3.4 Eigenvectors

Definition 3.4.1. *Let* $T : V \to V$ *be a linear transformation, where* V *is a finite dimensional vector space. A non-zero element* v *of* V *is a* eigenvector *of* T *if* $T(v)$ *is a scalar multiple of* v*.*

If v is an eigenvector of T, then the 1-dimensional subspace of V spanned by v, which consists of all scalar multiples of v, is mapped to itself by T. It is said to be a T-*invariant* line.

Definition 3.4.2. *If* v *is an eigenvector of* T*, then* $T(v) = \lambda v$ *for some scalar* λ*, and* λ *is called the* eigenvalue *of* T *to which* v *corresponds.*

Here is the matrix version.

Definition 3.4.3. *Let* $A \in M_n(\mathbb{F})$*. A vector* $v \in \mathbb{F}^n$ *is an* eigenvector *of* A *if* $Av = \lambda v$ *for some scalar* $\lambda \in \mathbb{F}$*, called the* eigenvalue *of* A *to which* v *corresponds.*

Given a matrix A and a vector v, it is quite a straightforward task to determine whether v is an eigenvector of A, and to determine the corresponding eigenvalue if so - just calculate the matrix-vector product Av and see if it is a scalar multiple of v. In fact, given a vector v, we can construct a matrix that has v as an eigenvector, with our favourite scalar as an eigenvalue.

Example 3.4.4. *Find a matrix* $A \in M_3(\mathbb{R})$ *that has* $v = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ *as an eigenvector, corresponding to the eigenvalue 28.*

To do this, write u_1, u_2, u_3 as the three rows of A. What we need is that $u_1 v = 28(1) = 28$, $u_2 v = 28(2) = 56$, $u_3 v = 28(3) = 84$. The easy way to arrange this is to choose $u_1 = (28\ \ 0\ \ 0)$, $u_2 = (0\ \ 28\ \ 0)$, $u_3 = (0\ \ 0\ \ 28)$, so that $A = 28I_3$. This answer is correct but we can find others, and the conditions on u_1, u_2, u_3 are independent. For example we can choose

$$u_1 = (3\ \ 2\ \ 3),\ \ u_2 = (0\ \ -2\ \ 20),\ \ u_3 = (5\ \ 2\ \ 25)$$

to get $A = \begin{bmatrix} 3 & 2 & 3 \\ 0 & -2 & 20 \\ 5 & 2 & 25 \end{bmatrix}$, and it is easily confirmed that $Av = 28v$.

Exercise: Show that the set of matrices in $M_3(\mathbb{R})$ that satisfy $Mv = 28v$ is a *subspace* of $M_3(\mathbb{R})$.

Example 3.4.5. Show that $\begin{bmatrix} 3 \\ 4 \end{bmatrix}$ is an eigenvector of $\begin{bmatrix} -2 & 9 \\ 8 & 4 \end{bmatrix}$ and find the corresponding eigenvalue.

$$\begin{bmatrix} -2 & 9 \\ 8 & 4 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 30 \\ 40 \end{bmatrix} = 10 \begin{bmatrix} 3 \\ 4 \end{bmatrix}.$$

The corresponding eigenvalue is 10.

Diagonal matrices.

A harder problem is to find the eigenvectors of a matrix or linear trasformation, given only the matrix or linear transformation itself. For example, suppose that

$$B = \begin{bmatrix} 5 & 6 & 2 \\ 0 & -1 & -8 \\ 1 & 0 & -2 \end{bmatrix}.$$

Finding an eigenvector of B means finding solutions for x, y, z and λ, to the following system of equations, where the values of x, y, z are not all zero.

$$\begin{bmatrix} 5 & 6 & 2 \\ 0 & -1 & -8 \\ 1 & 0 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

If λ is regarded as a variable, this is not a system of linear equations. Where to begin?

It turns out that the key to making progress is to find the *eigenvalues* first, even if it's the eigenvectors that we want. To see why, we show that the number of distinct eigenvalues of a $n \times n$ matrix cannot exceed n.

