CHAPTER 6

EIGENVALUES AND EIGENVECTORS

The first major problem of linear algebra is to understand how to solve the basic linear system $Ax = b$ and what the solution means. We have explored this system from three points of view: In Chapter 1 we approached the problem from an operational point of view and learned the mechanics of computing solutions. In Chapter 2, we took a more sophisticated look at the system from the perspective of matrix theory. Finally, in Chapter 3, we viewed the problem from the vantage of vector space theory. Now we begin a study of the second major problem of linear algebra, namely the eigenvalue problem. We had to tackle linear systems first because the eigenvalue problem is more sophisticated and will require most of the tools that we have thus far developed. This subject has many important applications in areas such as computer graphics, mechanical vibrations, heat flow, population dynamics, quantum mechanics, and economics.

6.1. Eigenvalues and Eigenvectors

In this section we will define the notions of eigenvalue and eigenvector, and discuss some of their basic properties.

- **Eigenvalue and eigenvector**

  *Let A be a square $n \times n$ matrix. An **eigenvector** of A is a nonzero vector $\mathbf{v}$ in $\mathbb{R}^n$ such that for some scalar λ, we have*

  $$A\mathbf{v} = \lambda\mathbf{v}.$$

  *The scalar λ is called an **eigenvalue** of the matrix A, and we say that the vector $\mathbf{v}$ is an eigenvector belonging to the eigenvalue λ. The pair $\{\lambda, \mathbf{v}\}$ is called an **eigenpair** for the matrix A.*

We note that the eigenvalue λ is allowed to be the 0 scalar, but an eigenvector $\mathbf{v}$ is, by definition, never the zero vector $\mathbf{0}$.

Example: Let $A = \begin{bmatrix} 7 & 4 \\ 3 & 6 \end{bmatrix}$, $\mathbf{u} = (-1, 1)$, and $\mathbf{v} = (4, 3)$. It is easy to check that $A\mathbf{u} = (-3, 3) = 3\mathbf{u}$ and $A\mathbf{v} = (40, 30) = 10\mathbf{v}$. It follows that $\mathbf{u}$ and $\mathbf{v}$ are eigenvectors corresponding to eigenvalues 3 and 10, respectively.$\blacksquare$

Knowledge of eigenvectors and eigenvalues gives us deep insights into the structure of the matrix A. For example, we would like to have a better understanding of the effect of multiplication of a vector $\mathbf{v}$ by powers of the matrix A, that is, of $A^k\mathbf{v}$. If we knew that $\mathbf{v}$ were an eigenvector of A, then we would have that for some scalar λ,

$$\begin{aligned} A\mathbf{v} &= \lambda\mathbf{v} \\ A^2\mathbf{v} &= A(A\mathbf{v}) = A\lambda\mathbf{v} = \lambda A\mathbf{v} = \lambda^2\mathbf{v} \\ &\vdots \\ A^k\mathbf{v} &= A\left(A^{k-1}\mathbf{v}\right) = \cdots = \lambda^k\mathbf{v}. \end{aligned}$$

This reduces something complicated, namely matrix-vector multiplication, to something simple, namely scalar-vector multiplication.

These quantities are very useful. We consider some of the basic points about eigenvalues and eigenvectors.

- **Finding eigenvalues**

  For an $n \times n$ square matrix A the eigenvalues of A are all the scalars λ that are solutions to the nth degree polynomial equation $\det(\lambda I - A) = 0$.

Proof: We note that $\lambda\mathbf{v} = \lambda I\mathbf{v}$. A has eigenvalue λ if and only if $A\mathbf{v} = \lambda\mathbf{v}$, for some nonzero vector $\mathbf{v}$, which is true if and only if

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

for some nonzero vector $\mathbf{v}$. The matrix $\lambda I - A$ is square, so it has a nontrivial null space precisely when it is singular. This occurs only when $\det(\lambda I - A) = 0$. If we expand this determinant down the first column, we see that the highest-order term involving λ that occurs is the product of the diagonal terms $(\lambda - a_{ii})$, so that the degree of the expression $\det(\lambda I - A)$ as a polynomial in λ is n. This proves the result.■

We call a polynomial **monic** if the leading coefficient is 1. For example, $\lambda^2 + 2\lambda + 3$ is a monic polynomial in λ while $2\lambda^2 + \lambda + 1$ is not.

- **Characteristic polynomial**

  Given a square $n \times n$ matrix A, the equation $\det(\lambda I - A) = 0$ *is called the* ***characteristic equation*** *of A, and the nth-degree monic polynomial $p(\lambda) =$* $\det(\lambda I - A)$ *is called the* ***characteristic polynomial*** *of A.*

Example: Find all eigenvalues of the matrix

$$A = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}.$$

Solution: It follows that the eigenvalues of A are the solutions of the equation $\det(\lambda I - A) = 0$, which we can write as

$$\begin{vmatrix} \lambda - 3 & 0 \\ -8 & \lambda + 1 \end{vmatrix} = 0$$

from which we obtain $(\lambda - 3)(\lambda + 1) = 0$. This shows that the eigenvalues of A are $\lambda = 3$ and $\lambda = -1$.■

Example: Find the eigenvalues of

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 4 & -17 & 8 \end{bmatrix}.$$

Solution: The characteristic polynomial of A is

$$\det(\lambda I - A) = \det \begin{bmatrix} \lambda & -1 & 0 \\ 0 & \lambda & -1 \\ -4 & 17 & \lambda - 8 \end{bmatrix} = \lambda^3 - 8\lambda^2 + 17\lambda - 4.$$

