$$f' = \lambda f$$

for some real λ. This is a first order linear differential equation. All its solutions are given by the formula

$$f(x) = ce^{\lambda x},$$

where c is an arbitrary real constant. Therefore the eigenvectors of D are all exponential functions $f(x) = ce^{\lambda x}$ with $c \neq 0$. The eigenvalue corresponding to $f(x) = ce^{\lambda x}$ is λ. In examples like this one where V is a function space the eigenvectors are called *eigenfunctions.*

EXAMPLE 7. *The integration operator.* Let V be the linear space of all real functions continuous on a finite interval $[a, b]$. If $f \in V$ define $g = T(f)$ to be that function given by

$$g(x) = \int_a^x f(t)\, dt \qquad \text{if} \quad a \leq x \leq b.$$

The eigenfunctions of T (if any exist) are those nonzerofsatisfying an equation of the form

$$\int_a^x f(t)\, dt = \lambda f(x) \tag{4.3}$$

for some real 4. If an eigenfunction exists we may differentiate this equation to obtain the relation $f(x) = \lambda f'(x)$, from which we find $f(x) = ce^{x/\lambda}$, provided $\lambda \neq 0$. In other words, the only candidates for eigenfunctions are those exponential functions of the form $f(x) = ce^{x/\lambda}$ with $c \neq 0$ and $\lambda \neq 0$. However, if we put $x = a$ in (4.3) we obtain

$$0 = \lambda f(a) = Ice”‘“.$$

Since $e^{a/\lambda}$ is never zero we see that the equation $T(f) = \lambda f$ cannot be satisfied with a nonzero f, so Thas no eigenfunctions and no eigenvalues.

EXAMPLE 8. *The subspace spanned by an eigenvector.* Let $T\colon S \to V$ be a linear transformation having an eigenvalue 1. Let x be an eigenvector belonging to λ and let $L(x)$ be the subspace spanned by x. That is, $L(x)$ is the set of all scalar multiples of x. It is easy to show that T maps $L(x)$ into itself. In fact, if $y = cx$ we have

$$T(y) = T(cx) = cT(x) = c(\lambda x) = \lambda(cx) = \lambda y.$$

If $c \neq 0$ then $y \neq 0$ so every nonzero element y of $L(x)$ is also an eigenvector belonging to λ.

A subspace U of S is called *invariant* under Tif Tmaps each element of U onto an element of U. We have j ust shown that the subspace spanned by an eigenvector is invariant under T.

4.3 Linear independence of eigenvectors corresponding to distinct eigenvalues

One of the most important properties of eigenvalues is described in the next theorem. As before, S denotes a subspace of a linear space V.

THEOREM 4.2. *Let $u_1, \ldots, u_k$ be eigenvectors of a linear transformation $T: S \to V$, and assume that the corresponding eigenvalues $\lambda_1, \ldots, \lambda_k$ are distinct. Then the eigenvectors $u_1, \ldots, u_k$ are independent.*

Proof. The proof is by induction on k. The result is trivial when $k = 1$. Assume, then, that it has been proved for every set of $k - 1$ eigenvectors. Let $u_1, \ldots, u_k$ be k eigenvectors belonging to distinct eigenvalues, and assume that scalars c_i exist such that

$$\sum_{i=1}^{k} c_i u_i = O\,. \tag{4.4}$$

Applying T to both members of (4.4) and using the fact that $T(u_i) = \lambda_i u_i$ we find

$$\sum_{i=1}^{k} c_i \lambda_i u_i = O\,. \tag{4.5}$$

Multiplying (4.4) by λ_k and subtracting from (4.5) we obtain the equation

$$\sum_{i=1}^{k-1} c_i(\lambda_i - \lambda_k) u_i = O\,.$$

But since $u_1, \ldots, u_{k-1}$ are independent we must have $c_i(\lambda_i - \lambda_k) = 0$ for each $i = 1, 2, \ldots, k - 1$. Since the eigenvalues are distinct we have $\lambda_i \neq \lambda_k$ for $i \neq k$ so $c_i = 0$ for $i = 1, 2, \ldots, k - 1$. From (4.4) we see that c_k is also 0 so the eigenvectors $u_1, \ldots, u_k$ are independent.

Note that Theorem 4.2 would not be true if the zero element were allowed to be an eigenvector. This is another reason for excluding O as an eigenvector.

> *Warning:* The converse of Theorem 4.2 does not hold. That is, if T has independent eigenvectors $u_1, \ldots, u_k$, then the corresponding eigenvalues $\lambda_1, \ldots, \lambda_k$ need not be distinct. For example, if T is the identity transformation, $T(x) = x$ for all x, then every $x \neq O$ is an eigenvector but there is only one eigenvalue, $\lambda = 1$.

