

lyryx with Open Texts

LINEAR ALGEBRA with Applications

Open Edition



ADAPTABLE | ACCESSIBLE | AFFORDABLE

Adapted for

Emory University

Math 221

Linear Algebra

Sections 1 & 2

Lectured and adapted by

Le Chen

April 15, 2021

le.chen@emory.edu

Course page

http://math.emory.edu/~lchen41/teaching/2021_Spring_Math221

by W. Keith Nicholson

Creative Commons License (CC BY-NC-SA)

Contents

1	Systems of Linear Equations	5
1.1	Solutions and Elementary Operations	6
1.2	Gaussian Elimination	16
1.3	Homogeneous Equations	28
	Supplementary Exercises for Chapter 1	37
2	Matrix Algebra	39
2.1	Matrix Addition, Scalar Multiplication, and Transposition	40
2.2	Matrix-Vector Multiplication	53
2.3	Matrix Multiplication	72
2.4	Matrix Inverses	91
2.5	Elementary Matrices	109
2.6	Linear Transformations	119
2.7	LU-Factorization	135
3	Determinants and Diagonalization	147
3.1	The Cofactor Expansion	148
3.2	Determinants and Matrix Inverses	163
3.3	Diagonalization and Eigenvalues	178
	Supplementary Exercises for Chapter 3	201
4	Vector Geometry	203
4.1	Vectors and Lines	204
4.2	Projections and Planes	223
4.3	More on the Cross Product	244
4.4	Linear Operators on \mathbb{R}^3	251
	Supplementary Exercises for Chapter 4	260
5	Vector Space \mathbb{R}^n	263
5.1	Subspaces and Spanning	264
5.2	Independence and Dimension	273
5.3	Orthogonality	287
5.4	Rank of a Matrix	297

5.5	Similarity and Diagonalization	307
	Supplementary Exercises for Chapter 5	320
6	Vector Spaces	321
6.1	Examples and Basic Properties	322
6.2	Subspaces and Spanning Sets	333
6.3	Linear Independence and Dimension	342
6.4	Finite Dimensional Spaces	354
	Supplementary Exercises for Chapter 6	364
7	Linear Transformations	365
7.1	Examples and Elementary Properties	366
7.2	Kernel and Image of a Linear Transformation	374
7.3	Isomorphisms and Composition	385
8	Orthogonality	399
8.1	Orthogonal Complements and Projections	400
8.2	Orthogonal Diagonalization	410
8.3	Positive Definite Matrices	421
8.4	QR-Factorization	427
8.5	Computing Eigenvalues	431
8.6	The Singular Value Decomposition	436
8.6.1	Singular Value Decompositions	436
8.6.2	Fundamental Subspaces	442
8.6.3	The Polar Decomposition of a Real Square Matrix	445
8.6.4	The Pseudoinverse of a Matrix	447

3.3 Diagonalization and Eigenvalues

The world is filled with examples of systems that evolve in time—the weather in a region, the economy of a nation, the diversity of an ecosystem, etc. Describing such systems is difficult in general and various methods have been developed in special cases. In this section we describe one such method, called *diagonalization*, which is one of the most important techniques in linear algebra. A very fertile example of this procedure is in modelling the growth of the population of an animal species. This has attracted more attention in recent years with the ever increasing awareness that many species are endangered. To motivate the technique, we begin by setting up a simple model of a bird population in which we make assumptions about survival and reproduction rates.

Example 3.3.1

Consider the evolution of the population of a species of birds. Because the number of males and females are nearly equal, we count only females. We assume that each female remains a juvenile for one year and then becomes an adult, and that only adults have offspring. We make three assumptions about reproduction and survival rates:

1. The number of juvenile females hatched in any year is twice the number of adult females alive the year before (we say the **reproduction rate** is 2).
2. Half of the adult females in any year survive to the next year (the **adult survival rate** is $\frac{1}{2}$).
3. One quarter of the juvenile females in any year survive into adulthood (the **juvenile survival rate** is $\frac{1}{4}$).

If there were 100 adult females and 40 juvenile females alive initially, compute the population of females k years later.

Solution. Let a_k and j_k denote, respectively, the number of adult and juvenile females after k years, so that the total female population is the sum $a_k + j_k$. Assumption 1 shows that $j_{k+1} = 2a_k$, while assumptions 2 and 3 show that $a_{k+1} = \frac{1}{2}a_k + \frac{1}{4}j_k$. Hence the numbers a_k and j_k in successive years are related by the following equations:

$$\begin{aligned} a_{k+1} &= \frac{1}{2}a_k + \frac{1}{4}j_k \\ j_{k+1} &= 2a_k \end{aligned}$$

If we write $\mathbf{v}_k = \begin{bmatrix} a_k \\ j_k \end{bmatrix}$ and $A = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} \\ 2 & 0 \end{bmatrix}$ these equations take the matrix form

$$\mathbf{v}_{k+1} = A\mathbf{v}_k, \quad \text{for each } k = 0, 1, 2, \dots$$

Taking $k = 0$ gives $\mathbf{v}_1 = A\mathbf{v}_0$, then taking $k = 1$ gives $\mathbf{v}_2 = A\mathbf{v}_1 = A^2\mathbf{v}_0$, and taking $k = 2$ gives $\mathbf{v}_3 = A\mathbf{v}_2 = A^3\mathbf{v}_0$. Continuing in this way, we get

$$\mathbf{v}_k = A^k\mathbf{v}_0, \quad \text{for each } k = 0, 1, 2, \dots$$

Since $\mathbf{v}_0 = \begin{bmatrix} a_0 \\ j_0 \end{bmatrix} = \begin{bmatrix} 100 \\ 40 \end{bmatrix}$ is known, finding the population profile \mathbf{v}_k amounts to computing A^k for all $k \geq 0$. We will complete this calculation in Example 3.3.12 after some new techniques have been developed.

Let A be a fixed $n \times n$ matrix. A sequence $\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2, \dots$ of column vectors in \mathbb{R}^n is called a **linear dynamical system**⁸ if \mathbf{v}_0 is known and the other \mathbf{v}_k are determined (as in Example 3.3.1) by the conditions

$$\mathbf{v}_{k+1} = A\mathbf{v}_k \text{ for each } k = 0, 1, 2, \dots$$

These conditions are called a **matrix recurrence** for the vectors \mathbf{v}_k . As in Example 3.3.1, they imply that

$$\mathbf{v}_k = A^k \mathbf{v}_0 \text{ for all } k \geq 0$$

so finding the columns \mathbf{v}_k amounts to calculating A^k for $k \geq 0$.

Direct computation of the powers A^k of a square matrix A can be time-consuming, so we adopt an indirect method that is commonly used. The idea is to first **diagonalize** the matrix A , that is, to find an invertible matrix P such that

$$P^{-1}AP = D \text{ is a diagonal matrix} \tag{3.8}$$

This works because the powers D^k of the diagonal matrix D are easy to compute, and Equation 3.8 enables us to compute powers A^k of the matrix A in terms of powers D^k of D . Indeed, we can solve Equation 3.8 for A to get $A = PDP^{-1}$. Squaring this gives

$$A^2 = (PDP^{-1})(PDP^{-1}) = PD^2P^{-1}$$

Using this we can compute A^3 as follows:

$$A^3 = AA^2 = (PDP^{-1})(PD^2P^{-1}) = PD^3P^{-1}$$

Continuing in this way we obtain Theorem 3.3.1 (even if D is not diagonal).

Theorem 3.3.1

If $A = PDP^{-1}$ then $A^k = PD^kP^{-1}$ for each $k = 1, 2, \dots$

Hence computing A^k comes down to finding an invertible matrix P as in equation Equation 3.8. To do this it is necessary to first compute certain numbers (called eigenvalues) associated with the matrix A .

⁸More precisely, this is a *linear discrete* dynamical system. Many models regard \mathbf{v}_t as a continuous function of the time t , and replace our condition between \mathbf{b}_{k+1} and $A\mathbf{v}_k$ with a differential relationship viewed as functions of time.

Eigenvalues and Eigenvectors

Definition 3.4 Eigenvalues and Eigenvectors of a Matrix

If A is an $n \times n$ matrix, a number λ is called an **eigenvalue** of A if

$$A\mathbf{x} = \lambda\mathbf{x} \text{ for some column } \mathbf{x} \neq \mathbf{0} \text{ in } \mathbb{R}^n$$

In this case, \mathbf{x} is called an **eigenvector** of A corresponding to the eigenvalue λ , or a **λ -eigenvector** for short.

Example 3.3.2

If $A = \begin{bmatrix} 3 & 5 \\ 1 & -1 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$ then $A\mathbf{x} = 4\mathbf{x}$ so $\lambda = 4$ is an eigenvalue of A with corresponding eigenvector \mathbf{x} .

