Characteristic Vectors

If \( V \) is a finite-dimensional vector space and \( T \) is a linear transformation from \( V \) into \( V \), then, as was shown in Section 1, \( T \) can be represented, in terms of coordinates with respect to a given basis, as multiplication by a matrix \( A \). The choice of a different basis results in a different matrix \( B \), which is similar to \( A \). In the next two sections, we shall discuss the problem of finding, if possible, a basis such that the matrix \( B \) is diagonal. In this section a concept is presented that will be useful in those later discussions.

We say that a vector \( \vec{v} \) is a characteristic vector for \( T \) and that \( \lambda \) is a characteristic value of \( T \) if

\[
T \vec{v} = \lambda \vec{v} \quad \text{and} \quad \vec{v} \neq \vec{0}
\]

In other words, a nonzero vector \( \vec{v} \) is a characteristic vector if \( T \vec{v} \) is a multiple of \( \vec{v} \).

This concept gives a useful formulation of the problem of finding a diagonal matrix representation for \( T \), for

Suppose \( T \) is a linear transformation from \( V \) into \( V \) and that the matrix \( B \) of \( T \) with respect to the basis \( \vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n \) is diagonal. Then each \( \vec{u}_i \) is a characteristic vector for \( T \), and the diagonal entries of \( B \) are the corresponding characteristic values.
To simplify the proof of this result assume that \( n = 2 \). The columns of
the matrix \( B \) are the coordinates of \( T\tilde{u}_1 \) and \( T\tilde{u}_2 \) with respect to the basis
\( \tilde{u}_1, \tilde{u}_2 \). (See formula 6, page 239.) Therefore

\[
B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \quad \text{where} \quad T\tilde{u}_1 = B_{11}\tilde{u}_1 + B_{21}\tilde{u}_2 \quad T\tilde{u}_2 = B_{12}\tilde{u}_1 + B_{22}\tilde{u}_2
\]

Thus, if \( B \) is diagonal, then \( B_{12} = B_{21} = 0 \), and therefore

\[
2 \quad T\tilde{u}_1 = B_{11}\tilde{u}_1 \quad \text{and} \quad T\tilde{u}_2 = B_{22}\tilde{u}_2
\]

The vectors \( \tilde{u}_1 \) and \( \tilde{u}_2 \) are independent and hence nonzero. We therefore
conclude that \( \tilde{u}_1 \) and \( \tilde{u}_2 \) are characteristic vectors, with corresponding
characteristic values \( B_{11} \) and \( B_{22} \), the diagonal entries of \( B \).

The converse of statement 1 is also true, because

If \( V \) has a basis \( \tilde{u}_1, \tilde{u}_2, \ldots, \tilde{u}_n \) consisting of characteristic
vectors for \( T \), the matrix of \( T \) with respect to this basis is
diagonal.

To simplify the proof of this, assume that \( n = 2 \). Suppose \( \tilde{u}_1 \) and \( \tilde{u}_2 \) are
a basis for \( V \) consisting of characteristic vectors for \( T \). Therefore \( T\tilde{u}_1 \) is a
multiple of \( \tilde{u}_1 \) and \( T\tilde{u}_2 \) is a multiple of \( \tilde{u}_2 \), so we can write

\[
T\tilde{u}_1 = a\tilde{u}_1 \quad \text{and} \quad T\tilde{u}_2 = b\tilde{u}_2
\]

Rewrite this in the form

\[
T\tilde{u}_1 = a\tilde{u}_1 + 0\tilde{u}_2 \quad \text{and} \quad T\tilde{u}_2 = 0\tilde{u}_1 + b\tilde{u}_2
\]

Thus the matrix of \( T \) with respect to \( \tilde{u}_1 \) and \( \tilde{u}_2 \) is the diagonal matrix

\[
\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}
\]

Some examples will now be given to illustrate these concepts, after which
a method for finding characteristic vectors and characteristic values will
be discussed.

**Example 1**

Geometric methods can be used to find characteristic vectors and character-
istic values for rotations, reflections, and projections. These make use
of the fact that a nonzero vector \( \tilde{v} \) is a characteristic vector for \( T \) if and
only if \( T\tilde{v} \) is a multiple of \( \tilde{v} \).