Theorem 4.2. has important consequences for the finite-dimensional case.

THEOREM 4.3. *If $\dim V = n$, every linear transformation $T: V \to V$ has at most n distinct eigenvalues. If T has exactly n distinct eigenvalues, then the corresponding eigenvectors form a basis for V and the matrix of T relative to this basis is a diagonal matrix with the eigenvalues as diagonal entries.*

Proof. If there were $n + 1$ distinct eigenvalues then, by Theorem 4.2, V would contain $n + 1$ independent elements. This is not possible since $\dim V = n$. The second assertion follows from Theorems 4.1 and 4.2.

Note: Theorem 4.3 tells us that the existence of n distinct eigenvalues is a ***sufficient*** condition for Tto have a diagonal matrix representation. This condition is not necessary. There exist linear transformations with less than n distinct eigenvalues that can be represented by diagonal matrices. The identity transformation is an example. All its eigenvalues are equal to 1 but it can be represented by the identity matrix. Theorem 4.1 tells us that the existence of n independent eigenvectors is ***necessary and sufficient*** for T to have a diagonal matrix representation.

4.4 Exercises

1. (a) If T has an eigenvalue λ, prove that aT has the eigenvalue $a\lambda$.
   (b) If x is an eigenvector for both T_1 and T_2, prove that x is also an eigenvector for $aT_1 + bT_2$. How are the eigenvalues related?
2. Assume 'T: $V \to V$ has an eigenvector x belonging to an eigenvalue 1. Prove that x is an eigenvector of T^2 belonging to λ^2 and, more generally, x is an eigenvector of T^n belonging to λ^n. Then use the result of Exercise 1 to show that if P is a polynomial, then x is an eigenvector of $P(T)$ belonging to $P(\lambda)$.
3. Consider the plane as a real linear space, $V = V_2(\mathbf{R})$, and let T be a rotation of V through an angle of $\pi/2$ radians. Although T has no eigenvectors, prove that every nonzero vector is an eigenvector for T^2.
4. If $T: V \to V$ has the property that T^2 has a nonnegative eigenvalue λ^2, prove that at least one of λ or $-\lambda$ is an eigenvalue for T. *[Hint:* $T^2 - \lambda^2 I = (T + \lambda I)(T - \lambda I)$.]
5. Let V be the linear space of all real functions differentiable on (0, 1). If $f \in V$, define $g = T(f)$ to mean that $g(t) = tf'(t)$ for all t in (0, 1). Prove that every real λ is an eigenvalue for T, and determine the eigenfunctions corresponding to A.
6. Let V be the linear space of all real polynomials p(x) of degree $\leq n$. If $p \in V$, define $q = T(p)$ to mean that $q(t) = p(t + 1)$ for all real t. Prove that T has only the eigenvalue 1. What are the eigenfunctions belonging to this eigenvalue?
7. Let V be the linear space of all functions continuous on $(-\infty, +\infty)$ and such that the integral $\int_{-\infty}^{x} f(t)\,dt$ exists for all real x. If $f \in V$, let $g = T(f)$ be defined by the equation g(x) = $\int_{-\infty}^{x} f(t)\,dt$. Prove that every positive I is an eigenvalue for Tand determine the eigenfunctions corresponding to λ.
8. Let V be the linear space of all functions continuous on $(-\infty, +\infty)$ and such that the integral $\int_{-\infty}^{x} t\,f(t)\,dt$ exists for all real x. If $f \in V$ let $g = T(f)$ be defined by the equation $g(x) = \int_{-\infty}^{x} t\,f(t)\,dt$. Prove that every negative λ is aneigenvalue for T and determine the eigenfunctions corresponding to 1.
9. Let $V = C(0, \pi)$ be the real linear space of all real functions continuous on the interval $[0, \pi]$. Let S be the subspace of all functions f which have a continuous second derivative in linear and which also satisfy the boundary conditions $f(0) = f(\pi) = 0$. Let $T: S \to V$ be the linear transformation which maps each f onto its second derivative, $T(f) = f''$. Prove that the eigenvalues of *Tare* the numbers of the form $-n^2$, where $n = 1, 2, \ldots$, and that the eigenfunctions corresponding to $-n^2$ are $f(t) = c_n \sin nt$, where $c_n \neq 0$.
10. Let V be the linear space of all real convergent sequences {x,}. Define $T: V \to V$ as follows: If x = {x,} is a convergent sequence with limit a, let $T(x) = \{y_n\}$, where $y_n = a - x_n$ for $n \geq 1$. Prove that Thas only two eigenvalues, $\lambda = 0$ and $\lambda = -1$, and determine the eigenvectors belonging to each such λ.
11. Assume that a linear transformation T has two eigenvectors x and y belonging to distinct eigenvalues λ and μ. If $ax + by$ is an eigenvector of T, prove that $a = 0$ or $b = 0$.
12. Let $T: S \to V$ be a linear transformation such that every nonzero element of S is an eigenvector. Prove that there exists a scalar c such that $T(x) = cx$. In other words, the only transformation with this property is a scalar times the identity. *[Hint:* Use Exercise 11.]