Theorem 3.4.6. *Let $A \in M_n(\mathbb{F})$ and let $v_1, \dots, v_k$ be eigenvectors of A in $\mathbb{F}^n$, corresponding to* distinct *eigenvalues $\lambda_1, \dots, \lambda_k$ of A. Then $\{v_1, \dots, v_k\}$ is a linearly independent subset of $\mathbb{F}^n$.*

Idea of Proof: First suppose that $k = 2$, and suppose that $a_1v_1 + a_2v_2 = 0$, for scalars a and b in $\mathbb{F}$. We need to show that $a_1 = a_2 = 0$. Multiplying the expression $a_1v_1 + a_2v_2$ on the left by A, we have

$$a_1Av_1 + a_1Av_2 = 0 \Longrightarrow a_1\lambda_1v_1 + a_2\lambda_2v_2 = 0.$$

Multiplying the same expression by the scalar λ_1 gives

$$a_1\lambda_1v_1 + a_2\lambda_1v_2 = 0.$$

Subtracting one of these expressions from the other gives

$$a_2(\lambda_1 - \lambda_2)v_2 = 0.$$

Now v_2 is not the zero vector because it is an eigenvector of A, and $\lambda_1 - \lambda_2$ is not zero, because λ_1 and λ_2 are distinct eigenvalues. So it must be that $a_2 = 0$. Since $a_1v_1 + a_2v_2 = 0$, it follows that $a_1 = 0$ also, since v_1 is not the zero vector (begin an eigenvector of A). We conclude that the zero vector can be written as a linear combination of v_1 and v_2 only if both coefficients are zero, which means that $\{v_1, v_2\}$ is a linearly independent set.

The proof in the general situation uses exactly this idea.

Proof. If $\{v_1, \dots, v_k\}$ is linearly dependent, then there are expressions for the zero vector as a linear combination of $v_1, \dots, v_k$ in which the coefficients are *not* all zero. Let d be the least number of non-zero coefficients in any such expression, and (after reordering the v_i and λ_i if necessary), suppose that

$$a_1v_1 + \cdots + a_dv_d = 0,$$

with $d \geqslant 2$ and each a_i is a non-zero element of $\mathbb{F}$. Multiplying this equation respectively by A (on the left) and by λ_1 gives

$$\begin{aligned} a_1\lambda_1v_1 + a_2\lambda_2v_2 + \cdots + a_d\lambda_dv_d &= 0 \\ a_1\lambda_1v_1 + a_2\lambda_1v_2 + \cdots + a_d\lambda_1v_d &= 0 \end{aligned}$$

Subtracting the second equation from the first gives

$$a_2(\lambda_2 - \lambda_1)v_1 + a_3(\lambda_3 - \lambda_1)v_2 + \cdots + a_d(\lambda_d - \lambda_1)v_d = 0.$$

None of the coefficients in this linear combination of $v_2, \dots, v_d$ are zero, since the a_i are all non-zero and the λ_i are all distinct. So this is a non-trivial expression for the zero vector as a linear combination of $v_1, \dots, v_k$ with fewer than d non-zero coefficients, which contradicts the choice of d. We conclude that $\{v_1, \dots, v_k\}$ is a linearly independent subset of $\mathbb{F}^n$. □

The following consequence of Theorem 3.4.6 suggests that we may have some chance of being able to find the eigenvalues of a $n \times n$ matrix, or at least that there are not too many of them.

Corollary 3.4.7. *Let $A \in M_n(\mathbb{F})$. Then A has at most n distinct eigenvalues in $\mathbb{F}$.*

Proof. If A has k distinct eigenvalues, with corresponding eigenvectors $v_1, \dots, v_k$ in $\mathbb{F}^n$, then k cannot exceed the dimension of $\mathbb{F}^n$, since $\{v_1, \dots, v_k\}$ is a linearly indepedent set in $\mathbb{F}^n$. Hence $k \leqslant n$. □

The following consequence is also important and useful.

Corollary 3.4.8. *Let $A \in M_n(\mathbb{F})$ and suppose that A has n distinct eigenvalues $\lambda_1, \dots, \lambda_n$ in $\mathbb{F}$. Then A is diagonalizable, and A is similar to the matrix* $\mathrm{diag}(\lambda_1, \dots, \lambda_n)$.

