LECTURE 4

Eigenvalues and eigenvectors

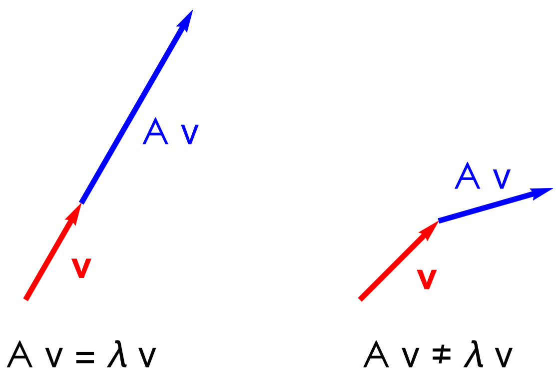
In the previous lecture, we asked: when is a matrix transformation guaranteed to preserve the lengths of, and relative angles between, any vectors it acts upon. This lead to the idea of orthogonal matrices, which are used extensively in chemical applications of group theory.

Today, we'll be thinking about matrices that transform vectors into *multiples of themselves*. Specifically, given some matrix A, can we find (non-zero) vectors $\mathbf{v}_i$ and scalars λ_i that satisfy the equation below?

$$\mathsf{A}\mathbf{v}_i = \lambda_i \mathbf{v}_i. \tag{4.1}$$

Such vectors are called the *eigenvectors* of A, and the corresponding scalar multiples are called the *eigenvalues*. They turn out to be immensely important for quantum mechanics.

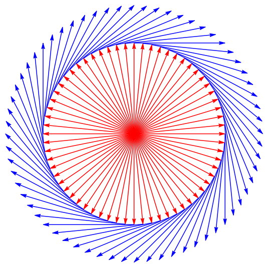
A geometric way to tackle the question is as follows. Given a matrix A, pick some vector $\mathbf{v}$ and calculate $\mathsf{A}\mathbf{v}$. Draw $\mathbf{v}$ (shown in red below) and $\mathsf{A}\mathbf{v}$ (shown in blue), head-to-tail. If $\mathbf{v}$ is an eigenvector of A, then $\mathsf{A}\mathbf{v} = \lambda\mathbf{v}$ and so the two arrows will be aligned (or anti-aligned, if $\lambda < 0$). Otherwise, $\mathsf{A}\mathbf{v}$ will point in a different direction from $\mathbf{v}$. The left-hand figure below illustrates an eigenvector of A, and the right-hand figure illustrates a vector that is not an eigenvector of A.



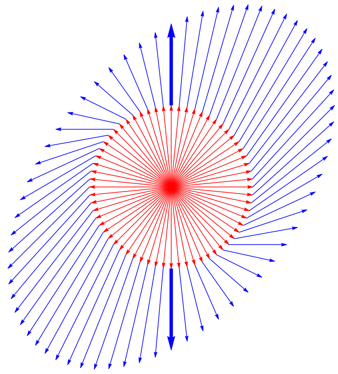
Using this method, we can search for eigenvectors of a particular matrix, by trying all possible non-zero vectors and seeing whether the red and blue arrows are aligned. I've done that below, for some 2×2 matrices. [For clarity, I've drawn the blue arrow thicker if it is aligned (or anti-aligned) with its red arrow.]

Notice that I've deliberately chosen some familiar matrices here. The first example is a rotation matrix ($\pi/2$ clockwise). It's therefore obvious that it will have no eigenvectors. The third matrix reflects in the line $y = x$. From that alone, you could figure out the eigenvectors. And obviously the fourth matrix is the identity.

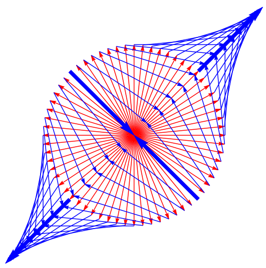
$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$



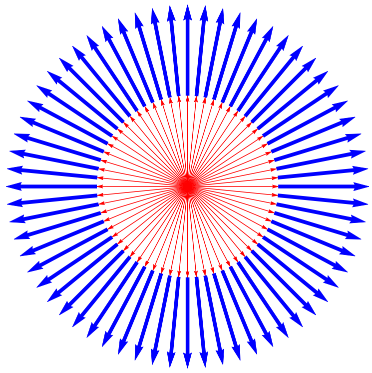
$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$



$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$



$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$



We can see that

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

has no eigenvectors. On the other hand, *every* non-zero vector is an eigenvector of the identity matrix,

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

For the other two matrices, we see that $\mathsf{A}\mathbf{v}$ is sometimes aligned with $\mathbf{v}$, but mostly not. Clearly, the general eigenvector question is subtle, and it will take us some time to understand it properly.

What counts as an eigenvector?