4.5 The finite-dimensional case. Characteristic polynomials

If dim $V = n$, the problem of finding the eigenvalues of a linear transformation $T: V \to V$ can be solved with the help of determinants. We wish to find those scalars λ such that the equation $T(x) = \lambda x$ has a solution x with x $\neq$ 0. The equation ***T(x)*** = λx can be written in the form

$$(\lambda I - \mathrm{T})(\mathrm{x}) = 0,$$

where I is the identity transformation. If we let $T_\lambda = \lambda I - T$, then λ is an eigenvalue if and only if the equation

$$T,(x) = 0 \tag{4.6}$$

has a nonzero solution x, in which case T_λ is not invertible (because of Theorem 2.10). Therefore, by Theorem 2.20, a nonzero solution of (4.6) exists if and only if the matrix of T_λ is singular. If A is a matrix representation for ***T,*** then λI — A is a matrix representation for T_λ. By Theorem 3.13, the matrix λI — A is singular if and only if det (λI — A) = 0. Thus, if λ is an eigenvalue for ***Tit*** satisfies the equation

$$\det\ (\lambda I - \mathrm{A}) = 0. \tag{4.7}$$

Conversely, any λ ***in the underlying field of scalars*** which satisfies (4.7) is an eigenvalue. This suggests that we should study the determinant det (λI — A) as a function of λ.

THEOREM 4.4. ***If A is any $n \times n$ matrix and if I is the $n \times n$ identity matrix, the function f defined by the equation***

$$f(\lambda) = \det\ (\mathrm{AZ} - \mathrm{A})$$

is a polynomial in λ of degree n. Moreover, the term of highest degree is λ^n, and the constant term is $f(0)$ = det (-A) = (- 1)" det A.

Proof. The statement f(0) = det ***(-A)*** follows at once from the definition of *f.* We prove thatf is a polynomial of degree n only for the case $n \leq 3$. The proof in the general case can be given by induction and is left as an exercise. (See Exercise 9 in Section 4.8.)

For $n = 1$ the determinant is the linear polynomial $f(\lambda) = \lambda - a_{11}$. For $n = 2$ we have

$$\det\ (\lambda I - A) = \begin{vmatrix} \lambda - a_{11} & -a_{12} \\ -a_{21} & \lambda - a_{22} \end{vmatrix} = (\lambda - a_{11})(\lambda - a_{22}) - a_{12}a_{21}$$

$$= \lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}),$$

a quadratic polynomial in λ. For n = ***3 we*** have

$$\det\ (\lambda I - A) = \begin{vmatrix} \lambda - a_{11} & -a_{12} & -a_{13} \\ -a_{21} & I - a_{22} & -a_{23} \\ -a_{31} & -a_{32} & \lambda - a_{33} \end{vmatrix}$$

$$= (\lambda - a_{11})\begin{vmatrix} \lambda - a_{22} & -a_{23} \\ -a_{32} & \lambda - a_{33} \end{vmatrix} + a_{12}\begin{vmatrix} -a_{21} & -a_{23} \\ -a_{31} & \lambda - a_{33} \end{vmatrix} - a_{13}\begin{vmatrix} -a_{21} & \lambda - a_{22} \\ -a_{31} & -a_{32} \end{vmatrix}$$

The last two terms are linear polynomials in λ. The first term is a cubic polynomial, the term of highest degree being λ^3.

DEFINITION. *If A is an n x n matrix the determinant*

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

is called the characteristic polynomial of A.

The roots of the characteristic polynomial of A are complex numbers, some of which may be real. If we let F denote either the real field **R** or the complex field C, we have the following theorem,.

THEOREM 4.5. ***Let $T: V \to V$ be a linear transformation, where V has scalars in F, and*** dim $V = n$. ***Let A be a matrix representation of T. Then the set of eigenvalues of T consists of those roots of the characteristic polynomial of A which lie in F.***

Proof. The discussion preceding Theorem 4.4 shows that every eigenvalue of T satisfies the equation $\det (\lambda I - A) = 0$ and that any root of the characteristic polynomial of A which lies in F is an eigenvalue of T.