Suppose \( T \) is a counterclockwise rotation in \( R^2 \) through the angle \( \theta \) and
that \( \theta \) is not a multiple of \( \pi \). If \( \tilde{v} \neq \tilde{0} \), \( T\tilde{v} \) cannot be a multiple of \( \tilde{v} \), as
shown in Figure 1.
This fact establishes that

A rotation through \( \theta \) has no characteristic vectors, if \( \theta \) is not a multiple of \( \pi \).

Suppose \( \theta = \pi \). Then, as Figure 2 indicates, for every nonzero vector \( \vec{v} \) we have \( T\vec{v} = -\vec{v} \). Since this is also true if \( \theta \) is any odd multiple of \( \pi \), we know that

If \( \theta \) is an odd multiple of \( \pi \), then every nonzero vector is a characteristic vector belonging to the characteristic value \( \lambda = -1 \).

If \( \theta \) is an even multiple of \( \pi \), then \( T \) is just the identity operator. Thus for every nonzero vector \( \vec{v} \) we have \( T\vec{v} = \vec{v} \). In other words,

If \( \theta \) is an even multiple of \( \pi \) then every nonzero vector is a characteristic vector belonging to the characteristic value \( \lambda = 1 \).

**Example 2** Suppose \( T \) is the projection in \( \mathbb{R}^2 \) onto the nonzero vector \( \vec{w} \). A vector parallel to \( \vec{w} \) is left fixed by \( T \); that is, if \( \vec{v} \) is a multiple of \( \vec{w} \), \( T\vec{v} = \vec{v} \). This shows that every nonzero multiple of \( \vec{w} \) is a characteristic vector belonging to the characteristic value \( \lambda = 1 \).

If \( \vec{v} \) is orthogonal to \( \vec{w} \), then \( T\vec{v} = 0 \). Since \( 0 = 0\vec{v} \), this statement shows that every nonzero vector \( \vec{v} \) which is orthogonal to \( \vec{w} \) is a characteristic vector belonging to the characteristic value \( \lambda = 0 \).

If \( \vec{v} \) is neither parallel nor orthogonal to \( \vec{w} \), then, as Figure 3 indicates, \( T\vec{v} \) is not parallel to \( \vec{v} \), and such a vector \( \vec{v} \) cannot be a characteristic vector.

In summary, we have shown that

The projection \( T \) onto \( \vec{w} \) has the characteristic values \( \lambda = 0 \) and \( \lambda = 1 \). The nonzero multiples of \( \vec{w} \) are the characteristic vectors belonging to \( \lambda = 1 \), and the nonzero vectors orthogonal to \( \vec{w} \) are the characteristic vectors belonging to \( \lambda = 0 \).
We extend our terminology to matrices by saying that a vector $\vec{v}$ (written as a column matrix) is a **characteristic vector** for $A$ belonging to the **characteristic value** $\lambda$ if

$$A\vec{v} = \lambda\vec{v} \quad \text{and} \quad \vec{v} \neq \vec{0}$$

Since $I\vec{v} = \vec{v}$, we can rewrite the equation $A\vec{v} = \lambda\vec{v}$ as

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

Therefore, if $\vec{v}$ is a characteristic vector for $A$ belonging to $\lambda$, it follows that $\vec{v}$ is a nonzero solution to $(\lambda I - A)\vec{u} = \vec{0}$; that is, $\vec{v}$ is a nonzero vector in the null space of $\lambda I - A$. Such a $\vec{v}$ can exist only if $\lambda I - A$ is not invertible (Theorem 11, page 188). We therefore have established that

4. If $\vec{v}$ is a characteristic vector for $A$ belonging to $\lambda$, $\lambda I - A$

is not invertible.

If $\lambda I - A$ is not invertible and $\vec{v}$ is a nonzero vector such that $(\lambda I - A)\vec{v} = \vec{0}$, we can then rewrite this equation to conclude that $A\vec{v} = \lambda\vec{v}$ and $\vec{v} \neq \vec{0}$. This shows that

5. If $\lambda I - A$ is not invertible, any nonzero vector $\vec{v}$ in the null space of $\lambda I - A$ is a characteristic vector belonging to the characteristic value $\lambda$.

We know from Theorem 14 that $\lambda I - A$ is not invertible if and only if $\det(\lambda I - A) = 0$. The function

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

is called the **characteristic polynomial** of $A$. In summary:

**Theorem 15**

The characteristic values of $A$ are the roots of the characteristic polynomial $f(\lambda) = \det(\lambda I - A)$. If $\lambda$ is a root of this polynomial, then any nonzero vector in the null space of $\lambda I - A$ is a characteristic vector belonging to $\lambda$.