The eigenvalues of A must therefore satisfy the cubic equation $\lambda^3 - 8\lambda^2 + 17\lambda - 4 = 0$. Since

$$\lambda^3 - 8\lambda^2 + 17\lambda - 4 = (\lambda - 4)\left(\lambda^2 - 4\lambda + 1\right) = (\lambda - 4)\left[\lambda - \left(2 + \sqrt{3}\right)\right]\left[\lambda - \left(2 - \sqrt{3}\right)\right] = 0,$$

the eigenvalues are $\lambda = 4$, $\lambda = 2 + \sqrt{3}$, and $\lambda = 2 - \sqrt{3}$.■

Example: Find the eigenvalues of the upper triangular matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{bmatrix}.$$

Solution: We obtain

$$\det(\lambda I - A) = \det \begin{bmatrix} \lambda - a_{11} & -a_{12} & -a_{13} & -a_{14} \\ 0 & \lambda - a_{22} & -a_{23} & -a_{24} \\ 0 & 0 & \lambda - a_{33} & -a_{34} \\ 0 & 0 & 0 & \lambda - a_{44} \end{bmatrix} = (\lambda - a_{11})(\lambda - a_{22})(\lambda - a_{33})(\lambda - a_{44}).$$

Thus the characteristic equation is $(\lambda - a_{11})(\lambda - a_{22})(\lambda - a_{33})(\lambda - a_{44}) = 0$, and the eigenvalues are $\lambda = a_{11}$, $\lambda = a_{22}$, $\lambda = a_{33}$, and $\lambda = a_{44}$, which are precisely the diagonal entries of A.■

This result is a special case of the following general result.

- **Eigenvalues of a triangular matrix**

  If A is an $n \times n$ triangular matrix, then the eigenvalues of A are the entries on the main diagonal of A.

The following result gives some alternative ways of describing eigenvalues.

- **Finding eigenvalues**

  If A is an $n \times n$ matrix, the following statements are equivalent.

  (a) *λ is an eigenvalue of A.*

  (b) *λ is a solution of the characteristic equation* $\det(\lambda I - A) = 0$.

  (c) *The system of equations $(\lambda I - A)x = \mathbf{0}$ has nontrivial solutions.*

  (d) *There is a nonzero vector $\mathbf{v}$ such that $A\mathbf{v} = \lambda\mathbf{v}$.*

Given an eigenvalue λ of the matrix A, the **eigenspace** corresponding to λ is the null subspace $N(\lambda I - A)$ of $\mathbb{R}^n$. By an **eigensystem** of the matrix A, we mean a list of all the eigenvalues of A and, for each eigenvalue λ, a complete description of the eigenspace corresponding to λ.

The usual way to give a complete description of an eigenspace is to list a basis for the space. Remember that there is one vector in the eigenspace $N(\lambda I - A)$ that is not an eigenvector, namely $\mathbf{0}$. In any case, the computational route is now clear.

- **Eigensystem algorithm**

  Let A be an $n \times n$ matrix. To find an eigensystem of A:

  (1) *Find the scalars that are roots to the characteristic equation* $\det(\lambda I - A) = 0$.

  (2) *For each scalar λ in* (1), *use the null space algorithm to find a basis of the eigenspace $N(\lambda I - A)$.*

Example: Find an eigensystem for the matrix

$$A = \begin{bmatrix} 7 & 4 \\ 3 & 6 \end{bmatrix}.$$

Solution: First solve the characteristic equation

$$0 = \det(\lambda I - A) = \lambda^2 - 13\lambda + 30 = (\lambda - 3)(\lambda - 10).$$

Hence the eigenvalues are $\lambda = 3$ and $\lambda = 10$. Next, for each eigenvector calculate the corresponding eigenspace. If $\lambda = 3$, then

$$3I - A = \begin{bmatrix} 3-7 & -4 \\ -3 & 3-6 \end{bmatrix} = \begin{bmatrix} -4 & -4 \\ -3 & -3 \end{bmatrix}$$

and row reduction gives

$$\begin{bmatrix} -4 & -4 \\ -3 & -3 \end{bmatrix} \xrightarrow{\left(-\frac{1}{4}\right)I} \begin{bmatrix} 1 & 1 \\ -3 & -3 \end{bmatrix} \xrightarrow{II+3I} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

so the general solution is

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Therefore, a basis for the eigenspace for $\lambda = 3$ is $\{(-1, 1)\}$. If $\lambda = 10$, then

$$10I - A = \begin{bmatrix} 10-7 & -4 \\ -3 & 10-6 \end{bmatrix} = \begin{bmatrix} 3 & -4 \\ -3 & 4 \end{bmatrix}$$

and row reduction gives

$$\begin{bmatrix} 3 & -4 \\ -3 & 4 \end{bmatrix} \xrightarrow{II+I} \begin{bmatrix} 3 & -4 \\ 0 & 0 \end{bmatrix} \xrightarrow{\frac{1}{3}I} \begin{bmatrix} 1 & -\frac{4}{3} \\ 0 & 0 \end{bmatrix}$$

so the general solution is

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \frac{4}{3}x_2 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} \frac{4}{3} \\ 1 \end{bmatrix}.$$

Therefore, a basis for the eigenspace for $\lambda = 10$ is $\left\{\left(\frac{4}{3}, 1\right)\right\}$.■

The next result establishes a relationship between the eigenvalues and the invertibility of a matrix.

