

Chapter 10

Eigenvectors and Eigenvalues

10.1 Definitions

Let A be an $n \times n$ matrix. An *eigenvector* of A is a vector $\mathbf{v} \in \mathbb{R}^n$, with $\mathbf{v} \neq \mathbf{0}_n$, such that $A\mathbf{v} = \lambda\mathbf{v}$ for some scalar $\lambda \in \mathbb{R}$ (which might be 0). The scalar λ is called the *eigenvalue* of A associated to the eigenvector $\mathbf{v}$.

Example. Let $A = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$. The vector $\mathbf{v} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ is an eigenvector of A , with corresponding eigenvalue 5, since

$$\begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 5 \\ 10 \end{pmatrix} = 5 \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Geometrically, it is clear that the eigenvectors of the linear transformation $t_A : \mathbf{x} \mapsto A\mathbf{x}$ are the position vectors of points on fixed lines through the origin (except for the origin itself), and the eigenvalues are the corresponding stretch factors, at least in the case of eigenvalues $\lambda \neq 0$.

10.2 Eigenvectors corresponding to the eigenvalue 0

From the definitions, saying that 0 is an eigenvalue of A means the same as saying that there exists a nonzero vector $\mathbf{v}$ such that $A\mathbf{v} = \mathbf{0}_n$. But this implies that A is not invertible (since if A is invertible then $A\mathbf{x} = \mathbf{0}_n$ has *unique* solution $\mathbf{x} = A^{-1}\mathbf{0}_n = \mathbf{0}_n$). In fact the converse is also true, namely that if A is not invertible then $A\mathbf{x} = \mathbf{0}_n$ has a nonzero solution $\mathbf{x}$ (we are not going to prove this here, but we saw it was true in the case $n = 2$, when we examined *singular* 2×2 matrices). Thus we have the

Useful Fact. An $n \times n$ matrix has 0 as one of its eigenvalues if and only if A is not invertible.

10.3 Finding all eigenvalues

Given a matrix, it turns out to be easiest to first calculate the eigenvalues, and then the eigenvectors.

Theorem 10.1. Let A be an $n \times n$ matrix. Then the eigenvalues of A are those real numbers λ which have the property that $\det(A - \lambda I_n) = 0$.

Proof. Strictly speaking, we have only proved all properties of determinants (namely that A is invertible if and only if $\det A \neq 0$) we need for $n \leq 2$, and stated that the same holds for $n = 3$, and for general n if you read the optional section (§8.4) of Chapter 8. So strictly speaking we will only have proved this theorem for $n \leq 2$, though the same argument works for general n , as you will see in MTH5112: Linear Algebra I:

$$\begin{aligned} \lambda \text{ is an eigenvalue of } A &\iff \exists \mathbf{v} \neq \mathbf{0}_n \text{ such that } A\mathbf{v} - \lambda\mathbf{v} = (A - \lambda I_n)\mathbf{v} = \mathbf{0}_n \\ &\iff (A - \lambda I_n) \text{ is not invertible} \\ &\iff \det(A - \lambda I_n) = 0. \end{aligned} \quad \square$$

The expression $\det(A - xI_n)$ is called the *characteristic polynomial* of A (it is a polynomial in x , of degree n), so another way to state the Theorem above is to say that *the eigenvalues of A are the zeros of the characteristic polynomial of A* . (Recall that the *zeros* of a function $f(x)$ are the values of x such that $f(x) = 0$.) Eigenvalues are usually counted by multiplicity, so that if the characteristic polynomial is, say, $-(x - 4)^2(x + 3)$ we would say that the eigenvalues are $-3, 4, 4$.

Example 1. Let $A = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$. Then the characteristic polynomial of A is

$$\begin{aligned} f(x) &= \det(A - xI_2) = \det\left(\begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix} - \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix}\right) = \det\begin{pmatrix} 1-x & 2 \\ 4 & 3-x \end{pmatrix} \\ &= (1-x)(3-x) - 8 = x^2 - 4x - 5 = (x+1)(x-5). \end{aligned}$$

The only zeros of $f(x)$ are -1 and 5 , so by Theorem 10.1 the only eigenvalues of A are -1 and 5 .