Proof. Let $v_1, \dots, v_n$ be eigenvectors of A in $\mathbb{F}^m$, corresponding to $\lambda_1, \dots, \lambda_n$ respectively. Then $(v_1, \dots, v_n)$ is an (ordered) basis of $\mathbb{F}^n$, by Theorem 3.4.6. If P is the matrix with columns $v_1, \dots, v_n$, then $P^{-1}AP = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$. □

If a $n \times n$ matrix has fewer than n distinct eigenvalues, then it may or may not be diagonalizable. The two examples below indicate two ways in which a matrix in $M_n(\mathbb{F})$ could fail to be diagonalizable in $M_n(\mathbb{F})$.

Example 3.4.9. Let $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ in $M_2(\mathbb{R})$.

Suppose that $\begin{bmatrix} x \\ y \end{bmatrix}$ is an eigenvector of A. Then

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \end{bmatrix} \Longrightarrow \begin{array}{rcl} x + y & = & \lambda x \\ y & = & \lambda y \end{array}$$

The second equation says that $y = 0$ or $\lambda = 1$. If $y = 0$, then the first equation says $x = \lambda x$. Since x and y cannot both be 0 in an eigenvector, it follows that $\lambda = 1$ anyway. Thus $\lambda = 1$ is the *only* possible eigenvalue of A. The non-zero vector $\begin{bmatrix} x \\ y \end{bmatrix}$ is an eigenvector of A corresponding to $\lambda = 1$ if and only if $x + y = x$ and $y = y$. The first equation says $y = 0$, and x may have any value. The eigenvectors of A are all vectors of the form $\begin{bmatrix} x \\ 0 \end{bmatrix}$, where $x \neq 0$ in $\mathbb{R}$, i.e. all scalar multiples of $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. These comprise only a 1-dimensional subspace of $\mathbb{R}^2$, os $\mathbb{R}^2$ does not have a basis consisting of eigenvectors of A, and A is not diagonalizable.

The point of the following example is to show that if A is a matrix in $M_n(\mathbb{F})$, the eigenvalues of A may not be in $\mathbb{F}$ but in a bigger field.

Example 3.4.10. Let $B = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ in $M_2(\mathbb{R})$.

Note that B is the matrix of a counter-clockwise rotation through $\frac{pi}{2}$ about the origin in $\mathbb{R}^2$. From that geometric intepretation we can see that B has no eigenvector in $\mathbb{R}^2$, since no line in $\mathbb{R}^2$ is preserved by this rotation. We can also see this algebraically.

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \end{bmatrix} \Longrightarrow \begin{array}{rcl} y & = & \lambda x \\ -x & = & \lambda y \end{array}.$$

Looking at both of these equations, we have $y = \lambda x = \lambda(-\lambda y) \longrightarrow y = -\lambda^2 y$.

If $y = 0$, then $x = 0$ which does not give an eigenvector. If $y \neq 0$, then $y = -\lambda^2 y$ means $\lambda^2 = -1$, which is not satisfied by any real number λ. This means that B has no real eigenvalue and no eigenvector in $\mathbb{R}^2$. However, if we allow complex values for λ, then $\lambda = i$ and $\lambda = -i$ satisfy $\lambda^2 = -1$. To find corresponding eigenvectors:

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = i \begin{bmatrix} x \\ y \end{bmatrix} \Longrightarrow \begin{array}{rcl} y & = & ix \\ -x & = & iy \end{array}$$

So $\begin{bmatrix} 1 \\ i \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue i.

For the eigenvalue $-i$:

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = -i \begin{bmatrix} x \\ y \end{bmatrix} \Longrightarrow \begin{array}{rcl} y & = & -ix \\ -x & = & -iy \end{array}$$

So $\begin{bmatrix} 1 \\ -i \end{bmatrix}$ is an eigenvector corresponding to the eigenvalue i.

We conclude that B is not diagonalizable in $M_2(\mathbb{R})$ but that it is similar in $M_2(\mathbb{C})$ to the diagonal matrix $\begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$.

Learning Outcomes for Section 3.4

1. To define an eigenvector of a linear transformation or of a square matrix.
2. To know that eigenvectors corresponding to different eigenvalues are linearly independent.
3. And that this means a $n \times n$ matrix can have at most n distinct eigenvalues
4. and that it is diagonalizable if it does have n distinct eigenvalues.