- **Eigenvalues and invertibility**

  A square matrix A is invertible if and only if $\lambda = 0$ is not an eigenvalue of A.

Proof: Assume that A is an $n \times n$ matrix and observe first that $\lambda = 0$ is a solution of the characteristic equation

$$\lambda^n + a_1\lambda^{n-1} + \cdots + a_n = 0$$

if and only if the constant term $a_n = 0$. Thus, it suffices to prove that A is invertible if and only if $a_n \neq 0$. Since

$$\det(\lambda I - A) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_n,$$

letting $\lambda = 0$, we find that $\det(-A) = a_n$, or $(-1)^n \det(A) = a_n$. It then follows that $\det(A) = 0$ if and only if $a_n = 0$, and hence A is invertible if and only if $a_n \neq 0$.■

As a result in this section, we can add one more equivalent condition to the list of equivalences for invertibility.

- **Equivalent statements VI**

  If A is an $n \times n$ matrix, then the following statements are equivalent:

  (1) *A is invertible.*

  (2) *$Ax = \mathbf{0}$ has only the trivial solution.*

  (3) *The reduced row echelon form of A is I_n.*

  (4) *A is expressible as a product of elementary matrices.*

  (5) *$Ax = b$ is consistent for every $n \times 1$ matrix b.*

  (6) *$Ax = b$ has exactly one solution for every $n \times 1$ matrix b.*

  (7) $\det(A) \neq 0$.

  (8) $N(A) = \{\mathbf{0}\}$.

  (9) The kernel of T_A is $\{\mathbf{0}\}$.

  (10) *The range of T_A is $\mathbb{R}^n$.*

  (11) *T_A is one-to-one.*

  (12) *$\lambda = 0$ is not an eigenvalue of A.*

The following definition extends the notion of eigenvalues and eigenvectors to general vector spaces.

- **Eigenvalues of general linear transformations**

  If $T : V \to V$ is a linear operator on a vector space V, then a nonzero vector $\mathbf{v}$ in V is called an ***eigenvector*** *of T if*

  $$T(\mathbf{v}) = \lambda \mathbf{v}$$

  for some scalar λ. The scalar λ is called an ***eigenvalue*** *of T, and $\mathbf{v}$ is said to be an* ***eigenvector corresponding to*** *λ.*

Example: If $D : C^\infty(-\infty, \infty) \to C^\infty(-\infty, \infty)$ is the derivative operator on the vector space of functions with continuous derivatives of all orders on the interval $(-\infty, \infty)$, and if λ is a constant, then $D\left(\mathrm{e}^{\lambda x}\right) = \lambda \mathrm{e}^{\lambda x}$ so that λ is an eigenvalue of D and $\mathrm{e}^{\lambda x}$ is a corresponding eigenvector.■

6.2. Diagonalization

In this section we will be concerned with the problem of finding a basis for $\mathbb{R}^n$ that consists of eigenvectors of an $n \times n$ matrix A. Such bases can be used to study geometric properties of A and to simplify various numerical computations. These bases are also of physical significance in a wide variety of applications.

We start by showing the importance of eigenvalues, namely we first consider a discrete linear dynamical system.

Example: By some unfortunate accident a new species of frog has been introduced into an area where it has too few natural predators. In an attempt to restore the ecological balance, a team of scientists is considering introducing a species of bird that feeds on this frog. Experimental data suggests that the population of frogs and birds from one year to the next can be modeled by the linear relationships

$$\begin{aligned} B_{k+1} &= 0.6B_k + 0.4F_k \\ F_{k+1} &= -0.35B_k + 1.4F_k \end{aligned}$$

where the quantities F_k and B_k represent the populations of the frogs and birds in the kth year. This system models their joint behavior reasonably well. To determine that the introduction of the birds reduce or eliminate growth of the frog population, we let the population vector in the kth year be $\mathbf{x}^{(k)} = (B_k, F_k)$. Then the linear relationship above becomes

$$\begin{bmatrix} B_{k+1} \\ F_{k+1} \end{bmatrix} = \begin{bmatrix} 0.6 & 0.4 \\ -0.35 & 1.4 \end{bmatrix} \begin{bmatrix} B_k \\ F_k \end{bmatrix},$$

which is a discrete linear dynamical system. Notice that this is different from the Markov chains we studied earlier, since one of the entries of the coefficient matrix is negative. Letting $A = \begin{bmatrix} 0.6 & 0.4 \\ -0.35 & 1.4 \end{bmatrix}$, we observe that $\mathbf{x}^{(k)} = A^k \mathbf{x}^{(0)}$. So what we really need to know is how the powers of the transition matrix A behave. In general, this is very hard.■

But if the matrix A in the above example is a diagonal matrix, then things are relatively simple.

- **Powers of diagonal matrices**

  If D is an $n \times n$ diagonal matrix, then for a positive integer k, D^k is the diagonal matrix with entries kth powers of the entries of D.

Proof: Let

$$D = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}.$$

Then

$$D^2 = DD = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} a_{11}^2 & 0 & \cdots & 0 \\ 0 & a_{22}^2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn}^2 \end{bmatrix}.$$

Iterating we find that

$$D^k = \begin{bmatrix} a_{11}^k & 0 & \cdots & 0 \\ 0 & a_{22}^k & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn}^k \end{bmatrix}$$

for a positive integer k.■

This theorem has a very interesting consequence.

- **Polynomial of diagonal matrices**

  If D is an $n \times n$ diagonal matrix and $f(x)$ is a polynomial, then $f(D)$ is the diagonal matrix with entries that are values of the polynomial $f(x)$ at the entries of D.