Example 2. Let $A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 4 \\ 0 & 0 & 3 \end{pmatrix}$. Then the characteristic polynomial of A is

$$\begin{aligned} f(x) &= \det(A - xI_3) = \begin{vmatrix} 1-x & 2 & -1 \\ 0 & 1-x & 4 \\ 0 & 0 & 3-x \end{vmatrix} = (1-x) \begin{vmatrix} 1-x & 4 \\ 0 & 3-x \end{vmatrix} - 0 + 0 \\ &= (1-x)((1-x)(3-x) - 0) = -(x-1)^2(x-3). \end{aligned}$$

The only zeros of $f(x)$ are 1 and 3 , so by Theorem 10.1 the only distinct eigenvalues of A are 1 and 3 . In this example, the characteristic polynomial has a repeated root, and

the eigenvalues of A are actually 1, 1, 3. This does have some mathematical significance. Using language that will be defined in MTH5112: Linear Algebra I, this means that the eigenvectors having eigenvalue 1, together with the zero vector, could potentially form a 2-dimensional subspace of $\mathbb{R}^3$, though in this example they only form a 1-dimensional subspace.

Remarks.

1. Let $n = 2$ or 3 , and let A be an $n \times n$ matrix, with characteristic polynomial $f(x)$. It is not difficult to see that $f(x)$ is a polynomial with coefficients in $\mathbb{R}$, and that the degree of $f(x)$ is n (the highest degree term in $f(x)$ comes from the product of the terms down the main diagonal of A , so the coefficient of x^n is ± 1 [$(-1)^n$ in fact]). Such an $f(x)$ has at most n real zeros, and so A has at most n distinct real eigenvalues.
2. A polynomial of odd degree with coefficients in $\mathbb{R}$ always has at least one real zero. (This follows from the Intermediate Value Theorem [see your MTH4100: Calculus I notes], since the graph of $y = f(x)$ is either above the x -axis as x tends to $+\infty$ and below the x -axis as x tends to $-\infty$, or vice versa.) Hence it follows that every 3×3 (real) matrix has at least one (real) eigenvalue, and so every linear transformation of $\mathbb{R}^3$ (except for zero transformation) fixes at least one line through the origin.

10.4 Finding eigenvectors

Let $n = 2$ or 3 , and let A be an $n \times n$ matrix with an eigenvalue λ . How do we find one (or all) eigenvectors $\mathbf{v}$ of A with $A\mathbf{v} = \lambda\mathbf{v}$? The answer is that we solve a system of n equations in n unknowns. We have:

$$\begin{aligned} A\mathbf{v} = \lambda\mathbf{v} &\iff A\mathbf{v} - \lambda\mathbf{v} = \mathbf{0}_n \\ &\iff A\mathbf{v} - \lambda I_n \mathbf{v} = \mathbf{0}_n \\ &\iff (A - \lambda I_n)\mathbf{v} = \mathbf{0}_n. \end{aligned}$$

If $n = 2$, to obtain the eigenvectors of A corresponding to the eigenvalue λ we solve $(A - \lambda I_2) \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ for $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ with $\begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

If $n = 3$, to obtain the eigenvectors of A corresponding to the eigenvalue λ we solve $(A - \lambda I_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ for $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$ with $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$.

Example 1. Let

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

In the previous subsection we found that the eigenvalues of A are -1 and 5 . We now determine the eigenvectors corresponding to eigenvalue -1 . We solve:

$$(A - (-1)I_2) \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

that is to say,

$$\left(\begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix} - \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \right) \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

in other words,

$$\begin{pmatrix} 2 & 2 \\ 4 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This is the system of equations:

$$\left. \begin{array}{l} 2x + 2y = 0 \\ 4x + 4y = 0 \end{array} \right\},$$

which, when reduced to echelon form is:

$$\left. \begin{array}{l} 2x + 2y = 0 \\ 0 = 0 \end{array} \right\}.$$

Thus y can be any real number r , and then $2x + 2r = 0$, so $x = -r$. Thus the set of all eigenvectors of A corresponding to the eigenvalue -1 is

$$\left\{ \begin{pmatrix} -r \\ r \end{pmatrix} : r \in \mathbb{R} \mid r \neq 0 \right\}.$$

One such eigenvector is $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$. We can check our calculation, as shown below.

$$A \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} = (-1) \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

It is left as an exercise for you to compute the eigenvectors of A corresponding to the other eigenvalue, 5 .