The matrix A in Example 3.3.2 has another eigenvalue in addition to $\lambda = 4$. To find it, we develop a general procedure for *any* $n \times n$ matrix A .

By definition a number λ is an eigenvalue of the $n \times n$ matrix A if and only if $A\mathbf{x} = \lambda\mathbf{x}$ for some column $\mathbf{x} \neq \mathbf{0}$. This is equivalent to asking that the homogeneous system

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

of linear equations has a nontrivial solution $\mathbf{x} \neq \mathbf{0}$. By Theorem 2.4.5 this happens if and only if the matrix $\lambda I - A$ is not invertible and this, in turn, holds if and only if the determinant of the coefficient matrix is zero:

$$\det(\lambda I - A) = 0$$

This last condition prompts the following definition:

Definition 3.5 Characteristic Polynomial of a Matrix

If A is an $n \times n$ matrix, the **characteristic polynomial** $c_A(x)$ of A is defined by

$$c_A(x) = \det(xI - A)$$

Note that $c_A(x)$ is indeed a polynomial in the variable x , and it has degree n when A is an $n \times n$ matrix (this is illustrated in the examples below). The above discussion shows that a number λ is an eigenvalue of A if and only if $c_A(\lambda) = 0$, that is if and only if λ is a **root** of the characteristic polynomial $c_A(x)$. We record these observations in

Theorem 3.3.2

Let A be an $n \times n$ matrix.

1. The eigenvalues λ of A are the roots of the characteristic polynomial $c_A(x)$ of A .
2. The λ -eigenvectors \mathbf{x} are the nonzero solutions to the homogeneous system

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

of linear equations with $\lambda I - A$ as coefficient matrix.

In practice, solving the equations in part 2 of Theorem 3.3.2 is a routine application of gaussian elimination, but finding the eigenvalues can be difficult, often requiring computers (see Section 8.5). For now, the examples and exercises will be constructed so that the roots of the characteristic polynomials are relatively easy to find (usually integers). However, the reader should not be misled by this into thinking that eigenvalues are so easily obtained for the matrices that occur in practical applications!

Example 3.3.3

Find the characteristic polynomial of the matrix $A = \begin{bmatrix} 3 & 5 \\ 1 & -1 \end{bmatrix}$ discussed in Example 3.3.2, and then find all the eigenvalues and their eigenvectors.

Solution. Since $xI - A = \begin{bmatrix} x & 0 \\ 0 & x \end{bmatrix} - \begin{bmatrix} 3 & 5 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} x-3 & -5 \\ -1 & x+1 \end{bmatrix}$ we get

$$c_A(x) = \det \begin{bmatrix} x-3 & -5 \\ -1 & x+1 \end{bmatrix} = x^2 - 2x - 8 = (x-4)(x+2)$$

Hence, the roots of $c_A(x)$ are $\lambda_1 = 4$ and $\lambda_2 = -2$, so these are the eigenvalues of A . Note that $\lambda_1 = 4$ was the eigenvalue mentioned in Example 3.3.2, but we have found a new one: $\lambda_2 = -2$.

To find the eigenvectors corresponding to $\lambda_2 = -2$, observe that in this case

$$(\lambda_2 I - A)\mathbf{x} = \begin{bmatrix} \lambda_2 - 3 & -5 \\ -1 & \lambda_2 + 1 \end{bmatrix} = \begin{bmatrix} -5 & -5 \\ -1 & -1 \end{bmatrix}$$

so the general solution to $(\lambda_2 I - A)\mathbf{x} = \mathbf{0}$ is $\mathbf{x} = t \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ where t is an arbitrary real number.

Hence, the eigenvectors \mathbf{x} corresponding to λ_2 are $\mathbf{x} = t \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ where $t \neq 0$ is arbitrary.

Similarly, $\lambda_1 = 4$ gives rise to the eigenvectors $\mathbf{x} = t \begin{bmatrix} 5 \\ 1 \end{bmatrix}$, $t \neq 0$ which includes the observation in Example 3.3.2.

Note that a square matrix A has *many* eigenvectors associated with any given eigenvalue λ .

In fact *every* nonzero solution \mathbf{x} of $(\lambda I - A)\mathbf{x} = \mathbf{0}$ is an eigenvector. Recall that these solutions are all linear combinations of certain basic solutions determined by the gaussian algorithm (see Theorem 1.3.2). Observe that any nonzero multiple of an eigenvector is again an eigenvector,⁹ and such multiples are often more convenient.¹⁰ Any set of nonzero multiples of the basic solutions of $(\lambda I - A)\mathbf{x} = \mathbf{0}$ will be called a set of **basic eigenvectors** corresponding to λ .

Example 3.3.4

Find the characteristic polynomial, eigenvalues, and basic eigenvectors for

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

Solution. Here the characteristic polynomial is given by

$$c_A(x) = \det \begin{bmatrix} x-2 & 0 & 0 \\ -1 & x-2 & 1 \\ -1 & -3 & x+2 \end{bmatrix} = (x-2)(x-1)(x+1)$$

so the eigenvalues are $\lambda_1 = 2$, $\lambda_2 = 1$, and $\lambda_3 = -1$. To find all eigenvectors for $\lambda_1 = 2$, compute

$$\lambda_1 I - A = \begin{bmatrix} \lambda_1 - 2 & 0 & 0 \\ -1 & \lambda_1 - 2 & 1 \\ -1 & -3 & \lambda_1 + 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 1 \\ -1 & -3 & 4 \end{bmatrix}$$

We want the (nonzero) solutions to $(\lambda_1 I - A)\mathbf{x} = \mathbf{0}$. The augmented matrix becomes

$$\left[\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & -3 & 4 & 0 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

using row operations. Hence, the general solution \mathbf{x} to $(\lambda_1 I - A)\mathbf{x} = \mathbf{0}$ is $\mathbf{x} = t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ where t

is arbitrary, so we can use $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ as the basic eigenvector corresponding to $\lambda_1 = 2$. As

the reader can verify, the gaussian algorithm gives basic eigenvectors $\mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ and

$\mathbf{x}_3 = \begin{bmatrix} 0 \\ \frac{1}{3} \\ 1 \end{bmatrix}$ corresponding to $\lambda_2 = 1$ and $\lambda_3 = -1$, respectively. Note that to eliminate

fractions, we could instead use $3\mathbf{x}_3 = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$ as the basic λ_3 -eigenvector.

⁹In fact, any nonzero linear combination of λ -eigenvectors is again a λ -eigenvector.

¹⁰Allowing nonzero multiples helps eliminate round-off error when the eigenvectors involve fractions.

Example 3.3.5

If A is a square matrix, show that A and A^T have the same characteristic polynomial, and hence the same eigenvalues.

Solution. We use the fact that $xI - A^T = (xI - A)^T$. Then

$$c_{A^T}(x) = \det(xI - A^T) = \det[(xI - A)^T] = \det(xI - A) = c_A(x)$$

by Theorem 3.2.3. Hence $c_{A^T}(x)$ and $c_A(x)$ have the same roots, and so A^T and A have the same eigenvalues (by Theorem 3.3.2).

The eigenvalues of a matrix need not be distinct. For example, if $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ the characteristic polynomial is $(x - 1)^2$ so the eigenvalue 1 occurs twice. Furthermore, eigenvalues are usually not computed as the roots of the characteristic polynomial. There are iterative, numerical methods (for example the QR-algorithm in Section 8.5) that are much more efficient for large matrices.

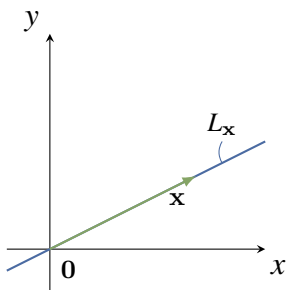
A-Invariance

If A is a 2×2 matrix, we can describe the eigenvectors of A geometrically using the following concept. A line L through the origin in \mathbb{R}^2 is called **A-invariant** if $A\mathbf{x}$ is in L whenever \mathbf{x} is in L . If we think of A as a linear transformation $\mathbb{R}^2 \rightarrow \mathbb{R}^2$, this asks that A carries L into itself, that is the image $A\mathbf{x}$ of each vector \mathbf{x} in L is again in L .