The matrix A depends on the choice of basis for V, but the eigenvalues of T were defined without reference to a basis. Therefore, the ***set*** of roots of the characteristic polynomial of A must be independent of the choice of basis. More than this is true. In a later section we shall prove that the characteristic polynomial itself is independent of the choice of basis. We turn now to the problem of actually calculating the eigenvalues and eigenvectors in the finite-dimensional case.

4.6 Calculation of' eigenvalues and eigenvectors in the finite-dimensional case

In the finite-dimensional case the eigenvalues and eigenvectors of a linear transformation Tare also called eigenvalues and eigenvectors of each matrix representation of T. Thus, the eigenvalues of a square matrix A are the roots of the characteristic polynomial $f(\lambda) = \det (\lambda I - A)$. The eigenvectors corresponding to an eigenvalue λ are those nonzero vectors $X = (x_1, \ldots, x_n)$ regarded as n x 1 column matrices satisfying the matrix equation

$$AX = \lambda X, \qquad \text{or} \qquad (\lambda I - A)X = 0.$$

This is a system of n linear equations for the components $x_1, \ldots, x_n$. Once we know λ we can obtain the eigenvectors by solving this system. We illustrate with three examples that exhibit different features.

EXAMPLE 1. *A matrix with all its eigenvalues distinct.* The matrix

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 4 \\ -1 & -1 & -2 \end{bmatrix}$$

has the characteristic polynomial

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

so there are three distinct eigenvalues: 1, -1, and 3. To find the eigenvectors corresponding to $\lambda = 1$ we solve the system $AX = X$, or

$$\begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 4 \\ -1 & -1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

This gives us

$$\begin{aligned} 2x_1 + x_2 + x_3 &= x_1 \\ 2x_1 + 3x_2 + 4x_3 &= x_2 \\ -x_1 - x_2 - 2x_3 &= x_3, \end{aligned}$$

which can be rewritten as

$$\begin{aligned} x_1 + x_2 + x_3 &= 0 \\ 2x_1 + 2x_2 + 4x_3 &= 0 \\ -x_1 - x_2 - 3x_3 &= 0. \end{aligned}$$

Adding the first and third equations we find $x_3 = 0$, and all three equations then reduce to $x_1 + x_2 = 0$. Thus the eigenvectors corresponding to $\lambda = 1$ are $X = t(1, -1, 0)$, where t is any nonzero scalar.

By similar calculations we find the eigenvectors $X = t(0, 1, -1)$ corresponding to $\lambda = -1$, and $X = t(2, 3, -1)$ corresponding to $\lambda = 3$, with t any nonzero scalar. Since the eigenvalues are distinct the corresponding eigenvectors $(1, -1, 0)$, $(0, 1, -1)$, and $(2, 3, -1)$ are independent. The results can be summarized in tabular form as follows. In the third column we have listed the dimension of the eigenspace $E(1)$.

Eigenvalue λ	Eigenvectors	dim $E(\lambda)$
1	$t(1, -1, 0), \quad t \neq 0$	1
-1	$t(0, 1, -1), \quad t \neq 0$	1
3	$t(2, 3, -1), \quad t \neq 0$	1

EXAMPLE 2. *A matrix with repeated eigenvalues.* The matrix

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

has the characteristic polynomial

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

The eigenvalues are 2, 2, and 4. (We list the eigenvalue 2 twice to emphasize that it is a double root of the characteristic polynomial.) To find the eigenvectors corresponding to $\lambda = 2$ we solve the system $AX = 2X$, which reduces to

$$\begin{aligned} -x_2 + x_3 &= 0 \\ x_2 - x_3 &= 0 \\ 2x_1 + x_2 + x_3 &= 0. \end{aligned}$$

This has the solution $x_2 = x_3 = -x_1$ so the eigenvectors corresponding to $\lambda = 2$ are $t(-1, 1, 1)$, where $t \neq 0$. Similarly we find the eigenvectors $t(1, -1, 1)$ corresponding to the eigenvalue: $\lambda = 4$. The results can be summarized as follows:

Eigenvalue	Eigenvectors	dim $E(\lambda)$
2, 2	$t(-1, 1, 1), \quad t \neq 0$	1
4	$t(1, -1, 1), \quad t \neq 0$	1