Proof: Observe that if $f(x) = a_0 + a_1 x + \cdots + a_n x^n$, then $f(D) = a_0 I + a_1 D + \cdots + a_n D^n$. Applying the preceding result to each monomial D^k and adding up the resulting terms in $f(D)$, we have the result.■

Now for the powers of a more general matrix A, we consider a 3×3 matrix for simplicity. If $\lambda_1, \lambda_2, \lambda_3$ are eigenvalues of A and $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ are linearly independent eigenvectors, then $A\mathbf{v}_1 = \lambda_1 \mathbf{v}_1$, $A\mathbf{v}_2 = \lambda_2 \mathbf{v}_2$, $A\mathbf{v}_3 = \lambda_3 \mathbf{v}_3$. In matrix form, we write

$$A \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}.$$

Now we set $P = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix}$ and $D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$. Then P is invertible since the columns of P are linearly independent. So the equation $AP = PD$, if multiplied on the left by P^{-1}, gives the equation

$$P^{-1}AP = D.$$

This equation makes the powers of A simple to understand. The procedure we just went through is reversible as well. In other words, if P is an invertible matrix such that $P^{-1}AP = D$, then we deduce that $AP = PD$, identify the columns of P by the equation $P = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix}$, and conclude that the columns of P are linearly independent eigenvectors of A.

- **Similar matrices**

  *A matrix A is said to be **similar** to matrix B if there exists an invertible matrix P such that*

  $$P^{-1}AP = B.$$

  *The matrix P is called a **similarity transformation matrix**.*

It is easy to show that similar matrices have to be square and of the same size. Furthermore, if A is similar to B, then B is similar to A. To see this, suppose that $P^{-1}AP = B$ and multiply by P on the left and P^{-1} on the right to obtain that

$$A = PP^{-1}APP^{-1} = PBP^{-1} = \left(P^{-1}\right)^{-1} BP^{-1}.$$

Similar matrices have much in common. For example, suppose that $B = P^{-1}AP$ and λ is an eigenvalue of A, say $A\mathbf{v} = \lambda\mathbf{v}$. Since

$$A\mathbf{v} = \lambda P^{-1}\mathbf{v} = P^{-1}A\mathbf{v} = P^{-1}AP\left(P^{-1}\mathbf{v}\right),$$

we see that λ is an eigenvalue of B. In the following we give a slightly stronger statement.

- **Similar matrices and characteristic polynomials**

  Suppose that A is similar to B. Then the matrices A and B have the same characteristic polynomial, and hence the same eigenvalues.

Proof: Since the determinant distributes over products, we have

$$\begin{aligned} \det(\lambda I - B) &= \det\left(\lambda P^{-1}IP - P^{-1}AP\right) = \det\left(P^{-1}(\lambda I - A)P\right) \\ &= \det\left(P^{-1}\right)\det(\lambda I - A)\det(P) = \det(\lambda I - A)\det\left(P^{-1}P\right) \\ &= \det(\lambda I - A). \end{aligned}$$

This proves the result.■

Now we can see the significance of the equation $P^{-1}AP = D$, where D is diagonal. It follows from this equation that for any positive integer k, we have $P^{-1}A^kP = D^k$, so multiplying on the left by P and on the right by P^{-1} yields

$$A^k = PD^kP^{-1}.$$

Since the term PD^kP^{-1} is easily computed, this gives us a way of constructing a formula for A^k.

Example: Let $A = \begin{bmatrix} 2 & 1 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & 2 \end{bmatrix}$. Since

$$\det(\lambda I - A) = \begin{vmatrix} \lambda - 2 & -1 & -2 \\ 0 & \lambda - 1 & 2 \\ 0 & 0 & \lambda - 2 \end{vmatrix} = (\lambda - 1)(\lambda - 2)^2 = 0,$$

$\lambda = 1, 2, 2$ are eigenvalues. For $\lambda = 2$, we have

$$\begin{bmatrix} 0 & -1 & -2 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{I+II} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{I \leftrightarrow II} \begin{bmatrix} 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

so that the free variables are x_1, x_3 and the general solution is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ -2x_3 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix}.$$

Thus a basis for the corresponding eigenspace is $\{(1,0,0), (0,-2,1)\}$. On the other hand, for $\lambda = 1$, we have

$$\begin{bmatrix} -1 & -1 & -2 \\ 0 & 0 & 2 \\ 0 & 0 & -1 \end{bmatrix} \xrightarrow{I+II} \begin{bmatrix} -1 & -1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & -1 \end{bmatrix} \xrightarrow[\frac{1}{2}II]{(-1)I} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \end{bmatrix} \xrightarrow{III+II} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix},$$

which gives a general eigenvector of the form

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_2 \\ 0 \end{bmatrix} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}.$$

Hence the corresponding eigenspace has the basis $\{(-1,1,0)\}$. Therefore, we set $\mathbf{v}_1 = (1,0,0)$, $\mathbf{v}_2 = (0,-2,1)$, $\mathbf{v}_3 = (-1,1,0)$, and form the matrix

$$P = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & -2 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

This matrix is invertible since $\det(P) = -1$, and we have

$$P^{-1} = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix}.$$

P is then a similarity transformation matrix that diagonalizes A, that is,

$$P^{-1}AP = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = D.$$

This means that for any positive integer k, we have

$$A^k = PD^kP^{-1} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & -2 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2^k & 0 & 0 \\ 0 & 2^k & 0 \\ 0 & 0 & 2^k \end{bmatrix} \begin{bmatrix} 1 & 1 & 2 \\ 0 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 2^k & 2^k - 1 & 2^{k+1} - 2 \\ 0 & 1 & -2^{k+1} + 2 \\ 0 & 0 & 2^k \end{bmatrix},$$

which is much more easier than calculating A^k directly.■

This example shows some very nice calculations. Given a general matrix we can apply same sort of calculation.