Example 2. Let

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

We have found that the characteristic polynomial of A is $f(x) = -(x-1)^2(x-3)$, and thus that the eigenvalues of A are 1 and 3 (the real zeros of $f(x)$), with 1 being repeated. We now find all the eigenvectors of A corresponding to the eigenvalue 3 . We solve:

$$(A - 3I_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

that is to say,

$$\begin{pmatrix} 1-3 & 2 & -1 \\ 0 & 1-3 & 4 \\ 0 & 0 & 3-3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

which, in other words, is

$$\begin{pmatrix} -2 & 2 & -1 \\ 0 & -2 & 4 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

This is the system of equations:

$$\left. \begin{aligned} -2x + 2y - z &= 0 \\ -2y + 4z &= 0 \\ 0 &= 0 \end{aligned} \right\}.$$

These equations are already in echelon form, so we can solve them by setting z to be r (representing any real number) and deducing by back substitution that $y = 2r$ (from the second equation) and then that $x = \frac{3}{2}r$ from the first equation. Thus the set of all eigenvectors of A corresponding to the eigenvalue 3 is

$$\left\{ \begin{pmatrix} \frac{3}{2}r \\ 2r \\ r \end{pmatrix} : r \in \mathbb{R} \mid r \neq 0 \right\}.$$

One such eigenvector is $\mathbf{b} := \begin{pmatrix} 3 \\ 4 \\ 2 \end{pmatrix}$. (Check this!)

Thus the line through the origin with vector equation $\mathbf{r} = \mu\mathbf{b}$ ($\mu \in \mathbb{R}$) is fixed by the linear transformation t_A represented by A , and a point on this line with position vector $\mathbf{v}$ is mapped by t_A to the point that has position vector $3\mathbf{v}$.

It is left as an exercise for you to compute the eigenvectors of A corresponding to the other eigenvalue, 1.

10.5 Eigenvectors and eigenvalues for linear transformations of the plane

We revisit rotations and reflexions, axes stretches, dilations and shears in $\mathbb{R}^2$, to see how eigenvectors and eigenvalues are involved.

Rotations. No real eigenvalues, as no fixed lines, except for the case of a rotation through $m\pi$ for some $m \in \mathbb{Z}$, when every nonzero vector is an eigenvector with eigenvalue $(-1)^m$, and the rotation matrix is $(-1)^m I_2$. (In general, an anticlockwise rotation through θ has complex eigenvalues $e^{i\theta}$ and $e^{-i\theta}$.)

Reflexions. Here every (nonzero) vector in the direction $\mathbf{u}$ of the mirror is an eigenvector with eigenvalue $+1$, and every (nonzero) vector orthogonal to $\mathbf{u}$ is an eigenvector with eigenvalue -1 .

Axes stretches. The matrix $\begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}$ (with a and d nonzero and $a \neq d$) has the position vector of every point on the x -axis (except the origin) as an eigenvector with eigenvalue a , and the position vector of every point on the y -axis (except the origin) as an eigenvector with eigenvalue d .

Dilations. The matrix $\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$ (with $a > 0$) has a as an eigenvalue and every nonzero vector as an eigenvector corresponding to a .

Shears. Consider for example $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. By an easy calculation the only eigenvalue is $+1$, and the eigenvectors corresponding to this eigenvalue are the vectors $\left\{ \begin{pmatrix} t \\ 0 \end{pmatrix} : t \in \mathbb{R} \mid t \neq 0 \right\}$, that is, the position vectors of all points on the x -axis other than the origin. In general, all shears (in 2 dimensions) have characteristic polynomial $x^2 - 2x + 1 = (x - 1)^2$, and thus eigenvalues $1, 1$. However, all eigenvectors of a shear lie along a single line (origin excluded).

10.6 Rotations and reflexions in $\mathbb{R}^3$

It was mentioned in an earlier remark (without proof) that every rigid motion of $\mathbb{R}^3$ which fixes the origin is represented by an *orthogonal* matrix, that is to say a matrix A with the property that $AA^T = I_3 = A^T A$ (where A^T denotes the *transpose* of A). It follows that $(\det A)^2 = 1$ and hence that $\det A = \pm 1$.