We have previously noted that $f$ is indeed a polynomial whose degree is the size of $A$. (See Example 4, page 222.)
EXAMPLE 3  Theorem 15 provides a method for finding the characteristic values and corresponding characteristic vectors for a matrix. For example, suppose
\[ A = \begin{bmatrix} 3 & -1 \\ 2 & 0 \end{bmatrix} \]

Form the matrix
\[ \lambda I - A = \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 3 & -1 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} \lambda - 3 & 1 \\ -2 & \lambda \end{bmatrix} \]
and take its determinant to obtain the characteristic polynomial of A:
\[
\begin{align*}
f(\lambda) &= \det(\lambda I - A) \\
&= (\lambda - 3)\lambda + 2 \\
&= \lambda^2 - 3\lambda + 2 \\
&= (\lambda - 2)(\lambda - 1)
\end{align*}
\]
Since the roots of \(f(\lambda)\) are \(\lambda = 2\) and \(\lambda = 1\), Theorem 15 tells us that these are the characteristic values of A. We then find the corresponding characteristic vectors by finding the nonzero vectors in the null spaces of \(2I - A\) and \(1I - A\). We have
\[
\begin{align*}
2I - A &= 2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 3 & -1 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ -2 & 2 \end{bmatrix} \\
1I - A &= 1 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 3 & -1 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ -2 & 1 \end{bmatrix}
\end{align*}
\]
which reduce, respectively, to
\[
\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & 0 \end{bmatrix}
\]
We see that
\[
\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}
\]
are respective bases for the null spaces of \(2I - A\) and \(1I - A\). Therefore,
The characteristic values of A are \(\lambda = 2\) and \(\lambda = 1\).
The nonzero multiples of \(\vec{v}_1\) are the characteristic vectors belonging to \(\lambda = 2\), while the nonzero multiples of \(\vec{v}_2\) are the characteristic vectors belonging to \(\lambda = 1\).
This information will be used in the next section to show that $A$ is similar to
\[
\begin{bmatrix}
2 & 0 \\
0 & 1
\end{bmatrix}
\]

**Example 4** In using Theorem 15 it is necessary to calculate $\det(\lambda I - A)$. This can be done using the definition of determinant. (See formula 2, page 217.) The formulas of Exercise 10, page 230, are, however, worth remembering:

If
\[
A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}
\]

then $\det(\lambda I - A) = \lambda^2 - (a + d)\lambda + ad - bc$

If
\[
A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}
\]

then
\[
\det(\lambda I - A) = \lambda^3 - (A_{11} + A_{22} + A_{33})\lambda^2
\]
\[+ \left( \det \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} + \det \begin{bmatrix} A_{11} & A_{13} \\ A_{31} & A_{33} \end{bmatrix} + \det \begin{bmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{bmatrix} \right) \lambda
\]
\[- \det A
\]

For example, if
\[
A = \begin{bmatrix}
8 & 9 & 9 \\
3 & 2 & 3 \\
-9 & -9 & -10
\end{bmatrix}
\]

formula 7 gives
\[
\det(\lambda I - A) = \lambda^3 - (8 + 2 - 10)\lambda^2
\]
\[+ \left( \det \begin{bmatrix} 8 & 9 \\ 3 & 2 \end{bmatrix} + \det \begin{bmatrix} 8 & 9 \\ -9 & -10 \end{bmatrix} + \det \begin{bmatrix} 2 & 3 \\ -9 & -10 \end{bmatrix} \right) \lambda
\]
\[- \det A
\]

Calculation of these determinants gives
\[
\det(\lambda I - A) = \lambda^3 - 3\lambda - 2 = (\lambda + 1)^2(\lambda - 2)
\]
Theorem 15 then tells us that the characteristic values of $A$ are $\lambda = -1$ and $\lambda = 2$. Corresponding characteristic vectors are found by finding the nonzero vectors in the null spaces of $-1I - A$ and $2I - A$.

**Example 5** The characteristic values of an upper (or lower) triangular matrix are easy to find. For example, if

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

then

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

The determinant of an upper triangular matrix is the product of the diagonal entries. (See property 4c, page 218.) Therefore

$$f(\lambda) = (\lambda - 2)(\lambda - 2)(\lambda + 3)(\lambda - 4)$$

so that the characteristic values are 2, -3, and 4.