Example 3.3.6

The x axis $L = \left\{ \begin{bmatrix} x \\ 0 \end{bmatrix} \mid x \text{ in } \mathbb{R} \right\}$ is A -invariant for any matrix of the form

$$A = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \text{ because } \begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \begin{bmatrix} x \\ 0 \end{bmatrix} = \begin{bmatrix} ax \\ 0 \end{bmatrix} \text{ is } L \text{ for all } \mathbf{x} = \begin{bmatrix} x \\ 0 \end{bmatrix} \text{ in } L$$



To see the connection with eigenvectors, let $\mathbf{x} \neq \mathbf{0}$ be any nonzero vector in \mathbb{R}^2 and let $L_{\mathbf{x}}$ denote the unique line through the origin containing \mathbf{x} (see the diagram). By the definition of scalar multiplication in Section 2.6, we see that $L_{\mathbf{x}}$ consists of all scalar multiples of \mathbf{x} , that is

$$L_{\mathbf{x}} = \mathbb{R}\mathbf{x} = \{t\mathbf{x} \mid t \text{ in } \mathbb{R}\}$$

Now suppose that \mathbf{x} is an eigenvector of A , say $A\mathbf{x} = \lambda\mathbf{x}$ for some λ in \mathbb{R} . Then if $t\mathbf{x}$ is in $L_{\mathbf{x}}$ then

$$A(t\mathbf{x}) = t(A\mathbf{x}) = t(\lambda\mathbf{x}) = (t\lambda)\mathbf{x} \text{ is again in } L_{\mathbf{x}}$$

That is, $L_{\mathbf{x}}$ is A -invariant. On the other hand, if $L_{\mathbf{x}}$ is A -invariant then $A\mathbf{x}$ is in $L_{\mathbf{x}}$ (since \mathbf{x} is in $L_{\mathbf{x}}$). Hence $A\mathbf{x} = t\mathbf{x}$ for some t in \mathbb{R} , so \mathbf{x} is an eigenvector for A (with eigenvalue t). This proves:

Theorem 3.3.3

Let A be a 2×2 matrix, let $\mathbf{x} \neq \mathbf{0}$ be a vector in \mathbb{R}^2 , and let $L_{\mathbf{x}}$ be the line through the origin in \mathbb{R}^2 containing \mathbf{x} . Then

\mathbf{x} is an eigenvector of A if and only if $L_{\mathbf{x}}$ is A -invariant

Example 3.3.7

1. If θ is not a multiple of π , show that $A = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$ has no real eigenvalue.
2. If m is real show that $B = \frac{1}{1+m^2} \begin{bmatrix} 1-m^2 & 2m \\ 2m & m^2-1 \end{bmatrix}$ has a 1 as an eigenvalue.

Solution.

1. A induces rotation about the origin through the angle θ (Theorem 2.6.4). Since θ is not a multiple of π , this shows that no line through the origin is A -invariant. Hence A has no eigenvector by Theorem 3.3.3, and so has no eigenvalue.
2. B induces reflection Q_m in the line through the origin with slope m by Theorem 2.6.5. If \mathbf{x} is any nonzero point on this line then it is clear that $Q_m \mathbf{x} = \mathbf{x}$, that is $Q_m \mathbf{x} = 1\mathbf{x}$. Hence 1 is an eigenvalue (with eigenvector \mathbf{x}).

If $\theta = \frac{\pi}{2}$ in Example 3.3.7, then $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ so $c_A(x) = x^2 + 1$. This polynomial has no root in \mathbb{R} , so A has no (real) eigenvalue, and hence no eigenvector. In fact its eigenvalues are the complex numbers i and $-i$, with corresponding eigenvectors $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ and $\begin{bmatrix} 1 \\ i \end{bmatrix}$. In other words, A has eigenvalues and eigenvectors, just not real ones.

Note that *every* polynomial has complex roots,¹¹ so every matrix has complex eigenvalues. While these eigenvalues may very well be real, this suggests that we really should be doing linear algebra over the complex numbers. Indeed, everything we have done (gaussian elimination, matrix algebra, determinants, etc.) works if all the scalars are complex.

¹¹This is called the *Fundamental Theorem of Algebra* and was first proved by Gauss in his doctoral dissertation.

Diagonalization

An $n \times n$ matrix D is called a **diagonal matrix** if all its entries off the main diagonal are zero, that is if D has the form

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

where $\lambda_1, \lambda_2, \dots, \lambda_n$ are numbers. Calculations with diagonal matrices are very easy. Indeed, if $D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ and $E = \text{diag}(\mu_1, \mu_2, \dots, \mu_n)$ are two diagonal matrices, their product DE and sum $D+E$ are again diagonal, and are obtained by doing the same operations to corresponding diagonal elements:

$$\begin{aligned} DE &= \text{diag}(\lambda_1\mu_1, \lambda_2\mu_2, \dots, \lambda_n\mu_n) \\ D+E &= \text{diag}(\lambda_1+\mu_1, \lambda_2+\mu_2, \dots, \lambda_n+\mu_n) \end{aligned}$$

Because of the simplicity of these formulas, and with an eye on Theorem 3.3.1 and the discussion preceding it, we make another definition:

Definition 3.6 Diagonalizable Matrices

An $n \times n$ matrix A is called **diagonalizable** if

$$P^{-1}AP \text{ is diagonal for some invertible } n \times n \text{ matrix } P$$

Here the invertible matrix P is called a **diagonalizing matrix** for A .

To discover when such a matrix P exists, we let $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ denote the columns of P and look for ways to determine when such \mathbf{x}_i exist and how to compute them. To this end, write P in terms of its columns as follows:

$$P = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]$$

Observe that $P^{-1}AP = D$ for some diagonal matrix D holds if and only if

$$AP = PD$$

If we write $D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, where the λ_i are numbers to be determined, the equation $AP = PD$ becomes

$$A[\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

By the definition of matrix multiplication, each side simplifies as follows

$$[A\mathbf{x}_1 \quad A\mathbf{x}_2 \quad \cdots \quad A\mathbf{x}_n] = [\lambda_1\mathbf{x}_1 \quad \lambda_2\mathbf{x}_2 \quad \cdots \quad \lambda_n\mathbf{x}_n]$$

Comparing columns shows that $A\mathbf{x}_i = \lambda_i\mathbf{x}_i$ for each i , so

$$P^{-1}AP = D \quad \text{if and only if } A\mathbf{x}_i = \lambda_i\mathbf{x}_i \text{ for each } i$$

In other words, $P^{-1}AP = D$ holds if and only if the diagonal entries of D are eigenvalues of A and the columns of P are corresponding eigenvectors. This proves the following fundamental result.

Theorem 3.3.4

Let A be an $n \times n$ matrix.

1. A is diagonalizable if and only if it has eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ such that the matrix $P = [\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_n]$ is invertible.
2. When this is the case, $P^{-1}AP = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ where, for each i , λ_i is the eigenvalue of A corresponding to \mathbf{x}_i .

Example 3.3.8

Diagonalize the matrix $A = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 2 & -1 \\ 1 & 3 & -2 \end{bmatrix}$ in Example 3.3.4.

Solution. By Example 3.3.4, the eigenvalues of A are $\lambda_1 = 2$, $\lambda_2 = 1$, and $\lambda_3 = -1$, with corresponding basic eigenvectors $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$, and $\mathbf{x}_3 = \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}$ respectively.

Since the matrix $P = [\mathbf{x}_1 \ \mathbf{x}_2 \ \mathbf{x}_3] = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 3 \end{bmatrix}$ is invertible, Theorem 3.3.4 guarantees that

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = D$$

The reader can verify this directly—easier to check $AP = PD$.

In Example 3.3.8, suppose we let $Q = [\mathbf{x}_2 \ \mathbf{x}_1 \ \mathbf{x}_3]$ be the matrix formed from the eigenvectors $\mathbf{x}_1, \mathbf{x}_2$, and \mathbf{x}_3 of A , but in a *different order* than that used to form P . Then $Q^{-1}AQ = \text{diag}(\lambda_2, \lambda_1, \lambda_3)$ is diagonal by Theorem 3.3.4, but the eigenvalues are in the *new order*. Hence we can choose the diagonalizing matrix P so that the eigenvalues λ_i appear in any order we want along the main diagonal of D .

In every example above each eigenvalue has had only one basic eigenvector. Here is a diagonalizable matrix where this is not the case.