- **Diagonalizable matrix**

  A matrix A is ***diagonalizable*** *if it is similar to a diagonal matrix, that is, there is an invertible matrix P and diagonal matrix D such that $P^{-1}AP = D$. In this case we say that P is a* ***diagonalizing matrix*** *for A or that P* ***diagonalizes*** *A.*

We can be more specific about when a matrix is diagonalizable. As a first step, we notice that the calculations in the above example can easily be written in terms of an $n \times n$ matrix instead of a 3×3 matrix. These calculations prove the following basic fact.

- **Diagonalization theorem**

  An $n \times n$ matrix A is diagonalizable if and only if there exists a linearly independent set of eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ of A, in which case $P = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix}$ is a diagonalizing matrix for A.

Example: Diagonalize the matrix

$$A = \begin{bmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 24 & -12 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

Solution: Since

$$\det(\lambda I - A) = \begin{vmatrix} \lambda+2 & 0 & 0 & 0 \\ 0 & \lambda+2 & 0 & 0 \\ -24 & 12 & \lambda-2 & 0 \\ 0 & 0 & 0 & \lambda-2 \end{vmatrix} = (\lambda+2)^2(\lambda-2)^2,$$

the eigenvalues are $\lambda = -2, -2, 2, 2$. For $\lambda = 2$, we have

$$\begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ -24 & 12 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4x_1 \\ 4x_2 \\ -24x_1+12x_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 4x_1 \\ 0 \\ -24x_1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 4x_2 \\ 12x_2 \\ 0 \end{bmatrix} = 4x_1\begin{bmatrix} 1 \\ 0 \\ -6 \\ 0 \end{bmatrix} + 4x_2\begin{bmatrix} 0 \\ 1 \\ 3 \\ 0 \end{bmatrix},$$

so a basis for the corresponding eigenspace is $\{(1,0,-6,0),(0,1,3,0)\}$. On the other hand, for $\lambda = -2$, we have

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -24 & 12 & -4 & 0 \\ 0 & 0 & 0 & -4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -4x_3 \\ -4x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -4x_3 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ -4x_4 \end{bmatrix} = -4x_3\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} - 4x_4\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

so the basis for the corresponding eigenspace is $\{(0,0,1,0),(0,0,0,1)\}$. Since $\mathbf{v}_1 = (1,0,-6,0), \mathbf{v}_2 = (0,1,3,0), \mathbf{v}_3 = (0,0,1,0), \mathbf{v}_4 = (0,0,0,1)$ are linearly independent, we obtain that

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -6 & 3 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Since

$$P^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 6 & -3 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

we find that

$$D = P^{-1}AP = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 6 & -3 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 24 & -12 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -6 & 3 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix},$$

which diagonalizes the matrix A.■

It would be very useful to have some working criterion for when we can manufacture linearly independent sets of eigenvectors. The next theorem gives us such a criterion.

- **Linear independence of eigenvectors**

  Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ be a set of eigenvectors of a matrix A such that corresponding eigenvalues are all distinct. Then the set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is linearly independent.

Proof: Suppose the set is linearly dependent. Discarding linear combinations, we deduce a smallest linearly dependent subset such as $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ with $\mathbf{v}_i$ belonging to λ_i. All the vectors have nonzero coefficients in a linear combination that sums to zero, for we could discard the ones that have zero coefficient in the linear combination and still have a linearly dependent set. So there is some linear combination of the form

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_m\mathbf{v}_m = \mathbf{0}$$

with each $c_j \neq 0$ and $\mathbf{v}_j$ belonging to the eigenvalue λ_j. We multiply this equation by λ_1 to obtain the equation

$$c_1\lambda_1\mathbf{v}_1 + c_2\lambda_1\mathbf{v}_2 + \cdots + c_m\lambda_1\mathbf{v}_m = \mathbf{0}.$$

Next we multiply last equation on the left by A to obtain

$$\mathbf{0} = A\left(c_1\lambda_1\mathbf{v}_1 + c_2\lambda_1\mathbf{v}_2 + \cdots + c_m\lambda_1\mathbf{v}_m\right) = c_1A\mathbf{v}_1 + c_2A\mathbf{v}_2 + \cdots + c_mA\mathbf{v}_m,$$

that is,

$$c_1\lambda_1\mathbf{v}_1 + c_2\lambda_2\mathbf{v}_2 + \cdots + c_k\lambda_m\mathbf{v}_m = \mathbf{0}.$$

The two equations for λ's give

$$0\mathbf{v}_1 + c_2\left(\lambda_1 - \lambda_2\right)\mathbf{v}_2 + \cdots + c_k\left(\lambda_1 - \lambda_m\right)\mathbf{v}_m = \mathbf{0}.$$

This is a new nontrivial linear combination (since $c_2\left(\lambda_1 - \lambda_2\right) \neq 0$) of fewer terms, that contradicts our choice of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$. It follows that the original set of vectors must be linearly independent.■

An application of the theorem is as in the following that is useful for many problems.