Rigid motions of $\mathbb{R}^3$ represented by matrices A which have $\det A = +1$ are called *orientation-preserving* (they send right-handed triples of vectors to right-handed triples of vectors), and those which have $\det A = -1$ are called *orientation-reversing*.

As was mentioned in another earlier remark, every linear transformation of $\mathbb{R}^3$ has a fixed line (since the characteristic polynomial is a cubic and therefore has a real root). If the linear transformation is a rigid motion, the corresponding eigenvalue must be $+1$ or -1 , since rigid motions preserve distances.

Let A be a 3×3 matrix representing a rigid motion of $\mathbb{R}^3$ fixing the origin.

Case 1. A has $+1$ as an eigenvalue.

Let $\mathbf{u} \neq \mathbf{0}_3$ be an eigenvector with eigenvalue $+1$. Now A maps the plane Π through the origin orthogonal to $\mathbf{u}$ to itself. If this map of Π to itself is a *rotation* then A represents a rotation of $\mathbb{R}^3$ around the axis which has direction $\mathbf{u}$. If the map of Π to itself is a

reflexion, with mirror a line in Π having direction $\mathbf{v}$, then A represents a reflexion of $\mathbb{R}^3$ with mirror the plane containing $\mathbf{u}$ and $\mathbf{v}$.

Case 2. A has -1 as an eigenvalue.

Let $\mathbf{u} \neq \mathbf{0}_3$ be an eigenvector with eigenvalue -1 . Once again, A maps the plane Π through the origin orthogonal to $\mathbf{u}$ to itself. If this map of Π to itself is a reflexion, then Π contains an eigenvector $\mathbf{v}$ of A with eigenvalue $+1$, corresponding to the direction of the mirror in Π , and an eigenvector $\mathbf{w}$ orthogonal to $\mathbf{v}$ with eigenvalue -1 . Thus $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w}$ are an orthogonal set of vectors and A sends $\mathbf{u} \mapsto -\mathbf{u}$, $\mathbf{v} \mapsto \mathbf{v}$, $\mathbf{w} \mapsto -\mathbf{w}$, and so A is a *rotation* about the direction of $\mathbf{v}$ through an angle π . If, on the other hand, the map of Π to itself is a rotation, there is not much more to say except that A represents the composition of a reflexion in the plane Π followed by a rotation about the axis orthogonal to Π .

Remark. It follows from the analysis above that every orientation-preserving rigid motion of $\mathbb{R}^3$ which fixes the origin is a rotation about some axis, but that (unlike the case for $\mathbb{R}^2$) an orientation-reversing rigid motion of $\mathbb{R}^3$ which fixes the origin is not necessarily a reflexion.

We finish by exhibiting examples of 3×3 matrices representing rotations and reflexions.

Example 1. The 3×3 matrix which represents a rotation through angle θ around the z -axis is

$$\begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

and similarly the 3×3 matrices which represent rotations through angle θ around the x - and y -axes are respectively.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{pmatrix}.$$

In these examples we look at the origin from a point somewhere along the positive portion of the relevant axis. These are rotations through an angle of θ anticlockwise from this point of view. (If we look the origin from the opposite direction then these rotations have an anticlockwise angle of $-\theta$.)

Example 2. The reflexion $s_{\mathbf{n}} = s_{\Pi}$ in the plane Π through the origin with normal vector $\mathbf{n} \neq \mathbf{0} = \mathbf{0}_3$ sends each vector $\mathbf{x}$ to:

$$s_{\Pi}(\mathbf{x}) = \mathbf{x} - 2 \frac{(\mathbf{x} \cdot \mathbf{n})}{|\mathbf{n}|^2} \mathbf{n}.$$

To see this, draw a picture like the one we used in our calculation of the distance from the point X with position vector $\mathbf{x}$, to the plane Π (the point with position vector $\mathbf{x} - \frac{(\mathbf{x} \cdot \mathbf{n})}{|\mathbf{n}|^2} \mathbf{n}$ is the point of Π which is closest to X).