Example 3.3.9

Diagonalize the matrix $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$

Solution. To compute the characteristic polynomial of A first add rows 2 and 3 of $xI - A$ to row 1:

$$\begin{aligned} c_A(x) &= \det \begin{bmatrix} x & -1 & -1 \\ -1 & x & -1 \\ -1 & -1 & x \end{bmatrix} = \det \begin{bmatrix} x-2 & x-2 & x-2 \\ -1 & x & -1 \\ -1 & -1 & x \end{bmatrix} \\ &= \det \begin{bmatrix} x-2 & 0 & 0 \\ -1 & x+1 & 0 \\ -1 & 0 & x+1 \end{bmatrix} = (x-2)(x+1)^2 \end{aligned}$$

Hence the eigenvalues are $\lambda_1 = 2$ and $\lambda_2 = -1$, with λ_2 repeated twice (we say that λ_2 has *multiplicity two*). However, A is diagonalizable. For $\lambda_1 = 2$, the system of equations

$(\lambda_1 I - A)\mathbf{x} = \mathbf{0}$ has general solution $\mathbf{x} = t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ as the reader can verify, so a basic

λ_1 -eigenvector is $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

Turning to the repeated eigenvalue $\lambda_2 = -1$, we must solve $(\lambda_2 I - A)\mathbf{x} = \mathbf{0}$. By gaussian elimination, the general solution is $\mathbf{x} = s \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ where s and t are arbitrary.

Hence the gaussian algorithm produces *two* basic λ_2 -eigenvectors $\mathbf{x}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ and

$\mathbf{y}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ If we take $P = [\mathbf{x}_1 \ \mathbf{x}_2 \ \mathbf{y}_2] = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$ we find that P is invertible.

Hence $P^{-1}AP = \text{diag}(2, -1, -1)$ by Theorem 3.3.4.

Example 3.3.9 typifies every diagonalizable matrix. To describe the general case, we need some terminology.

Definition 3.7 Multiplicity of an Eigenvalue

An eigenvalue λ of a square matrix A is said to have **multiplicity** m if it occurs m times as a root of the characteristic polynomial $c_A(x)$.

For example, the eigenvalue $\lambda_2 = -1$ in Example 3.3.9 has multiplicity 2. In that example the gaussian algorithm yields two basic λ_2 -eigenvectors, the same number as the multiplicity. This

works in general.

Theorem 3.3.5

A square matrix A is diagonalizable if and only if every eigenvalue λ of multiplicity m yields exactly m basic eigenvectors; that is, if and only if the general solution of the system $(\lambda I - A)\mathbf{x} = \mathbf{0}$ has exactly m parameters.

One case of Theorem 3.3.5 deserves mention.

Theorem 3.3.6

An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

The proofs of Theorem 3.3.5 and Theorem 3.3.6 require more advanced techniques and are given in Chapter 5. The following procedure summarizes the method.

Diagonalization Algorithm

To diagonalize an $n \times n$ matrix A :

Step 1. Find the distinct eigenvalues λ of A .

Step 2. Compute a set of basic eigenvectors corresponding to each of these eigenvalues λ as basic solutions of the homogeneous system $(\lambda I - A)\mathbf{x} = \mathbf{0}$.

Step 3. The matrix A is diagonalizable if and only if there are n basic eigenvectors in all.

Step 4. If A is diagonalizable, the $n \times n$ matrix P with these basic eigenvectors as its columns is a diagonalizing matrix for A , that is, P is invertible and $P^{-1}AP$ is diagonal.

The diagonalization algorithm is valid even if the eigenvalues are nonreal complex numbers. In this case the eigenvectors will also have complex entries, but we will not pursue this here.

Example 3.3.10

Show that $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ is not diagonalizable.

Solution 1. The characteristic polynomial is $c_A(x) = (x - 1)^2$, so A has only one eigenvalue $\lambda_1 = 1$ of multiplicity 2. But the system of equations $(\lambda_1 I - A)\mathbf{x} = \mathbf{0}$ has general solution $t \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, so there is only one parameter, and so only one basic eigenvector $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Hence A is not diagonalizable.

Solution 2. We have $c_A(x) = (x - 1)^2$ so the only eigenvalue of A is $\lambda = 1$. Hence, if A were diagonalizable, Theorem 3.3.4 would give $P^{-1}AP = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$ for some invertible matrix

P . But then $A = PIP^{-1} = I$, which is not the case. So A cannot be diagonalizable.

Diagonalizable matrices share many properties of their eigenvalues. The following example illustrates why.

Example 3.3.11

If $\lambda^3 = 5\lambda$ for every eigenvalue of the diagonalizable matrix A , show that $A^3 = 5A$.

Solution. Let $P^{-1}AP = D = \text{diag}(\lambda_1, \dots, \lambda_n)$. Because $\lambda_i^3 = 5\lambda_i$ for each i , we obtain

$$D^3 = \text{diag}(\lambda_1^3, \dots, \lambda_n^3) = \text{diag}(5\lambda_1, \dots, 5\lambda_n) = 5D$$

Hence $A^3 = (PDP^{-1})^3 = PD^3P^{-1} = P(5D)P^{-1} = 5(PDP^{-1}) = 5A$ using Theorem 3.3.1. This is what we wanted.

If $p(x)$ is any polynomial and $p(\lambda) = 0$ for every eigenvalue of the diagonalizable matrix A , an argument similar to that in Example 3.3.11 shows that $p(A) = 0$. Thus Example 3.3.11 deals with the case $p(x) = x^3 - 5x$. In general, $p(A)$ is called the *evaluation* of the polynomial $p(x)$ at the matrix A . For example, if $p(x) = 2x^3 - 3x + 5$, then $p(A) = 2A^3 - 3A + 5I$ —note the use of the identity matrix.

In particular, if $c_A(x)$ denotes the characteristic polynomial of A , we certainly have $c_A(\lambda) = 0$ for each eigenvalue λ of A (Theorem 3.3.2). Hence $c_A(A) = 0$ for every diagonalizable matrix A . This is, in fact, true for *any* square matrix, diagonalizable or not, and the general result is called the Cayley-Hamilton theorem. It is proved in Section ?? and again in Section ??.

Linear Dynamical Systems

We began Section 3.3 with an example from ecology which models the evolution of the population of a species of birds as time goes on. As promised, we now complete the example—Example 3.3.12 below.

The bird population was described by computing the female population profile $\mathbf{v}_k = \begin{bmatrix} a_k \\ j_k \end{bmatrix}$ of the species, where a_k and j_k represent the number of adult and juvenile females present k years after the initial values a_0 and j_0 were observed. The model assumes that these numbers are related by the following equations:

$$\begin{aligned} a_{k+1} &= \frac{1}{2}a_k + \frac{1}{4}j_k \\ j_{k+1} &= 2a_k \end{aligned}$$

If we write $A = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} \\ 2 & 0 \end{bmatrix}$ the columns \mathbf{v}_k satisfy $\mathbf{v}_{k+1} = A\mathbf{v}_k$ for each $k = 0, 1, 2, \dots$

Hence $\mathbf{v}_k = A^k\mathbf{v}_0$ for each $k = 1, 2, \dots$. We can now use our diagonalization techniques to determine the population profile \mathbf{v}_k for all values of k in terms of the initial values.

Example 3.3.12

Assuming that the initial values were $a_0 = 100$ adult females and $j_0 = 40$ juvenile females, compute a_k and j_k for $k = 1, 2, \dots$

Solution. The characteristic polynomial of the matrix $A = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} \\ 2 & 0 \end{bmatrix}$ is $c_A(x) = x^2 - \frac{1}{2}x - \frac{1}{2} = (x-1)(x+\frac{1}{2})$, so the eigenvalues are $\lambda_1 = 1$ and $\lambda_2 = -\frac{1}{2}$ and gaussian elimination gives corresponding basic eigenvectors $\begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix}$ and $\begin{bmatrix} -\frac{1}{4} \\ 1 \end{bmatrix}$. For convenience, we can use multiples $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\mathbf{x}_2 = \begin{bmatrix} -1 \\ 4 \end{bmatrix}$ respectively. Hence a diagonalizing matrix is $P = \begin{bmatrix} 1 & -1 \\ 2 & 4 \end{bmatrix}$ and we obtain

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

This gives $A = PDP^{-1}$ so, for each $k \geq 0$, we can compute A^k explicitly:

$$\begin{aligned} A^k &= PD^kP^{-1} = \begin{bmatrix} 1 & -1 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & (-\frac{1}{2})^k \end{bmatrix} \frac{1}{6} \begin{bmatrix} 4 & 1 \\ -2 & 4 \end{bmatrix} \\ &= \frac{1}{6} \begin{bmatrix} 4 + 2(-\frac{1}{2})^k & 1 - (-\frac{1}{2})^k \\ 8 - 8(-\frac{1}{2})^k & 2 + 4(-\frac{1}{2})^k \end{bmatrix} \end{aligned}$$

Hence we obtain

$$\begin{aligned} \begin{bmatrix} a_k \\ j_k \end{bmatrix} &= \mathbf{v}_k = A^k \mathbf{v}_0 = \frac{1}{6} \begin{bmatrix} 4 + 2(-\frac{1}{2})^k & 1 - (-\frac{1}{2})^k \\ 8 - 8(-\frac{1}{2})^k & 2 + 4(-\frac{1}{2})^k \end{bmatrix} \begin{bmatrix} 100 \\ 40 \end{bmatrix} \\ &= \frac{1}{6} \begin{bmatrix} 440 + 160(-\frac{1}{2})^k \\ 880 - 640(-\frac{1}{2})^k \end{bmatrix} \end{aligned}$$

Equating top and bottom entries, we obtain exact formulas for a_k and j_k :

$$a_k = \frac{220}{3} + \frac{80}{3} \left(-\frac{1}{2}\right)^k \text{ and } j_k = \frac{440}{3} + \frac{320}{3} \left(-\frac{1}{2}\right)^k \text{ for } k = 1, 2, \dots$$

In practice, the exact values of a_k and j_k are not usually required. What is needed is a measure of how these numbers behave for large values of k . This is easy to obtain here. Since $(-\frac{1}{2})^k$ is nearly zero for large k , we have the following approximate values

$$a_k \approx \frac{220}{3} \text{ and } j_k \approx \frac{440}{3} \text{ if } k \text{ is large}$$

Hence, in the long term, the female population stabilizes with approximately twice as many juveniles as adults.