- **Eigenvalues and diagonalization**

  If an $n \times n$ matrix A has n distinct eigenvalues, then A is diagonalizable.

Proof: We can always find one nonzero eigenvector $\mathbf{v}_i$ for each eigenvalue λ_i of A. By the preceding result, the set $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ is linearly independent. Thus A is diagonalizable by the diagonalization theorem.■

Example: Show that the matrix

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ -3 & 5 & 2 \end{bmatrix}$$

is not diagonalizable.

Solution: Since

$$\det(\lambda I - A) = \begin{vmatrix} \lambda - 1 & 0 & 0 \\ -1 & \lambda - 2 & 0 \\ 3 & -5 & \lambda - 2 \end{vmatrix} = (\lambda - 1)(\lambda - 2)^2 = 0,$$

$\lambda = 1, 2, 2$ are eigenvalues, and the distinct eigenvalues are $\lambda = 1$ and $\lambda = 2$. For $\lambda = 1$, the base for the corresponding eigenspace is $\mathbf{v}_1 = \left(\frac{1}{8}, -\frac{1}{8}, 1\right)$ and for $\lambda = 2$, we have the base $\mathbf{v}_2 = (0, 0, 1)$. Since A is a 3×3 matrix and there are only two basis vectors in total, A is not diagonalizable.■

Example: Find A^7 if

$$A = \begin{bmatrix} 4 & 0 & -2 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

Solution: We have

$$\det(\lambda I - A) = \begin{vmatrix} \lambda - 4 & 0 & 2 \\ 0 & \lambda - 1 & 0 \\ -1 & 0 & \lambda - 1 \end{vmatrix} = (\lambda - 1)(\lambda - 2)(\lambda - 3),$$

so the eigenvalues are $\lambda = 1$, $\lambda = 2$, and $\lambda = 3$. For $\lambda = 1$, we have

$$I - A = \begin{bmatrix} -3 & 0 & 2 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix},$$

whose reduced-row echelon form is

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

This yields to an eigenvector of the form

$$\mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

For $\lambda = 2$, we have

$$2I - A = \begin{bmatrix} -2 & 0 & 2 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix},$$

whose reduced-row echelon form is

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

This yields to an eigenvector of the form

$$\mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

For $\lambda = 3$, we have

$$3I - A = \begin{bmatrix} -1 & 0 & 2 \\ 0 & 2 & 0 \\ -1 & 0 & 2 \end{bmatrix},$$

whose reduced-row echelon form is

$$\begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

This yields to an eigenvector of the form

$$\mathbf{v}_3 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}.$$

Thus, if

$$P = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix},$$

then $A = PDP^{-1}$, and hence $A^7 = PD^7P^{-1}$. Since

$$P^{-1} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 2 \\ 1 & 0 & -1 \end{bmatrix},$$

we find that

$$\begin{aligned} A^7 &= \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 128 & 0 \\ 0 & 0 & 2187 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 2 \\ 1 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 128 & 4374 \\ 1 & 0 & 0 \\ 0 & 128 & 2187 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 2 \\ 1 & 0 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 4246 & 0 & -4118 \\ 0 & 1 & 0 \\ 2059 & 0 & -1931 \end{bmatrix}. \blacksquare \end{aligned}$$

6.3. The Eigenvalues of a Linear Transformation

Recall that if $\mathbf{v}$ is an eigenvector of a matrix A, then $A\mathbf{v} = \lambda\mathbf{v}$ for some constant λ. Since $A\mathbf{v}$ can be regarded as an image of the matrix transformation T_A, we have the notion of eigenvalues and eigenvectors for matrix transformations. This concept can be generalized to any vector space. In the case of a finite dimensional vector space, we can generalize the idea of diagonalization as well.

- **Eigenvalues of a linear operator**

  Let V be a vector space and $T : V \to V$ be a linear operator. A scalar λ is called an ***eigenvalue of*** *T if there exists a nonzero element v of V such that*

  $$T(\mathbf{v}) = \lambda\mathbf{v}.$$

  Such an element v is called an ***eigenvector of*** *T*

If V consists of functions, then $\mathbf{v}$ in the above definition is called an **eigenfunction**, if V consists of matrices, then $\mathbf{v}$ is called an **eigenmatrix**, and so on.

Example: Consider the derivative operator D on $C^\infty(-\infty, \infty)$, the vector space of all functions defined on the entire real line which are continuously differentiable of any order. Since

$$D(\mathrm{e}^x) = (\mathrm{e}^x)' = \mathrm{e}^x = 1 \cdot \mathrm{e}^x,$$

we see that e^x is an eigenfunction of D associated with the eigenvalue 1. More generally,

$$D\left(\mathrm{e}^{kx}\right) = k\mathrm{e}^{kx}$$

shows that e^{kx} is an eigenfunction with associated eigenvalue k. Here, k can be any real number. So, any real number is an eigenvalue of D.■

- **Eigenbasis and diagonalization**

  *If V is a finite dimensional vector space, then a basis of V consisting of eigenvectors of the linear operator $T : V \to V$ is called an **eigenbasis for** T. We say that the operation T is **diagonalizable** if the matrix for T with respect to some eigenbasis is diagonal.*

Example: Consider the linear operator on M_{22} defined by

$$L(A) = A^T.$$

If A is symmetric, then

$$L(A) = A^T = A = 1 \cdot A,$$

so A is an eigenmatrix with eigenvalue 1. The symmetric 2×2 matrices $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$ form a three-dimensional space with basis

$$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}.$$

We need one more matrix to form an eigenbasis for L since M_{22} is four-dimensional.