In general,

8 The characteristic values of an upper (or lower) triangular matrix are the diagonal entries of the matrix.

**Exercises**

1. For each of the following operators $T$ describe the characteristic values and vectors. Figures may be helpful.
   a. $T$ is reflection in $R^2$ through the line through $\vec{w}$.
   b. $T$ is projection in $R^2$ orthogonal to $\vec{w}$.
   c. $T$ is reflection in $R^2$ through the line through $\vec{w}$.
   d. $T$ is projection in $R^3$ onto $\vec{w}$.
   e. $T$ is projection in $R^3$ orthogonal to $\vec{w}$.
   f. $T$ is counterclockwise rotation in $R^2$ through $\pi/4$ followed by reflection in the $x$-axis.

2. Suppose $Df = f'$. Show that every real number is a characteristic value of $D$. [Hint: Calculate $D(e^{\alpha t})$.]
3. Show that \( \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \) and \( \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \) are characteristic vectors for \( A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \).

What are the corresponding characteristic values?

4. Use Theorem 15 and formulas 6, 7, or 8 to find the characteristic polynomial and the characteristic values for each of the following matrices.

\[
\begin{align*}
a & = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \\
b & = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \\
c & = \begin{bmatrix} 3 & 1 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix} \\
d & = \begin{bmatrix} 5 & 1 & 1 \\ -3 & 1 & -3 \\ -2 & -2 & 2 \end{bmatrix} \\
e & = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix} \\
f & = \begin{bmatrix} 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}
\end{align*}
\]

5. For each of the matrices of Exercise 4 find at least one characteristic vector for each characteristic value.

6. Show that \( \begin{bmatrix} 1 & -2 \\ 1 & -1 \end{bmatrix} \) has no real characteristic values. (See also Exercise 11.)

7. a. Show that if \( A\vec{v} = \lambda \vec{v} \), then \( A^2\vec{v} = \lambda^2 \vec{v} \).

b. Suppose \( \lambda_1, \lambda_2, \ldots, \lambda_k \) are the characteristic values of \( A \). What are the characteristic values of \( A^2 \)?

c. Suppose \( A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{bmatrix} \)

What are the characteristic values of \( A^2 \) of \( A^3 \)?

8. a. Show that if \( A\vec{v} = \lambda \vec{v} \), then

\[
(A^3 - 3A^2 + A - 2I)\vec{v} = (\lambda^3 - 3\lambda^2 + \lambda - 2)\vec{v}
\]

b. Suppose \( q(\lambda) \) is a polynomial and \( \lambda_1, \lambda_2, \ldots, \lambda_k \) are the characteristic values of \( A \). What are the characteristic values of \( q(A) \)? (Hint: See part a.)

9. a. Show that similar matrices have the same characteristic polynomial.

[Hint: \( \lambda I - P^{-1}AP = P^{-1}(\lambda I - A)P \).]

b. How would you define the characteristic polynomial of a linear transformation \( T \) from \( V \) into \( V \)? Does your definition depend upon the choice of basis for \( V \)?
10 Suppose $A$ is similar to
\[
\begin{bmatrix}
2 & 0 & 0 \\
0 & -1 & 0 \\
0 & 0 & 3
\end{bmatrix}
\]
What are the characteristic values of $A$? (Hint: See Exercise 9a.)

□ 11 Show that \[
\begin{bmatrix}
2 \\
1 - i \\
1 + i
\end{bmatrix}
\]
and \[
\begin{bmatrix}
2 \\
1 - 2 \\
1 - 1
\end{bmatrix}
\]
are complex characteristic vectors for \[
\begin{bmatrix}
1 & -2 \\
1 & -1
\end{bmatrix}
\]. What are the corresponding characteristic values? Compare this with Exercise 6.

□ 12 Find the characteristic polynomial, the complex characteristic values, and at least one corresponding characteristic vector for each of the following.
\[
\begin{align*}
a & \quad \begin{bmatrix}
2 - i & 2 \\
1 + 3i & -1 + 3i
\end{bmatrix} \\
b & \quad \begin{bmatrix}
0 & 1 \\
1 & 0
\end{bmatrix}
\end{align*}
\]

□ 13 An important theorem, known as the Cayley-Hamilton Theorem, asserts that if $f(\lambda)$ is the characteristic polynomial of $A$, $f(A)$ is the zero matrix. Verify that this is so for
\[
A = \begin{bmatrix}
2 & -1 \\
1 & 3
\end{bmatrix}
\]