Definition 3.8 Linear Dynamical System

If A is an $n \times n$ matrix, a sequence $\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2, \dots$ of columns in \mathbb{R}^n is called a **linear dynamical system** if \mathbf{v}_0 is specified and $\mathbf{v}_1, \mathbf{v}_2, \dots$ are given by the matrix recurrence $\mathbf{v}_{k+1} = A\mathbf{v}_k$ for each $k \geq 0$. We call A the **migration** matrix of the system.

We have $\mathbf{v}_1 = A\mathbf{v}_0$, then $\mathbf{v}_2 = A\mathbf{v}_1 = A^2\mathbf{v}_0$, and continuing we find

$$\mathbf{v}_k = A^k \mathbf{v}_0 \text{ for each } k = 1, 2, \dots \quad (3.9)$$

Hence the columns \mathbf{v}_k are determined by the powers A^k of the matrix A and, as we have seen, these powers can be efficiently computed if A is diagonalizable. In fact Equation 3.9 can be used to give a nice “formula” for the columns \mathbf{v}_k in this case.

Assume that A is diagonalizable with eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ and corresponding basic eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$. If $P = [\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_n]$ is a diagonalizing matrix with the \mathbf{x}_i as columns, then P is invertible and

$$P^{-1}AP = D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

by Theorem 3.3.4. Hence $A = PDP^{-1}$ so Equation 3.9 and Theorem 3.3.1 give

$$\mathbf{v}_k = A^k \mathbf{v}_0 = (PDP^{-1})^k \mathbf{v}_0 = (PD^k P^{-1}) \mathbf{v}_0 = PD^k (P^{-1} \mathbf{v}_0)$$

for each $k = 1, 2, \dots$. For convenience, we denote the column $P^{-1} \mathbf{v}_0$ arising here as follows:

$$\mathbf{b} = P^{-1} \mathbf{v}_0 = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

Then matrix multiplication gives

$$\begin{aligned} \mathbf{v}_k &= PD^k (P^{-1} \mathbf{v}_0) \\ &= [\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_n] \begin{bmatrix} \lambda_1^k & 0 & \dots & 0 \\ 0 & \lambda_2^k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n^k \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \\ &= [\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_n] \begin{bmatrix} b_1 \lambda_1^k \\ b_2 \lambda_2^k \\ \vdots \\ b_n \lambda_n^k \end{bmatrix} \\ &= b_1 \lambda_1^k \mathbf{x}_1 + b_2 \lambda_2^k \mathbf{x}_2 + \dots + b_n \lambda_n^k \mathbf{x}_n \end{aligned} \quad (3.10)$$

for each $k \geq 0$. This is a useful **exact formula** for the columns \mathbf{v}_k . Note that, in particular,

$$\mathbf{v}_0 = b_1 \mathbf{x}_1 + b_2 \mathbf{x}_2 + \dots + b_n \mathbf{x}_n$$

However, such an exact formula for \mathbf{v}_k is often not required in practice; all that is needed is to *estimate* \mathbf{v}_k for large values of k (as was done in Example 3.3.12). This can be easily done if A has a largest eigenvalue. An eigenvalue λ of a matrix A is called a **dominant eigenvalue** of A if it has multiplicity 1 and

$$|\lambda| > |\mu| \text{ for all eigenvalues } \mu \neq \lambda$$

where $|\lambda|$ denotes the absolute value of the number λ . For example, $\lambda_1 = 1$ is dominant in Example 3.3.12.

Returning to the above discussion, suppose that A has a dominant eigenvalue. By choosing the order in which the columns \mathbf{x}_i are placed in P , we may assume that λ_1 is dominant among the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ of A (see the discussion following Example 3.3.8). Now recall the exact expression for \mathbf{v}_k in Equation 3.10 above:

$$\mathbf{v}_k = b_1 \lambda_1^k \mathbf{x}_1 + b_2 \lambda_2^k \mathbf{x}_2 + \cdots + b_n \lambda_n^k \mathbf{x}_n$$

Take λ_1^k out as a common factor in this equation to get

$$\mathbf{v}_k = \lambda_1^k \left[b_1 \mathbf{x}_1 + b_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k \mathbf{x}_2 + \cdots + b_n \left(\frac{\lambda_n}{\lambda_1} \right)^k \mathbf{x}_n \right]$$

for each $k \geq 0$. Since λ_1 is dominant, we have $|\lambda_i| < |\lambda_1|$ for each $i \geq 2$, so each of the numbers $(\lambda_i/\lambda_1)^k$ become small in absolute value as k increases. Hence \mathbf{v}_k is approximately equal to the first term $\lambda_1^k b_1 \mathbf{x}_1$, and we write this as $\mathbf{v}_k \approx \lambda_1^k b_1 \mathbf{x}_1$. These observations are summarized in the following theorem (together with the above exact formula for \mathbf{v}_k).

Theorem 3.3.7

Consider the dynamical system $\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2, \dots$ with matrix recurrence

$$\mathbf{v}_{k+1} = A \mathbf{v}_k \text{ for } k \geq 0$$

where A and \mathbf{v}_0 are given. Assume that A is a diagonalizable $n \times n$ matrix with eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ and corresponding basic eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$, and let

$P = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{bmatrix}$ be the diagonalizing matrix. Then an exact formula for \mathbf{v}_k is

$$\mathbf{v}_k = b_1 \lambda_1^k \mathbf{x}_1 + b_2 \lambda_2^k \mathbf{x}_2 + \cdots + b_n \lambda_n^k \mathbf{x}_n \text{ for each } k \geq 0$$

where the coefficients b_i come from

$$\mathbf{b} = P^{-1} \mathbf{v}_0 = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

Moreover, if A has dominant¹² eigenvalue λ_1 , then \mathbf{v}_k is approximated by

$$\mathbf{v}_k \approx b_1 \lambda_1^k \mathbf{x}_1 \text{ for sufficiently large } k.$$

Example 3.3.13

Returning to Example 3.3.12, we see that $\lambda_1 = 1$ is the dominant eigenvalue, with eigenvector $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Here $P = \begin{bmatrix} 1 & -1 \\ 2 & 4 \end{bmatrix}$ and $\mathbf{v}_0 = \begin{bmatrix} 100 \\ 40 \end{bmatrix}$ so $P^{-1}\mathbf{v}_0 = \frac{1}{3} \begin{bmatrix} 220 \\ -80 \end{bmatrix}$. Hence $b_1 = \frac{220}{3}$ in the notation of Theorem 3.3.7, so

$$\begin{bmatrix} a_k \\ j_k \end{bmatrix} = \mathbf{v}_k \approx b_1 \lambda_1^k \mathbf{x}_1 = \frac{220}{3} 1^k \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

where k is large. Hence $a_k \approx \frac{220}{3}$ and $j_k \approx \frac{440}{3}$ as in Example 3.3.12.

This next example uses Theorem 3.3.7 to solve a “linear recurrence.” See also Section ??.

Example 3.3.14

Suppose a sequence x_0, x_1, x_2, \dots is determined by insisting that

$$x_0 = 1, \quad x_1 = -1, \quad \text{and} \quad x_{k+2} = 2x_k - x_{k+1} \text{ for every } k \geq 0$$

Find a formula for x_k in terms of k .