If A is skew-symmetric, then

$$L(A) = A^T = -A = (-1)A,$$

so A is an eigenmatrix with eigenvalue -1. The skew-symmetric 2×2 matrices $\begin{bmatrix} 0 & a \\ -a & 0 \end{bmatrix}$ form a one-dimensional space with basis

$$\left\{ \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \right\}.$$

Thus, an eigenbasis of L is

$$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \right\},$$

so L is diagonalizable.■

Example: Consider the linear operator on P_2, the vector space of all polynomials of degree at most 2, defined by

$$T(p(x)) = p(2x - 1).$$

Show that the operator T is diagonal.

Solution: Let $p(x) = a + bx + cx^2$. Then

$$T(p(x)) = p(2x-1) = a + b(2x-1) + c(2x-1)^2 = a - b + c + (2b - 4c)x + 4cx^2.$$

The corresponding coefficient matrix is

$$\begin{bmatrix} a-b+c \\ 2b-4c \\ 4c \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & -4 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}.$$

Thus, we have

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & -4 \\ 0 & 0 & 4 \end{bmatrix}.$$

The eigenvalues are then $\lambda = 1$, $\lambda = 2$, and $\lambda = 4$, with corresponding eigenvectors

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix},$$

respectively. Transforming them into P_2, we find the eigenbasis for T as

$$B = \left\{1, x-1, (x-1)^2\right\}.$$

We have

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix} \quad \text{and} \quad S = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$$

as the diagonal and invertible matrices.■

6.4. Eigenvalues and Discrete Dynamical Systems

Recall that a discrete linear dynamical system is a sequence of vectors $\{\mathbf{x}_k\}$, which is defined by an initial vector $\mathbf{x}_0$ and by the rule $\mathbf{x}_{k+1} = A\mathbf{x}_k$, where A is called the transition matrix. Note that this equality implies $\mathbf{x}_k = A^k\mathbf{x}_0$. Eigenvalues and eigenvectors provide the key to understand the long term behavior, or evolution of a dynamical system.

Let A be an $n \times n$ matrix. If $\mathbf{v}$ is an eigenvector of A corresponding to the eigenvalue λ, then $A\mathbf{v} = \lambda\mathbf{v}$. Since

$$A^2\mathbf{v} = A(A\mathbf{v}) = A(\lambda\mathbf{v}) = \lambda(A\mathbf{v}) = \lambda(\lambda\mathbf{v}) = \lambda^2\mathbf{v},$$

we have

$$A^k\mathbf{v} = \lambda^k\mathbf{v}$$

by iteration. Thus, if $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an eigenbasis of Acorresponding to the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, respectively, and

$$\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$$

for some scalars $c_1, c_2, \ldots, c_n$,then

$$\mathbf{x}_k = c_1\lambda_1^k\mathbf{v}_1 + c_2\lambda_2^k\mathbf{v}_2 + \cdots + c_n\lambda_n^k\mathbf{v}_n$$

is a solution of $\mathbf{x}_{k+1} = A\mathbf{x}_k$.

Example: (A predator-prey system) Suppose there is a population of owls (the predators) living among a population of rats (the prey). A significant portion of owls' diet consists of rats. Let O_k and R_k be the number of owls and rats (in thousands) after k months, respectively. Suppose that

$$\begin{aligned} O_{k+1} &= (0.5)\,O_k + (0.4)\,R_k \\ R_{k+1} &= -\,(0.104)\,O_k + (1.1)\,R_k \end{aligned}$$

The $(0.5)\,R_k$ in the first equation states that with no rats for food, only 50% of owls will survive each month. $(1.1)\,R_k$ in the second equation states that with no owls as predators, the rat population will grow by 10% per month. If there are too many rats, then the $(0.4)\,R_k$ in the first equation will tend to make owl population rise. The negative term $-\,(0.104)\,O_k$ in the second equation measures the deaths of rats due to predation by owls (in fact, 104 is the average number of rats eaten by one owl in one month). Determine the evolution of this system.

Solution: Let $\mathbf{x}_k = \begin{bmatrix} O_k \\ R_k \end{bmatrix}$. The given system of equations gives

$$\mathbf{x}_{k+1} = \begin{bmatrix} O_{k+1} \\ R_{k+1} \end{bmatrix} = \begin{bmatrix} (0.5)\,O_k + (0.4)\,R_k \\ -\,(0.104)\,O_k + (1.1)\,R_k \end{bmatrix} = \begin{bmatrix} 0.5 & 0.4 \\ -0.104 & 1.1 \end{bmatrix} \begin{bmatrix} O_k \\ R_k \end{bmatrix} = A\mathbf{x}_k,$$

where $A = \begin{bmatrix} 0.5 & 0.4 \\ -0.104 & 1.1 \end{bmatrix}$. The eigenvalues of A are $\lambda = 1.02$ and $\lambda = 0.58$, and corresponding eigenvectors are

$$\mathbf{v}_1 = \begin{bmatrix} 10 \\ 13 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} 5 \\ 1 \end{bmatrix}.$$

An initial $\mathbf{x}_0$ can be written as $\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$. Then, for $k \geq 0$, we have

$$\mathbf{x}_k = c_1\,(1.02)^k\,\mathbf{v}_1 + c_2\,(0.58)^k\,\mathbf{v}_2.$$

As $k \to \infty$, $(0.58)^k$ approaches 0. If $c_1 > 0$, then for large values of k, we have

$$\mathbf{x}_k \approx c_1 (1.02)^k \begin{bmatrix} 10 \\ 13 \end{bmatrix} \Rightarrow \mathbf{x}_{k+1} \approx 1.02\mathbf{x}_k.$$