EXAMPLE 3. *Another matrix with repeated eigenvalues.* The matrix

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{bmatrix}$$

has the characteristic polynomial $(\lambda - 1)(\lambda - 1)(\lambda - 7)$. When $\lambda = 7$ the system $AX = 7X$ becomes

$$\begin{aligned} 5x_1 - x_2 - x_3 &= 0 \\ -2x_1 + 4x_2 - 2x_3 &= 0 \\ -3x_1 - 3x_2 + 3x_3 &= 0. \end{aligned}$$

This has the solution $x_2 = 2x_1$, $x_3 = 3x_1$, so the eigenvectors corresponding to $\lambda = 7$ are $t(1, 2, 3)$, where $t \neq 0$. For the eigenvalue $\lambda = 1$, the system $AX = X$ consists of the equation

$$x_1 + x_2 + x_3 = 0$$

repeated three times. To solve this equation we may take $x_1 = a$, $x_2 = b$, where a and b are arbitrary, and then take $x_3 = -a - b$. Thus every eigenvector corresponding to $\lambda = 1$ has the form

$$(a, b, -a - b) = a(1, 0, -1) + b(0, 1, -1),$$

where $a \neq 0$ or $b \neq 0$. This means that the vectors (1, 0, -1) and (0, 1, -1) form a basis for $E(1)$. Hence dim E(A) = 2 when $\lambda = 1$. The results can be summarized as follows:

Eigenvalue	Eigenvectors	dim $E(\lambda)$
7	$t(1, 2, 3),\quad t \neq 0$	1
1, 1	$a(1, 0, -1) + b(0, 1, -1)$, a, b not both 0.	2

Note that in this example there are three independent eigenvectors but only two distinct eigenvalues.

4.7 Trace of a matrix

Let $f(\lambda)$ be the characteristic polynomial of an $n \times n$ matrix A. We denote the n roots of $f(\lambda)$ by $\lambda_1, \ldots, \lambda_n$, with each root written as often as its multiplicity indicates. Then we have the factorization

$$f(\lambda) = (\lambda - \lambda_1) \cdots (\lambda - \lambda_n).$$

We can also write $f(\lambda)$ in decreasing powers of I as follows,

$$f(\lambda) = A" + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0.$$

Comparing this with the factored form we find that the constant term c_0 and thecoefficient of λ^{n-1} are given by the formulas

$$c_0 = (-1)^n\lambda_1 \cdots \lambda_n \qquad \text{and} \qquad c_{n-1} = -(\lambda_1 + \cdots + \lambda_n).$$

Since we also have c_0 = (-1)" det A, we see that

$$\lambda_1 \cdots 1, = \det A.$$

That is, ***theproduct** of **the roots** of **the characteristicpolynomialof A is equal to the determinant** of A.*

The ***sum*** of the roots off(A) is called the ***trace*** of ***A,*** denoted by tr ***A.*** Thus, by definition,

$$\operatorname{tr} A = \sum_{i=1}^{n} \lambda_i.$$

The coefficient of λ^{n-1} is given by c_{n-1} = -tr ***A.*** We can also compute this coefficient from the determinant form for $f(\lambda)$ and we find that

$$c_{n-1} = -(a_{11} + \cdots + a_{nn}).$$

(A proof of this formula is requested in Exercise 12 of Section 4.8.) The two formulas for c_{n-1} show that

$$\operatorname{tr} A = \sum_{i=1}^{n} a_{ii}\,.$$

That is, *the trace of A is also equal to the sum of the diagonal elements of A.*

Since the sum of the diagonal elements is easy to compute it can be used as a numerical check in calculations of eigenvalues. Further properties of the trace are described in the next set of exercises.

4.8 Exercises

Determine the eigenvalues and eigenvectors of each of the matrices in Exercises 1 through 3. Also, for each eigenvalue λ compute the dimension of the eigenspace $E(\lambda)$.

1. (a) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, (b) $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, (c) $\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$, (d) $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

2. $\begin{bmatrix} 1 & a \\ b & 1 \end{bmatrix}$, $a > 0$, $b > 0$. 3. $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$.

4. The matrices $P_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $P_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$, $P_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ occur in the quantum mechanical theory of electron spin and are called ***Pauli spin matrices***, in honor of the physicist Wolfgang Pauli (1900-1958). Verify that they all have eigenvalues 1 and -1. Then determine all 2 x 2 matrices with complex entries having the two eigenvalues 1 and -1 .
5. Determine all 2 x 2 matrices with real entries whose eigenvalues are (a) real and distinct, (b) real and equal, (c) complex conjugates.
6. Determine a, b, c, d, e, f, given that the vectors $(1, 1, 1)$, $(1, 0, -1)$, and $(1, -1, 0)$ are eigenvectors of the matrix

$$\begin{bmatrix} 1 & 1 & 1 \\ a & b & c \\ d & e & f \end{bmatrix}.$$

7. Calculate the eigenvalues and eigenvectors of each of the following matrices, Also, compute the dimension of the eigenspace $E(\lambda)$ for each eigenvalue λ.