Solution. Using the linear recurrence $x_{k+2} = 2x_k - x_{k+1}$ repeatedly gives

$$x_2 = 2x_0 - x_1 = 3, \quad x_3 = 2x_1 - x_2 = -5, \quad x_4 = 11, \quad x_5 = -21, \dots$$

so the x_i are determined but no pattern is apparent. The idea is to find $\mathbf{v}_k = \begin{bmatrix} x_k \\ x_{k+1} \end{bmatrix}$ for each k instead, and then retrieve x_k as the top component of \mathbf{v}_k . The reason this works is that the linear recurrence guarantees that these \mathbf{v}_k are a dynamical system:

$$\mathbf{v}_{k+1} = \begin{bmatrix} x_{k+1} \\ x_{k+2} \end{bmatrix} = \begin{bmatrix} x_{k+1} \\ 2x_k - x_{k+1} \end{bmatrix} = A\mathbf{v}_k \text{ where } A = \begin{bmatrix} 0 & 1 \\ 2 & -1 \end{bmatrix}$$

The eigenvalues of A are $\lambda_1 = -2$ and $\lambda_2 = 1$ with eigenvectors $\mathbf{x}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ and $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$,

so the diagonalizing matrix is $P = \begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix}$.

Moreover, $\mathbf{b} = P_0^{-1}\mathbf{v}_0 = \frac{1}{3} \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ so the exact formula for \mathbf{v}_k is

$$\begin{bmatrix} x_k \\ x_{k+1} \end{bmatrix} = \mathbf{v}_k = b_1 \lambda_1^k \mathbf{x}_1 + b_2 \lambda_2^k \mathbf{x}_2 = \frac{2}{3} (-2)^k \begin{bmatrix} 1 \\ -2 \end{bmatrix} + \frac{1}{3} 1^k \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Equating top entries gives the desired formula for x_k :

$$x_k = \frac{1}{3} \left[2(-2)^k + 1 \right] \text{ for all } k = 0, 1, 2, \dots$$

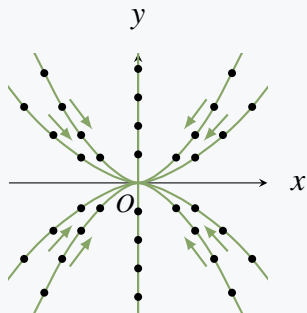
The reader should check this for the first few values of k .

¹²Similar results can be found in other situations. If for example, eigenvalues λ_1 and λ_2 (possibly equal) satisfy $|\lambda_1| = |\lambda_2| > |\lambda_i|$ for all $i > 2$, then we obtain $\mathbf{v}_k \approx b_1 \lambda_1^k \mathbf{x}_1 + b_2 \lambda_2^k \mathbf{x}_2$ for large k .

Graphical Description of Dynamical Systems

If a dynamical system $\mathbf{v}_{k+1} = A\mathbf{v}_k$ is given, the sequence $\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2, \dots$ is called the **trajectory** of the system starting at \mathbf{v}_0 . It is instructive to obtain a graphical plot of the system by writing $\mathbf{v}_k = \begin{bmatrix} x_k \\ y_k \end{bmatrix}$ and plotting the successive values as points in the plane, identifying \mathbf{v}_k with the point (x_k, y_k) in the plane. We give several examples which illustrate properties of dynamical systems. For ease of calculation we assume that the matrix A is simple, usually diagonal.

Example 3.3.15



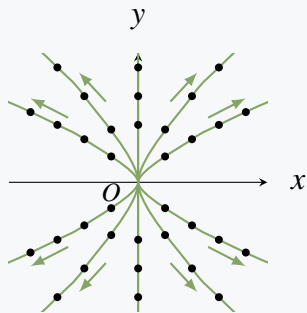
Let $A = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{3} \end{bmatrix}$. Then the eigenvalues are $\frac{1}{2}$ and $\frac{1}{3}$, with corresponding eigenvectors $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

The exact formula is

$$\mathbf{v}_k = b_1 \left(\frac{1}{2}\right)^k \begin{bmatrix} 1 \\ 0 \end{bmatrix} + b_2 \left(\frac{1}{3}\right)^k \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

for $k = 0, 1, 2, \dots$ by Theorem 3.3.7, where the coefficients b_1 and b_2 depend on the initial point \mathbf{v}_0 . Several trajectories are plotted in the diagram and, for each choice of \mathbf{v}_0 , the trajectories converge toward the origin because both eigenvalues are less than 1 in absolute value. For this reason, the origin is called an **attractor** for the system.

Example 3.3.16

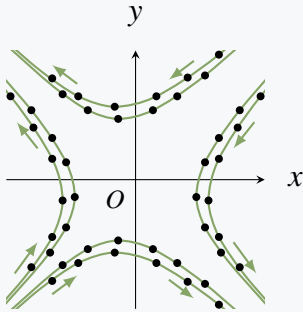


Let $A = \begin{bmatrix} \frac{3}{2} & 0 \\ 0 & \frac{4}{3} \end{bmatrix}$. Here the eigenvalues are $\frac{3}{2}$ and $\frac{4}{3}$, with corresponding eigenvectors $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ as before. The exact formula is

$$\mathbf{v}_k = b_1 \left(\frac{3}{2}\right)^k \begin{bmatrix} 1 \\ 0 \end{bmatrix} + b_2 \left(\frac{4}{3}\right)^k \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

for $k = 0, 1, 2, \dots$. Since both eigenvalues are greater than 1 in absolute value, the trajectories diverge away from the origin for every choice of initial point \mathbf{v}_0 . For this reason, the origin is called a **repellor** for the system.

Example 3.3.17



Let $A = \begin{bmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix}$. Now the eigenvalues are $\frac{3}{2}$ and $\frac{1}{2}$,

with corresponding eigenvectors $\mathbf{x}_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

The exact formula is

$$\mathbf{v}_k = b_1 \left(\frac{3}{2}\right)^k \begin{bmatrix} -1 \\ 1 \end{bmatrix} + b_2 \left(\frac{1}{2}\right)^k \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

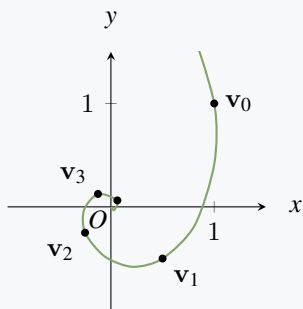
for $k = 0, 1, 2, \dots$. In this case $\frac{3}{2}$ is the dominant eigenvalue so, if $b_1 \neq 0$, we have $\mathbf{v}_k \approx b_1 \left(\frac{3}{2}\right)^k \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ for large k and \mathbf{v}_k is approaching the line $y = -x$.

However, if $b_1 = 0$, then $\mathbf{v}_k = b_2 \left(\frac{1}{2}\right)^k \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and so approaches the origin along the line $y = x$. In general the trajectories appear as in the diagram, and the origin is called a **saddle point** for the dynamical system in this case.

Example 3.3.18

Let $A = \begin{bmatrix} 0 & \frac{1}{2} \\ -\frac{1}{2} & 0 \end{bmatrix}$. Now the characteristic polynomial is $c_A(x) = x^2 + \frac{1}{4}$, so the eigenvalues are the complex numbers $\frac{i}{2}$ and $-\frac{i}{2}$ where $i^2 = -1$. Hence A is not diagonalizable as a real matrix. However, the trajectories are not difficult to describe. If we start with $\mathbf{v}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ then the trajectory begins as

$$\mathbf{v}_1 = \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -\frac{1}{4} \\ -\frac{1}{4} \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -\frac{1}{8} \\ \frac{1}{8} \end{bmatrix}, \mathbf{v}_4 = \begin{bmatrix} \frac{1}{16} \\ \frac{1}{16} \end{bmatrix}, \mathbf{v}_5 = \begin{bmatrix} \frac{1}{32} \\ -\frac{1}{32} \end{bmatrix}, \mathbf{v}_6 = \begin{bmatrix} -\frac{1}{64} \\ -\frac{1}{64} \end{bmatrix}, \dots$$



The first five of these points are plotted in the diagram.

Here each trajectory spirals in toward the origin, so the origin is an attractor. Note that the two (complex) eigenvalues have absolute value less than 1 here. If they had absolute value greater than 1, the trajectories would spiral out from the origin.

Google PageRank

Dominant eigenvalues are useful to the Google search engine for finding information on the Web. If an information query comes in from a client, Google has a sophisticated method of establishing the “relevance” of each site to that query. When the relevant sites have been determined, they are placed in order of importance using a ranking of *all* sites called the PageRank. The relevant sites with the highest PageRank are the ones presented to the client. It is the construction of the PageRank that is our interest here.