Before we move on, we need to be a bit more specific about how we count eigenvectors. Notice that (4.1) is trivially satisfied by the zero vector, $\mathbf{v} = \mathbf{0}$, whatever matrix A we choose. For this reason, **we do not count the zero vector as an eigenvector**.

Second, look at the diagrams above. In all cases the bold vectors appear in pairs. This is because any multiple of an eigenvector is also an eigenvector. To see why, suppose

$$\mathsf{A}\mathbf{v} = \lambda \mathbf{v}.$$

Then, if we multiply both sides by some scalar c, i.e.

$$c\mathsf{A}\mathbf{v} = c\lambda \mathbf{v},$$

we can use the linearity of the matrix multiplication to rewrite the equation as

$$\mathsf{A}(c\mathbf{v}) = \lambda(c\mathbf{v}).$$

If we call $\mathbf{u} = c\mathbf{v}$, we can see that $\mathbf{u}$ is also an eigenvector of A, because

$$\mathsf{A}\mathbf{u} = \lambda\mathbf{u}.$$

Indeed, $\mathbf{u}$ has the same eigenvalue, λ, as $\mathbf{v}$. So, in other words,

> If $\mathbf{v}$ is an eigenvector of A with eigenvalue λ, so is $c\mathbf{v}$ for any $c \neq 0$. (4.2)

For this reason, when counting eigenvectors, we will be mostly interested in the number of *distinct* eigenvectors: that is, eigenvectors that are not just multiples of each other.

In the example above, we see that

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

are both eigenvectors of

$$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

If you can't see this from the diagram, look back at how the arrows were defined on the previous page.

However, because these two vectors are multiples of each other, we count them as the same thing and say that this matrix has only one distinct eigenvector. Likewise

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

has only two distinct eigenvectors.

For those of you who are desperate for some algebra, here's a simple exercise.

Example 1. Show that

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

are (distinct) eigenvectors of

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

and identify their eigenvalues.

To do this, all we need to do is check that the eigenvector equation (4.1) is satisfied. Very straightforwardly, we find that

$$\mathsf{A}\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1 \times \mathbf{v}_1 \quad \text{and} \quad \mathsf{A}\mathbf{v}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix} = (-1) \times \mathbf{v}_2.$$

So $\lambda_1 = 1$ and $\lambda_2 = -1$. The eigenvectors are distinct, because they are not multiples of each other.

Determining the eigenvalues of a matrix

The geometric, trial-and-error method of finding eigenvectors is spectacularly beautiful, but it's not a useful way to solve the problem in general. For that, we need to start calculating.

We begin by rearranging (4.1). Obviously we can write

$$\mathsf{A}\mathbf{v}_i - \lambda_i \mathbf{v}_i = \mathbf{0}.$$

Perhaps less obviously, we can factorise the LHS by introducing an identity matrix:

$$(\mathsf{A} - \lambda_i \mathrm{I})\mathbf{v}_i = \mathbf{0}. \tag{4.3}$$

It's clear that this reproduces the equation above, if one expands the LHS of this equation. The identity matrix is absolutely needed, because $(\mathsf{A} - \lambda_i)\mathbf{v}_i$ does not make sense as an expression. Subtracting a number from a matrix is undefined.

Now, we notice that (4.3) is actually just a set of homogeneous linear equations. To see why, refer back to the section on 'Homogeneous linear equations: redux number 2' in Lecture 2.

As we've seen many times, (4.3) always has a trivial solution, $\mathbf{v}_i = \mathbf{0}$. We already said that this does not count as an eigenvector. As we've also seen, (4.3) can have non-trivial solutions; the condition for non-trivial solutions is

$$\det(\mathsf{A} - \lambda_i \mathrm{I}) = 0. \tag{4.4}$$

This equation can be used to determine the eigenvalues of a matrix. It is known as the *characteristic equation* of A.

Some physicists and chemists call the determinant here a 'secular determinant'. This terminology comes from early applications of the eigenvector problem to celestial mechanics.

Example 2. Determine the eigenvalues of the matrix

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Using the characteristic equation, (4.4),we need

$$\det\left[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right] = 0.$$

This simplifies to

$$\begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = 0,$$

which further simplifies to $\lambda^2 - 1 = 0$ and hence $\lambda = \pm 1$. These agree, as they should, with the two eigenvalues we determined in Example 1.

Determining the eigenvectors of a matrix

Let's return to (4.3):

$$(\mathsf{A} - \lambda_i \mathtt{I})\mathbf{v}_i = \mathbf{0}. \tag{4.5}$$

We've just seen how to find λ_i such that this equation has non-trivial solutions for $\mathbf{v}_i$. The remaining job is to find those non-trivial solutions. It's probably easiest to explain this with an example.

Example 3. Determine the eigenvectors of the matrix

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

We saw earlier that the eigenvalues are 1 and -1. We find the corresponding eigenvectors by considering each case in turn.