This approximation says that eventually both entries of $\mathbf{x}_k$ (the number of owls and rats) grow by a factor of almost 1.02 each month, a 2% monthly growth rate. $\mathbf{x}_k$ is approximately a multiple of $(10, 13)$, so the entries in $\mathbf{x}_k$ are nearly in the same ratio as 10 to 13. That is, for every 10 owls there are about 13 thousand rats.■

Example: (A population model) A population of insects consists of larvae, young adults, and older adults. Each year one percent of larvae grow into young adults and 30 percent of young adults survive to become older adults. Young adults have an average of 104 offspring each year, while older adults have an average of 160 offspring. Find the long-term evolution of the population.

Solution: Let L_t, Y_t, and O_t be the number of larvae, young adults, and older adults at year t, respectively. Then we have

$$\begin{aligned} L_{t+1} &= 104Y_t + 160O_t \\ Y_{t+1} &= 0.01L_t \\ O_{t+1} &= 0.3Y_t \end{aligned}$$

Letting $\mathbf{x}_k = \begin{bmatrix} L_t \\ Y_t \\ O_t \end{bmatrix}$, the above system can be written as $\mathbf{x}_{k+1} = A\mathbf{x}_k$, where

$$A = \begin{bmatrix} 0 & 104 & 160 \\ 0.01 & 0 & 0 \\ 0 & 0.3 & 0 \end{bmatrix}.$$

An eigenvalue of A is $\lambda = 1.2$ (other eigenvalues are not real). Thus,

$$\mathbf{x}_{k+1} \approx 1.2\mathbf{x}_k,$$

which implies that the number of larvae, young adults, and older adults grow by 20% yearly.■

Example: (Fibonacci numbers) Fibonacci numbers are denoted by F_n and defined recursively as

$$F_n = F_{n-1} + F_{n-2}$$

for $n \geq 2$ with the initial conditions $F_0 = 0$ and $F_1 = 1$. The sequence consisting of Fibonacci numbers is called the Fibonacci sequence:

$$\{0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, \ldots\}.$$

A formula for the nth Fibonacci number depending only the nonnegative integer n is due to Binet, and hence known as Binet's formula:

$$F_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right].$$

To prove Binet's formula we let $\mathbf{x}_k = \begin{bmatrix} F_k \\ F_{k+1} \end{bmatrix}$. Then by the recurrence of the Fibonacci numbers, we have

$$\mathbf{x}_{k+1} = \begin{bmatrix} F_{k+1} \\ F_{k+2} \end{bmatrix} = \begin{bmatrix} F_{k+1} \\ F_{k+1} + F_k \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} F_k \\ F_{k+1} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \mathbf{x}_k.$$

Thus, writing $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ and $\mathbf{x}_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, we have the discrete dynamical system $\mathbf{x}_{k+1} = A\mathbf{x}_k$. This gives $\mathbf{x}_k = A^k \mathbf{x}_0$. To find A^k, we diagonalize the matrix A. We have

$$\det(\lambda I - A) = \begin{vmatrix} \lambda & -1 \\ -1 & \lambda - 1 \end{vmatrix} = \lambda^2 - \lambda - 1 = 0,$$

so

$$\alpha = \frac{1+\sqrt{5}}{2} \quad \text{and} \quad \beta = \frac{1-\sqrt{5}}{2}$$

are eigenvalues of A (the constant α is known as the golden ratio). The corresponding eigenvectors are

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ \alpha \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ \beta \end{bmatrix}.$$

So we can diagonalize $A = PDP^{-1}$, where

$$P = \begin{bmatrix} 1 & 1 \\ \alpha & \beta \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix}.$$

Since

$$P^{-1} = -\frac{1}{\sqrt{5}} \begin{bmatrix} \beta & -1 \\ -\alpha & 1 \end{bmatrix},$$

we find that

$$\begin{aligned} A^k &= PD^kP^{-1} = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 1 \\ \alpha & \beta \end{bmatrix} \begin{bmatrix} \alpha^k & 0 \\ 0 & \beta^k \end{bmatrix} \begin{bmatrix} -\beta & 1 \\ \alpha & -1 \end{bmatrix} \\ &= \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 1 \\ \alpha & \beta \end{bmatrix} \begin{bmatrix} -\beta\alpha^k & \alpha^k \\ \alpha\beta^k & -\beta^k \end{bmatrix} \\ &= \frac{1}{\sqrt{5}} \begin{bmatrix} \alpha\beta^k - \beta\alpha^k & \alpha^k - \beta^k \\ \alpha\beta^{k+1} - \beta\alpha^{k+1} & \alpha^{k+1} - \beta^{k+1} \end{bmatrix}. \end{aligned}$$

We therefore have

$$\begin{bmatrix} F_k \\ F_{k+1} \end{bmatrix} = \mathbf{x}_k = A^k\mathbf{x}_0 = \frac{1}{\sqrt{5}} \begin{bmatrix} \alpha\beta^k - \beta\alpha^k & \alpha^k - \beta^k \\ \alpha\beta^{k+1} - \beta\alpha^{k+1} & \alpha^{k+1} - \beta^{k+1} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} \alpha^k - \beta^k \\ \alpha^{k+1} - \beta^{k+1} \end{bmatrix},$$

which yields to Binet's formula.■