The Web contains many links from one site to another. Google interprets a link from site j to site i as a “vote” for the importance of site i . Hence if site i has more links to it than does site j , then i is regarded as more “important” and assigned a higher PageRank. One way to look at this is to view the sites as vertices in a huge directed graph (see Section 2.2). Then if site j links to site i there is an edge from j to i , and hence the (i, j) -entry is a 1 in the associated adjacency matrix (called the *connectivity* matrix in this context). Thus a large number of 1s in row i of this matrix is a measure of the PageRank of site i .¹³

However this does not take into account the PageRank of the sites that link to i . Intuitively, the higher the rank of these sites, the higher the rank of site i . One approach is to compute a dominant eigenvector \mathbf{x} for the connectivity matrix. In most cases the entries of \mathbf{x} can be chosen to be positive with sum 1. Each site corresponds to an entry of \mathbf{x} , so the sum of the entries of sites linking to a given site i is a measure of the rank of site i . In fact, Google chooses the PageRank of a site so that it is proportional to this sum.¹⁴

Exercises for 3.3

Exercise 3.3.1 In each case find the characteristic polynomial, eigenvalues, eigenvectors, and (if possible) an invertible matrix P such that $P^{-1}AP$ is diagonal.

$$\text{a) } A = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$$

$$\text{b) } A = \begin{bmatrix} 2 & -4 \\ -1 & -1 \end{bmatrix}$$

$$\text{c) } A = \begin{bmatrix} 7 & 0 & -4 \\ 0 & 5 & 0 \\ 5 & 0 & -2 \end{bmatrix}$$

$$\text{d) } A = \begin{bmatrix} 1 & 1 & -3 \\ 2 & 0 & 6 \\ 1 & -1 & 5 \end{bmatrix}$$

$$\text{e) } A = \begin{bmatrix} 1 & -2 & 3 \\ 2 & 6 & -6 \\ 1 & 2 & -1 \end{bmatrix}$$

$$\text{f) } A = \begin{bmatrix} 0 & 1 & 0 \\ 3 & 0 & 1 \\ 2 & 0 & 0 \end{bmatrix}$$

$$\text{g) } A = \begin{bmatrix} 3 & 1 & 1 \\ -4 & -2 & -5 \\ 2 & 2 & 5 \end{bmatrix} \quad \text{h) } A = \begin{bmatrix} 2 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & -1 & 2 \end{bmatrix}$$

$$\text{i) } A = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \mu \end{bmatrix}, \lambda \neq \mu$$

$$\text{b. } (x-3)(x+2); 3; -2; \begin{bmatrix} 4 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix};$$

$$P = \begin{bmatrix} 4 & 1 \\ -1 & 1 \end{bmatrix}; P^{-1}AP = \begin{bmatrix} 3 & 0 \\ 0 & -2 \end{bmatrix}.$$

¹³For more on PageRank, visit <https://en.wikipedia.org/wiki/PageRank>.

¹⁴See the articles “Searching the web with eigenvectors” by Herbert S. Wilf, UMAP Journal 23(2), 2002, pages 101–103, and “The worlds largest matrix computation: Google’s PageRank is an eigenvector of a matrix of order 2.7 billion” by Cleve Moler, Matlab News and Notes, October 2002, pages 12–13.

d. $(x-2)^3; 2; \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$; No such P ; Not diagonalizable.

f. $(x+1)^2(x-2); -1, -2; \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$; No such P ; Not diagonalizable. Note that this matrix and the matrix in Example 3.3.9 have the same characteristic polynomial, but that matrix is diagonalizable.

h. $(x-1)^2(x-3); 1, 3; \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ No such P ; Not diagonalizable.

Exercise 3.3.2 Consider a linear dynamical system $\mathbf{v}_{k+1} = A\mathbf{v}_k$ for $k \geq 0$. In each case approximate \mathbf{v}_k using Theorem 3.3.7.

- a. $A = \begin{bmatrix} 2 & 1 \\ 4 & -1 \end{bmatrix}, \mathbf{v}_0 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$
- b. $A = \begin{bmatrix} 3 & -2 \\ 2 & -2 \end{bmatrix}, \mathbf{v}_0 = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$
- c. $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 3 \\ 1 & 4 & 1 \end{bmatrix}, \mathbf{v}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$
- d. $A = \begin{bmatrix} 1 & 3 & 2 \\ -1 & 2 & 1 \\ 4 & -1 & -1 \end{bmatrix}, \mathbf{v}_0 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$

b. $V_k = \frac{7}{3}2^k \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

d. $V_k = \frac{3}{2}3^k \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

Exercise 3.3.3 Show that A has $\lambda = 0$ as an eigenvalue if and only if A is not invertible.

Exercise 3.3.4 Let A denote an $n \times n$ matrix and put $A_1 = A - \alpha I$, α in \mathbb{R} . Show that λ is an eigenvalue of A if and only if $\lambda - \alpha$ is an eigenvalue of A_1 . (Hence, the eigenvalues of A_1 are just those of A

“shifted” by α .) How do the eigenvectors compare?

$A\mathbf{x} = \lambda\mathbf{x}$ if and only if $(A - \alpha I)\mathbf{x} = (\lambda - \alpha)\mathbf{x}$. Same eigenvectors.

Exercise 3.3.5 Show that the eigenvalues of $\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$ are $e^{i\theta}$ and $e^{-i\theta}$. (See Appendix ??)

Exercise 3.3.6 Find the characteristic polynomial of the $n \times n$ identity matrix I . Show that I has exactly one eigenvalue and find the eigenvectors.

Exercise 3.3.7 Given $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ show that:

- a. $c_A(x) = x^2 - \text{tr } A x + \det A$, where $\text{tr } A = a + d$ is called the **trace** of A .
- b. The eigenvalues are $\frac{1}{2} \left[(a+d) \pm \sqrt{(a-d)^2 + 4bc} \right]$.

Exercise 3.3.8 In each case, find $P^{-1}AP$ and then compute A^n .

- a. $A = \begin{bmatrix} 6 & -5 \\ 2 & -1 \end{bmatrix}, P = \begin{bmatrix} 1 & 5 \\ 1 & 2 \end{bmatrix}$
- b. $A = \begin{bmatrix} -7 & -12 \\ 6 & -10 \end{bmatrix}, P = \begin{bmatrix} -3 & 4 \\ 2 & -3 \end{bmatrix}$ [*Hint:* $(PDP^{-1})^n = PD^nP^{-1}$ for each $n = 1, 2, \dots$]

b. $P^{-1}AP = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$, so $A^n = P \begin{bmatrix} 1 & 0 \\ 0 & 2^n \end{bmatrix} P^{-1} = \begin{bmatrix} 9 - 8 \cdot 2^n & 12(1 - 2^n) \\ 6(2^n - 1) & 9 \cdot 2^n - 8 \end{bmatrix}$

Exercise 3.3.9

- a. If $A = \begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$ verify that A and B are diagonalizable, but AB is not.
- b. If $D = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ find a diagonalizable matrix A such that $D+A$ is not diagonalizable.

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

Exercise 3.3.10 If A is an $n \times n$ matrix, show that A is diagonalizable if and only if A^T is diagonalizable.

Exercise 3.3.11 If A is diagonalizable, show that each of the following is also diagonalizable.

- a. $A^n, n \geq 1$
- b. kA, k any scalar.
- c. $p(A), p(x)$ any polynomial (Theorem 3.3.1)
- d. $U^{-1}AU$ for any invertible matrix U .
- e. $kI + A$ for any scalar k .

- b. and d. $PAP^{-1} = D$ is diagonal, then
 - b. $P^{-1}(kA)P = kD$ is diagonal, and d. $Q(U^{-1}AU)Q = D$ where $Q = PU$.

Exercise 3.3.12 Give an example of two diagonalizable matrices A and B whose sum $A + B$ is not diagonalizable.

$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ is not diagonalizable by Example 3.3.8.
 But $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 0 & -1 \end{bmatrix} + \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}$ where $\begin{bmatrix} 2 & 1 \\ 0 & -1 \end{bmatrix}$ has diagonalizing matrix $P = \begin{bmatrix} 1 & -1 \\ 0 & 3 \end{bmatrix}$ and $\begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}$ is already diagonal.

Exercise 3.3.13 If A is diagonalizable and 1 and -1 are the only eigenvalues, show that $A^{-1} = A$.

Exercise 3.3.14 If A is diagonalizable and 0 and 1 are the only eigenvalues, show that $A^2 = A$.

We have $\lambda^2 = \lambda$ for every eigenvalue λ (as $\lambda = 0, 1$) so $D^2 = D$, and so $A^2 = A$ as in Example 3.3.9.

Exercise 3.3.15 If A is diagonalizable and $\lambda \geq 0$ for each eigenvalue of A , show that $A = B^2$ for some matrix B .