(a) $\begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 4 & -7 & 1 \end{bmatrix}$, (b) $\begin{bmatrix} 2 & 1 & 3 \\ 1 & 2 & 3 \\ 3 & 3 & 20 \end{bmatrix}$, (c) $\begin{bmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{bmatrix}$

8. Calculate the eigenvalues of each of the five matrices

(a) $\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$, (b) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$, (c) $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$,

(d) $\begin{bmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{bmatrix}$, (e) $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$.

These are called *Dirnc matrices* in honor of Paul A. M. Dirac (1902–), the English physicist. They occur in the solution of the relativistic wave equation in quantum mechanics.

9. If A and B are n x n matrices, with B a diagonal matrix, prove (by induction) that the determinant $f(\lambda)$ = det $(\lambda B - A)$ is a polynomial in λ with $f(0)$ = (-1)" det A, and with the coefficient of λ^n equal to the product of the diagonal entries of B.
10. Prove that a square matrix A and its transpose A^t have the same characteristic polynomial.
11. If A and B are n x n matrices, with A nonsingular, prove that AB and BA have the same set of eigenvalues. *Note:* It can be shown that AB and BA have the same characteristic polynomial, even if A is singular, but you are not required to prove this.
12. Let A be an n x n matrix with characteristic polynomial $f(\lambda)$. Prove (by induction) that the coefficient of λ^{n-1} in $f(\lambda)$ is -tr A.
13. Let A and B be n x n matrices with det A = det B and tr A = tr B. Prove that A and B have the same characteristic polynomial if n = 2 but that this need not be the case if n > 2.
14. Prove each of the following statements about the trace.

(a) tr (A + B) = tr A + tr B.
(b) tr (cA) = c tr A.
(c) tr (AB) = tr (BA).
(d) tr A^t = tr A.

4.9 Matrices representing the same linear transformation. Similar matrices

In this section we prove that two different matrix representations of a linear transformation have the same characteristic polynomial. To do this we investigate more closely the relation between matrices which represent the same transformation.

Let us recall how matrix representations are defined. Suppose $T: V \to W$ is a linear mapping of an n-dimensional space V into an m-dimensional space W. Let $(e_1, \ldots, e_n)$ and $(w_1, \ldots, w_m)$ be ordered bases for V and W respectively. The matrix representation of T relative to this choice of bases is the m x n matrix whose columns consist of the components of $T(e_1), \ldots, T(e_n)$ relative to the basis $(w_1, \ldots, w_m)$. Different matrix representations arise from different choices of the bases.

We consider now the case in which $V = W$, and we assume that the same ordered basis $(e_1, \ldots, e_n)$ is used for both V and W. Let A = (a_{ik}) be the matrix of T relative to this basis. This means that we have

$$T(e_k) = \sum_{i=1}^{n} a_{ik}e_i \qquad \text{for} \quad k = 1, 2, \ldots, n. \tag{4.8}$$

Now choose another ordered basis $(u_1, \ldots, u_n)$ for both V and *Wand* let $B = (b_{kj})$ be the matrix of T relative to this new basis. Then we have

(4.9) $$T(u_j) = \sum_{k=1}^{n} b_{kj}u_k \qquad \text{for } j = 1, 2, \ldots, n.$$

Since each u_j is in the space spanned by $e_1, \ldots, e_n$ we can write

(4.10) $$u_j = \sum_{k=1}^{n} c_{kj}e_k \qquad \text{for } j = 1, 2, \ldots, n,$$

for some set of scalars c_{kj}. The $n \times n$ matrix $C = (c_{kj})$ determined by these scalars is non-singular because it represents a linear transformation which maps a basis of V onto another basis of V. Applying T to both members of (4.10) we also have the equations

(4.11) $$T(u_j) = \sum_{k=1}^{n} c_{kj}T(e_k) \qquad \text{for } j = 1, 2, \ldots, n.$$

The systems of equations in (4.8) through (4.11) can be written more simply in matrix form by introducing matrices with vector entries. Let

$$E = [e_1, \ldots, e_n] \qquad \text{and} \qquad U = [u_1, \ldots, u_n]$$

be $1 \times n$ row matrices whose entries are the basis elements in question. Then the set of equations in (4.10) can be written as a single matrix equation,

