex Eigenvalues

In equation (9.12), we showed that the general solution for a discrete-time linear model can be written as $\vec{n}(t) = \mathbf{A}\,\mathbf{D}^t\,\mathbf{A}^{-1}\,\vec{n}(0)$. How do we interpret this solution if the eigenvalues are complex? Here, we show how the solution for a two-variable model can be written in terms of sine and cosine functions containing only real terms. To keep things general, we will work with the two-dimensional matrix $\mathbf{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. The eigenvalues and associated eigenvectors of this matrix are

$$\lambda_1 = \frac{a + d + \sqrt{(a-d)^2 + 4bc}}{2} \text{ with } \mathbf{u} = \left(1, \frac{-a + d + \sqrt{(a-d)^2 + 4bc}}{2b}\right), \quad (9.2.1a)$$

$$\lambda_2 = \frac{a + d - \sqrt{(a-d)^2 + 4bc}}{2} \text{ with } \mathbf{u} = \left(1, \frac{-a + d - \sqrt{(a-d)^2 + 4bc}}{2b}\right). \quad (9.2.1b)$$

If $(a - d)^2 + 4bc$ is positive, then the eigenvalues and associated eigenvectors are real, and (9.12) can be multiplied out to describe the system at any future point in time without any difficulties. If $(a - d)^2 + 4bc$ is negative, however, then the eigenvalues and eigenvectors are complex numbers involving $i = \sqrt{-1}$. How can we interpret these complex numbers? And how do we raise a complex eigenvalue to the tth power to evaluate $\mathbf{D}^t$?

Insights from geometry help. When an eigenvalue is complex, we can write it as

$$\lambda_1 = \alpha + \beta i$$

where

$$\alpha = \frac{a + d}{2} \text{ and } \beta = \frac{\sqrt{-(a-d)^2 - 4bc}}{2}.$$

By definition, α and β are real numbers, and βi is the imaginary part of the eigenvalue. Using Figure 9.2.1 as a guide, we can also write $\alpha + \beta i$ in terms of sine and cosine functions. This figure illustrates that

$$\alpha + \beta i = R(\cos(\theta) + i \sin(\theta)),$$

where the magnitude (i.e., absolute value) of the complex number is $R = \sqrt{\alpha^2 + \beta^2}$, which simplifies to $\sqrt{a\,d - b\,c}$, and θ is the angle between the number plotted on the complex plane and the horizontal axis;

$$\theta = \arctan\left(\frac{\beta}{\alpha}\right) = \arctan\left(\frac{\sqrt{-(a-d)^2 - 4bc}}{a + d}\right).$$

We have already encountered $(\cos\theta + i \sin\theta)$ in Euler's formula (Box 7.4). Euler showed that $(\cos\theta + i \sin\theta) = e^{i\theta}$. Consequently, we can rewrite $\alpha + \beta i$ as $R\, e^{i\theta}$. This is helpful because we can easily raise $R\, e^{i\theta}$ to the tth power, to get $\lambda^t = R^t e^{i\theta t}$. We can then apply Euler's formula again to write λ_1^t as $R^t(\cos(\theta t) + i$

sin(θt)). We can repeat this procedure for the second eigenvalue as well, the only difference being that the sign of β (and thus the sign of θ) changes.

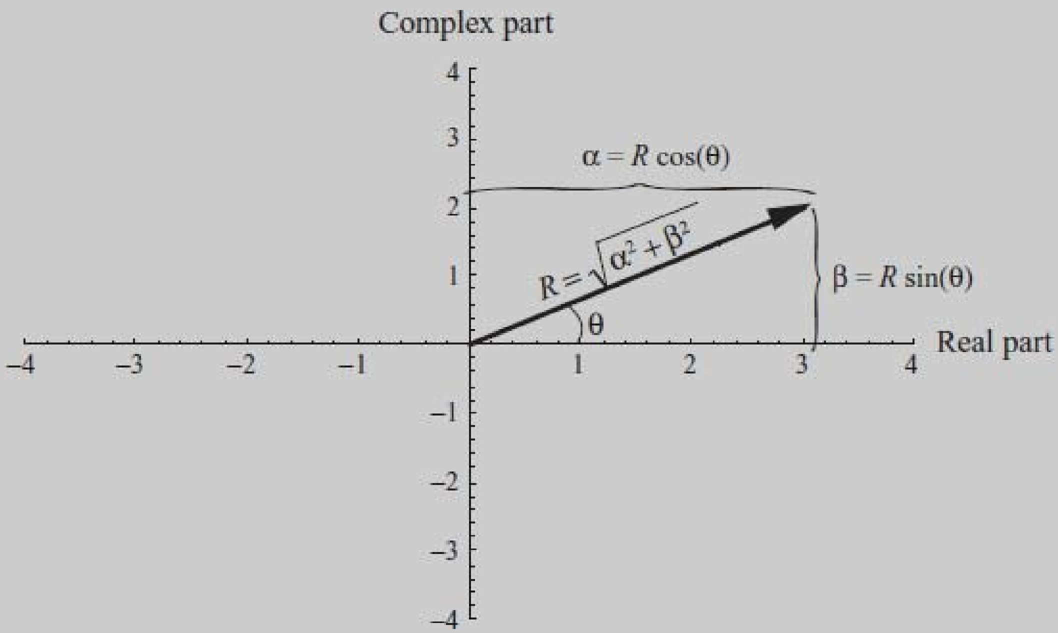

Figure 9.2.1: A complex number represented on the real-complex plane. Any complex number can be written as $\alpha + \beta\, i$ where both α and β are real numbers and $i = \sqrt{-1}$. Any number can thus be represented as a vector on a plot where the real part of the number (α) gives the position along the horizontal axis and the part multiplying $i(\beta)$ gives the position along the vertical axis. For example, we illustrate the number 3 + 2 i. From trigonometry, α must equal $R\cos(\theta)$, where θ is the angle between the vector and the horizontal axis and R is the length of the vector. Similarly, β must equal $R\sin(\theta)$. Thus, we can write $\alpha + \beta\, i$ in terms of sines and cosines as $R\cos(\theta) + R\sin(\theta)i$. The angle, θ, can be found using the trigonometric relationship, $\theta = \arctan(\beta/\alpha)$. The total length of the vector (its "magnitude") can be found from the theorem of Pythagoras, $R = \sqrt{\alpha^2 + \beta^2}$. For 3 + 2$i$, $\theta = \arctan(2/3) = 33.7°$ and $R = \sqrt{2^2 + 3^2} = 3.6$.

This procedure allows us to raise a complex number to the tth power, but the eigenvalue still involves a complex number. The beauty of using Euler's transformation is that when we multiply out $\mathbf{A}\,\mathbf{D}^t\,\mathbf{A}^{-1}$ and simplify the answer, we get a real matrix

$$\begin{pmatrix} n_1(t) \\ n_2(t) \end{pmatrix} = R^t \begin{pmatrix} \cos(\theta t) + \frac{(a-d)}{2\beta}\sin(\theta t) & \frac{b}{\beta}\sin(\theta t) \\ \frac{c}{\beta}\sin(\theta t) & \cos(\theta t) - \frac{(a-d)}{2\beta}\sin(\theta t) \end{pmatrix} \begin{pmatrix} n_1(0) \\ n_2(0) \end{pmatrix}, \tag{9.2.2}$$

where we have factored out $R^t = (a\,d - b\,c)^{t/2}$, which was present in every term in the matrix. This general solution tells us that the system cycles, with the matrix in (9.2.2) returning to the same value after a period of $\tau = 2\pi/\theta$. At the same time, the system expands or shrinks by a factor R every time step.