Exercise 3.3.16 If $P^{-1}AP$ and $P^{-1}BP$ are both diagonal, show that $AB = BA$. [*Hint*: Diagonal matrices commute.]

Exercise 3.3.17 A square matrix A is called **nilpotent** if $A^n = 0$ for some $n \geq 1$. Find all nilpotent diagonalizable matrices. [*Hint*: Theorem 3.3.1.]

Exercise 3.3.18 Let A be any $n \times n$ matrix and $r \neq 0$ a real number.

- a. Show that the eigenvalues of rA are precisely the numbers $r\lambda$, where λ is an eigenvalue of A .
- b. Show that $c_{rA}(x) = r^n c_A\left(\frac{x}{r}\right)$.

b. $c_{rA}(x) = \det [xI - rA]$
 $= r^n \det \left[\frac{x}{r}I - A \right] = r^n c_A \left[\frac{x}{r} \right]$

Exercise 3.3.19

- a. If all rows of A have the same sum s , show that s is an eigenvalue.
- b. If all columns of A have the same sum s , show that s is an eigenvalue.

Exercise 3.3.20 Let A be an invertible $n \times n$ matrix.

- a. Show that the eigenvalues of A are nonzero.
- b. Show that the eigenvalues of A^{-1} are precisely the numbers $1/\lambda$, where λ is an eigenvalue of A .
- c. Show that $c_{A^{-1}}(x) = \frac{(-x)^n}{\det A} c_A\left(\frac{1}{x}\right)$.

- b. If $\lambda \neq 0, A\mathbf{x} = \lambda\mathbf{x}$ if and only if $A^{-1}\mathbf{x} = \frac{1}{\lambda}\mathbf{x}$. The result follows.

Exercise 3.3.21 Suppose λ is an eigenvalue of a square matrix A with eigenvector $\mathbf{x} \neq \mathbf{0}$.

- a. Show that λ^2 is an eigenvalue of A^2 (with the same \mathbf{x}).
- b. Show that $\lambda^3 - 2\lambda + 3$ is an eigenvalue of $A^3 - 2A + 3I$.
- c. Show that $p(\lambda)$ is an eigenvalue of $p(A)$ for any nonzero polynomial $p(x)$.

b. $(A^3 - 2A - 3I)\mathbf{x} = A^3\mathbf{x} - 2A\mathbf{x} + 3\mathbf{x} = \lambda^3\mathbf{x} - 2\lambda\mathbf{x} + 3\mathbf{x} = (\lambda^3 - 2\lambda - 3)\mathbf{x}$.

Exercise 3.3.22 If A is an $n \times n$ matrix, show that $c_{A^2}(x^2) = (-1)^n c_A(x) c_A(-x)$.

Exercise 3.3.23 An $n \times n$ matrix A is called nilpotent if $A^m = 0$ for some $m \geq 1$.

- Show that every triangular matrix with zeros on the main diagonal is nilpotent.
 - If A is nilpotent, show that $\lambda = 0$ is the only eigenvalue (even complex) of A .
 - Deduce that $c_A(x) = x^n$, if A is $n \times n$ and nilpotent.
-

- If $A^m = 0$ and $A\mathbf{x} = \lambda\mathbf{x}$, $\mathbf{x} \neq \mathbf{0}$, then $A^2\mathbf{x} = A(\lambda\mathbf{x}) = \lambda A\mathbf{x} = \lambda^2\mathbf{x}$. In general, $A^k\mathbf{x} = \lambda^k\mathbf{x}$ for all $k \geq 1$. Hence, $\lambda^m\mathbf{x} = A^m\mathbf{x} = \mathbf{0}\mathbf{x} = \mathbf{0}$, so $\lambda = 0$ (because $\mathbf{x} \neq \mathbf{0}$).

Exercise 3.3.24 Let A be diagonalizable with real eigenvalues and assume that $A^m = I$ for some $m \geq 1$.

- Show that $A^2 = I$.
 - If m is odd, show that $A = I$. [*Hint*: Theorem ??]
-

- If $A\mathbf{x} = \lambda\mathbf{x}$, then $A^k\mathbf{x} = \lambda^k\mathbf{x}$ for each k . Hence $\lambda^m\mathbf{x} = A^m\mathbf{x} = \mathbf{x}$, so $\lambda^m = 1$. As λ is real, $\lambda = \pm 1$ by the Hint. So if $P^{-1}AP = D$ is diagonal, then $D^2 = I$ by Theorem 3.3.4. Hence $A^2 = PD^2P = I$.

Exercise 3.3.25 Let $A^2 = I$, and assume that $A \neq I$ and $A \neq -I$.

- Show that the only eigenvalues of A are $\lambda = 1$ and $\lambda = -1$.

- Show that A is diagonalizable. [*Hint*: Verify that $A(A+I) = A+I$ and $A(A-I) = -(A-I)$, and then look at nonzero columns of $A+I$ and of $A-I$.]

- If $Q_m: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is reflection in the line $y = mx$ where $m \neq 0$, use (b) to show that the matrix of Q_m is diagonalizable for each m .

- Now prove (c) geometrically using Theorem 3.3.3.

Exercise 3.3.26 Let $A = \begin{bmatrix} 2 & 3 & -3 \\ 1 & 0 & -1 \\ 1 & 1 & -2 \end{bmatrix}$ and $B =$

$\begin{bmatrix} 0 & 1 & 0 \\ 3 & 0 & 1 \\ 2 & 0 & 0 \end{bmatrix}$. Show that $c_A(x) = c_B(x) = (x+1)^2(x-2)$, but A is diagonalizable and B is not.

Exercise 3.3.27

- Show that the only diagonalizable matrix A that has only one eigenvalue λ is the scalar matrix $A = \lambda I$.
 - Is $\begin{bmatrix} 3 & -2 \\ 2 & -1 \end{bmatrix}$ diagonalizable?
-

- We have $P^{-1}AP = \lambda I$ by the diagonalization algorithm, so $A = P(\lambda I)P^{-1} = \lambda PP^{-1} = \lambda I$.

- No. $\lambda = 1$ is the only eigenvalue.

Exercise 3.3.28 Characterize the diagonalizable $n \times n$ matrices A such that $A^2 - 3A + 2I = 0$ in terms of their eigenvalues. [*Hint*: Theorem 3.3.1.]

Exercise 3.3.29 Let $A = \begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix}$ where B and C are square matrices.

- If B and C are diagonalizable via Q and R (that is, $Q^{-1}BQ$ and $R^{-1}CR$ are diagonal), show that A is diagonalizable via $\begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix}$

- Use (a) to diagonalize A if $B = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}$ and $C = \begin{bmatrix} 7 & -1 \\ -1 & 7 \end{bmatrix}$.

Exercise 3.3.30 Let $A = \begin{bmatrix} B & 0 \\ 0 & C \end{bmatrix}$ where B and C are square matrices.

- a. Show that $c_A(x) = c_B(x)c_C(x)$.
- b. If \mathbf{x} and \mathbf{y} are eigenvectors of B and C , respectively, show that $\begin{bmatrix} \mathbf{x} \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ \mathbf{y} \end{bmatrix}$ are eigenvectors of A , and show how every eigenvector of A arises from such eigenvectors.

Exercise 3.3.31 Referring to the model in Example 3.3.1, determine if the population stabilizes, becomes extinct, or becomes large in each case. Denote the adult and juvenile survival rates as A and J , and the reproduction rate as R .

	R	A	J
<i>a.</i>	2	$\frac{1}{2}$	$\frac{1}{2}$
<i>b.</i>	3	$\frac{1}{4}$	$\frac{1}{4}$
<i>c.</i>	2	$\frac{1}{4}$	$\frac{1}{3}$
<i>d.</i>	3	$\frac{3}{5}$	$\frac{1}{5}$

b. $\lambda_1 = 1$, stabilizes.

d. $\lambda_1 = \frac{1}{24}(3 + \sqrt{69}) = 1.13$, diverges.

Exercise 3.3.32 In the model of Example 3.3.1, does the final outcome depend on the initial population of adult and juvenile females? Support your answer.

Exercise 3.3.33 In Example 3.3.1, keep the same reproduction rate of 2 and the same adult survival rate of $\frac{1}{2}$, but suppose that the juvenile survival rate is ρ . Determine which values of ρ cause the population to become extinct or to become large.

Exercise 3.3.34 In Example 3.3.1, let the juvenile survival rate be $\frac{2}{5}$ and let the reproduction rate be 2. What values of the adult survival rate α will ensure that the population stabilizes?

Extinct if $\alpha < \frac{1}{5}$, stable if $\alpha = \frac{1}{5}$, diverges if $\alpha > \frac{1}{5}$.