(4.12) $$U = EC.$$

Similarly, if we introduce

$$E' = [T(e_1), \ldots, T(e_n)] \qquad \text{and} \qquad U' = [T(u_1), \ldots, T(u_n)],$$

Equations (4.8), (4.9), and (4.11) become, respectively,

(4.13) $$E' = EA, \qquad U' = UB, \qquad U' = E'C.$$

From (4.12) we also have

$$E = UC^{-1}.$$

To find the relation between A and B we express U' in two ways in terms of U. From (4.13) we have

$$U' = UB$$

and

$$U' = E'C = EAC = UC^{-1}AC.$$

Therefore $UB = UC^{-1}AC$. But each entry in this matrix equation is a linear combination of the basis vectors $u_1, \ldots, u_n$. Since the u_i are independent we must have

$$B = C^{-1}AC.$$

Thus, we have proved the following theorem.

THEOREM 4.6. *If two $n \times n$ matrices A and B represent the same linear transformation T, then there is a nonsingular matrix C such that*

$$B = C^{-1}AC.$$

Moreover, if A is the matrix of T relative to a basis $E = [e_1, \ldots, e_n]$ and if B is the matrix of T relative to a basis $U = [u_1, \ldots, u_n]$, then for C we can take the nonsingular matrix relating the two bases according to the matrix equation $U = EC$.

The converse of Theorem 4.6 is also true.

THEOREM 4.7. *Let A and B be two $n \times n$ matrices related by an equation of the form $B = C^{-1}AC$, where C is a nonsingular $n \times n$ matrix. Then A and B represent the same linear transformation.*

Proof. Choose a basis $E = [e_1, \ldots, e_n]$ for an n-dimensional space V. Let $u_1, \ldots, u_n$ be the vectors determined by the equations

$$u_j = \sum_{k=1}^{n} c_{kj} e_k \quad \text{for } j = 1, 2, \ldots, n, \tag{4.14}$$

where the scalars c_{kj} are the entries of C. Since C is nonsingular it represents an invertible linear transformation, so $U = [u_1, \ldots, u_n]$ is also a basis for V, and we have $U = EC$.

Let T be the linear transformation having the matrix representation A relative to the basis E, and let S be the transformation having the matrix representation B relative to the basis U. Then we have

$$T(e_k) = \sum_{i=1}^{n} a_{ik} e_i \quad \text{for} \quad k = 1, 2, \ldots, n \tag{4.15}$$

and

$$S(u_j) = \sum_{k=1}^{n} b_{kj} u_k \quad \text{for } j = 1, 2, \ldots, n. \tag{4.16}$$

We shall prove that $S = T$ by showing that $T(u_j) = S(u_j)$ for each j.

Equations (4.15) and (4.16) can be written more simply in matrix form as follows,

$$[T(e_1), \ldots, T(e_n)] = EA, \qquad [S(u_1), \ldots, S(u_n)] = UB.$$

Applying T to (4.14) we also obtain the relation $T(u_j) = \sum c_{kj}T(e_k)$, or

$$[T(u_1), \ldots, T(u_n)] = EAC.$$

But we have

$$UB = ECB = EC(C^{-1}AC) = EAC,$$

which shows that $T(u_j) = S(u_j)$ for each j. Therefore $T(x) = S(x)$ for each x in V, so $T = S$. In other words, the matrices A and B represent the same linear transformation.

DEFINITION. *'Two $n \times n$ matrices A and B are called similar if there is a nonsingular matrix C such that $B = C^{-1}AC$.*

Theorems 4.6 and 4.7 can be combined to give us

THEOREM 4.8. *Two $n \times n$ matrices are similar if and only if they represent the same linear transformation.*

Similar matrices share many properties. For example, they have the same determinant since

$$\det (C^{-1}AC) = \det (C^{-1})(\det A)(\det C) = \det A.$$

This property gives us the following theorem.

THEOREM 4.9. *Similar matrices have the same characteristic polynomial and therefore the same eigenvalues.*

Proof. If A and B are similar there is a nonsingular matrix C such that $B = C^{-1}AC$. Therefore we have

$$\lambda I - B = \lambda I - C^{-1}AC = \lambda C^{-1}IC - C^{-1}AC := C^{-1}(\lambda I - A)C.$$

This shows that $\lambda I - B$ and $\lambda I - A$ are similar, so $\det (\lambda I - B) = \det (\lambda I - A)$.