If the equation of the plane Π is $ax + by + cz = 0$, we can take $\mathbf{n} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$, and so:

$$s_{\Pi}(x\mathbf{i} + y\mathbf{j} + z\mathbf{k}) = x\mathbf{i} + y\mathbf{j} + z\mathbf{k} - 2\frac{ax + by + cz}{a^2 + b^2 + c^2}(a\mathbf{i} + b\mathbf{j} + c\mathbf{k}).$$

Thus in particular

$$s_{\Pi}(\mathbf{i}) = \mathbf{i} - 2\frac{a}{a^2 + b^2 + c^2}(a\mathbf{i} + b\mathbf{j} + c\mathbf{k}) = \frac{1}{a^2 + b^2 + c^2}((b^2 + c^2 - a^2)\mathbf{i} - 2ab\mathbf{j} - 2ac\mathbf{k}),$$

and so the first column of the matrix S_{Π} representing s_{Π} is

$$\frac{1}{a^2 + b^2 + c^2} \begin{pmatrix} b^2 + c^2 - a^2 \\ -2ab \\ -2ca \end{pmatrix}.$$

Similarly, one can compute the second and third columns, to give the following matrix, which you are not expected to memorise for the examination for this module!

$$S_{\Pi} = \frac{1}{a^2 + b^2 + c^2} \begin{pmatrix} b^2 + c^2 - a^2 & -2ab & -2ac \\ -2ab & c^2 + a^2 - b^2 & -2bc \\ -2ac & -2bc & a^2 + b^2 - c^2 \end{pmatrix}.$$

Note that

$$S_{\Pi}(\mathbf{n}) = S_{\Pi} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} -a \\ -b \\ -c \end{pmatrix} = -\mathbf{n},$$

which is exactly what one would expect since $\mathbf{n}$ is orthogonal to the plane Π . Now each vector $\mathbf{x}$ in $\mathbb{R}^3$ can be written uniquely as $\mathbf{x} = \mathbf{u} + \mathbf{v}$ where $\mathbf{u}$ and $\mathbf{n}$ are collinear and $\mathbf{v}$ is parallel to (and thus in) Π , so that $\mathbf{v}$ is orthogonal to $\mathbf{n}$. (We have $\mathbf{u} = \frac{(\mathbf{x} \cdot \mathbf{n})}{|\mathbf{n}|^2} \mathbf{n} = (\mathbf{x} \cdot \hat{\mathbf{n}}) \hat{\mathbf{n}}$ and $\mathbf{v} = \mathbf{x} - (\mathbf{x} \cdot \hat{\mathbf{n}}) \hat{\mathbf{n}}$, where $\hat{\mathbf{n}}$ is the unit vector in the direction of $\mathbf{n}$. You should prove the assertions of the previous sentence, and also try to prove that $\mathbf{u}$ and $\mathbf{v}$ are the *unique* vectors with this property.) Then $s_{\Pi}(\mathbf{u}) = -\mathbf{u}$ and $s_{\Pi}(\mathbf{v}) = \mathbf{v}$, so that $\mathbf{u}$ is an eigenvector of s_{Π} with corresponding eigenvalue -1 (unless $\mathbf{u} = \mathbf{0}$), and $\mathbf{v}$ is an eigenvector of s_{Π} with corresponding eigenvalue 1 (unless $\mathbf{v} = \mathbf{0}$). Moreover, $s_{\Pi}(\mathbf{x}) = s_{\Pi}(\mathbf{u} + \mathbf{v}) = -\mathbf{u} + \mathbf{v}$, so that (exercise for you) $\mathbf{x}$ is an eigenvector of s_{Π} when *exactly one* of the conditions $\mathbf{u} = \mathbf{0}$ and $\mathbf{v} = \mathbf{0}$ holds.

All of what we have done above generalises readily to reflexions in a hyperplane Π of $\mathbb{R}^n$ orthogonal to the vector $\mathbf{n} \neq \mathbf{0}_n$. I leave the reader to work out the necessary details.

Exercise. Check that the formula for S_{Π} gives the right answer when Π is the (x, y) plane (the plane defined by the equation $z = 0$). Also check that the transformation s_{Π} defined at the start of Example 2 is indeed a linear map.