Theorems 4.8 and 4.9 together show that all matrix representations of a given linear transformation T have the same characteristic polynomial. This polynomial is also called the characteristic polynomial of T.

The next theorem is a combination of Theorems 4.5, 4.2, and 4.6. In Theorem 4.10, F denotes either the real field R or the complex field C.

THEOREM 4.10. ***Let $T: V \to V$ be a linear transformation, where V has scalars in F, and*** $\dim V = n$. ***Assume that the characteristic polynomial of T has n distinct roots I„, . . . , λ_n in F. Then we have:***

(a) ***The corresponding eigenvectors $u_1, \ldots, u_n$ form a basis for V.***

(b) ***The matrix of T relative to the ordered basis U = [$u_1, \ldots,$ u,] is the diagonal matrix A having the eigenvalues as diagonal entries:***

$$\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n).$$

(c) ***If A is the matrix of T relative to another basis E = [$e_1, \ldots,$ e,], then***

$$\Lambda = C^{-1}AC,$$

where C is the nonsingular matrix relating the two bases by the equation

$$U = EC.$$

Proof. By Theorem 4.5 each root λ_i is an eigenvalue. Since there are n distinct roots, Theorem 4.2 tells us that the corresponding eigenvectors $u_1, \ldots, u_n$ are independent. Hence they form a basis for V. This proves (a). Since $T(u_i) = \lambda_i u_i$ the matrix of T relative to U is the diagonal matrix A, which proves (b). To prove (c) we use Theorem 4.6.

> ***Note:*** The nonsingular matrix C in Theorem 4.10 is called a ***diagonalizing matrix.*** If ($e_1, \ldots,$ e,) is the basis of unit coordinate vectors $(I_1, \ldots, I_n)$, then the equation $U = EC$ in Theorem 4.10 shows that the kth column of C consists of the components of the eigenvector u_k relative to $(I_1, \ldots, I_n)$.

If the eigenvalues of A are distinct then A is similar to a diagonal matrix. If the eigenvalues are not distinct then A still might be similar to a diagonal matrix. This will happen if and only if there are k independent eigenvectors corresponding to each eigenvalue of multiplicity k. Examples occur in the next set of exercises.

4.10 Exercises

1. Prove that the matrices $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ and [IO 0 1] have the same eigenvalues but are not similar.

2. In each case find a nonsingular matrix C such that $C^{-1}AC$ is a diagonal matrix or explain why no such C exists.

(a) $A = \begin{bmatrix} 1 & 0 \\ 1 & 3 \end{bmatrix}$, (b) $A = \begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix}$, (c) $A = \begin{bmatrix} 2 & 1 \\ -1 & 4 \end{bmatrix}$, (d) $A = \begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}$.

3. Three bases in the plane are given. With respect to these bases a point has components (x„ x_2), (y_1, y_2), and (z_1, z_2), respectively. Suppose that $[y_1, y_2] = [x_1, x_2]A$, $[z_1, z_2] = [x_1, x_2]B$, and [z„ z_2] $= [y_1, y_2]C$, where A, B, C are 2 x 2 matrices. Express C in terms of A and B.

4. In each case, show that the eigenvalues of A are not distinct but that A has three independent eigenvectors. Find a nonsingular matrix C such that $C^{-1}AC$ is a diagonal matrix.

$$\text{(a) } A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \qquad \text{(b) } A = \begin{bmatrix} 1 & -1 & -1 \\ 1 & 3 & 1 \\ -1 & -1 & 1 \end{bmatrix}.$$

5. Show that none of the following matrices is similar to a diagonal matrix, but that each is similar to a triangular matrix of the form $\begin{bmatrix} \lambda & 0 \\ 1 & \lambda \end{bmatrix}$, where λ is an eigenvalue.

$$\text{(a) } \begin{bmatrix} 2 & -1 \\ 0 & 2 \end{bmatrix}, \qquad \text{(b) } \begin{bmatrix} 2 & 1 \\ -1 & 4 \end{bmatrix}.$$

6. Determine the eigenvalues and eigenvectors of the matrix $\begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 1 \\ -1 & -3 & 3 \end{bmatrix}$ and thereby show that it is not similar to a diagonal matrix.

7. (a) Prove that a. square matrix A is nonsingular if and only if 0 is not an eigenvalue of A.
(b) If A is nonsingular, prove that the eigenvalues of A^{-1} are the reciprocals of the eigenvalues of A.

8. Given an $n \times n$ matrix A with real entries such that $A^2 = -I$. Prove the following statements about A.
(a) A is nonsingular.
(b) n is even.
(c) A has no real eigenvalues.
(d) det A = 1.