When $\lambda = 1$, (4.3) becomes

$$(\mathsf{A} - \mathtt{I})\mathbf{v} = \mathbf{0}.$$

If we denote the components of the eigenvector $\mathbf{v}$ as x and y, we can express this equation as

$$\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

which is equivalent to the simultaneous linear equations

$$\begin{aligned} -x + y &= 0 \\ x - y &= 0. \end{aligned}$$

Given that we are finding non-trivial solutions, we know from last Term that these two equations will be linearly-dependent. Therefore, we expect one equation to give us redundant information. In the present case, this fact is patently obvious, and we can see that the eigenvector satisfies a single constraint,

$$y = x.$$

Therefore, the eigenvector is of the form

$$\begin{pmatrix} x \\ x \end{pmatrix},$$

for any value of x. Given what we showed in (4.2), this makes sense. Taking out a factor of x, we see that the eigenvector is of the form

$$x \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

i.e. that it is any multiple of

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

For simplicity, let's give the vector above as our answer.

Now we turn to the second eigenvector, for $\lambda = -1$. We shall go through this in fewer steps.

When $\lambda = -1$, (4.3) becomes

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

which is equivalent to the simultaneous linear equations

$$\begin{aligned} x + y &= 0 \\ x + y &= 0. \end{aligned}$$

Hence, $y = -x$ and the eigenvector must be of the form

$$\begin{pmatrix} x \\ -x \end{pmatrix}.$$

Let's take $x = 1$ for convenience, to get the answer

$$\begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Over this example and the last, we've shown that the eigenvectors and eigenvalues of the given matrix are

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \text{and} \qquad \mathbf{v}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

with $\lambda_1 = 1$ and $\lambda_2 = -1$. This agrees precisely with Example 1.

Useful checks

There are two checks that one can do when finding eigenvalues and eigenvectors. The most obvious, which is always worth doing, is to check that your $\mathbf{v}_i$s and λ_is do actually satisfy (4.1).

Very few students do this obvious check when finding eigenvectors in Prelims. It takes only seconds to do, and can save a lot of marks if you have made a mistake somewhere.

Before reaching this stage, it is helpful to know if your eigenvalues are correct. We shall see, later in the course, that the sum of the eigenvalues is equal to the sum of the diagonal elements of the matrix. The latter is called the *trace* of the matrix.

In the previous example, the trace of the matrix was $0 + 0 = 0$. The eigenvalues were -1 and 1, which indeed add up to zero. It's well worth checking this once you've determined the eigenvalues, as everything will go horribly wrong if you try to find the eigenvectors using the wrong values of λ.

Actually, if you try to find eigenvectors with the wrong λ, you should find that the 'eigenvector' comes out to be the zero vector. This is because your λ will not give rise to non-trivial solutions of (4.3). This is another way to tell that your eigenvalues are incorrect.

A 3×3 example

To end the lecture, let's do another example. We'll consider a three-dimensional matrix, as students sometimes find it more difficult to solve the equations in these cases.

Example 4. Find the eigenvalues and eigenvectors of

$$\mathsf{A} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

The characteristic equation, $\det(\mathsf{A} - \lambda \mathrm{I}) = 0$, is

$$\begin{vmatrix} -\lambda & 1 & 0 \\ 1 & -\lambda & 0 \\ 0 & 0 & 2-\lambda \end{vmatrix} = 0.$$

In this case, expansion of the determinant is straightforward: we multiply out along the bottom row to get

$$(2-\lambda)\begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = 0.$$

and hence $(2-\lambda)(\lambda^2 - 1) = 0$. This leads to the eigenvalues

$$\lambda_1 = -1 \qquad \lambda_2 = 1 \qquad \lambda_3 = 2.$$

We carry out the check: the eigenvalues sum to 2, which is the same as the trace of the matrix. Good!

For $\lambda = \lambda_1 (= -1)$, the equation for $\mathbf{v}_1$ is

$$(\mathsf{A} - \lambda_1 \mathrm{I})\mathbf{v}_1 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

In simultaneous equation form, this is

$$\begin{aligned} x + y &= 0 \\ x + y &= 0 \\ 3z &= 0. \end{aligned}$$

As before, we notice that one equation is redundant. The other two give us $y = -x$ and $z = 0$, so the eigenvector must be of the form

$$\mathbf{v}_1 = \begin{pmatrix} x \\ -x \\ 0 \end{pmatrix}.$$

We take $x = 1$ for simplicity, to give

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}.$$

For $\lambda = \lambda_2 (= 1)$, the equation for $\mathbf{v}_1$ is

$$(\mathsf{A} - \lambda_2 \mathrm{I})\mathbf{v}_2 = \begin{pmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

This is very similar to the case above, and we find that $y = x$ and $z = 0$. Hence

$$\mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

The case of $\lambda = \lambda_3 (= 2)$, is a little different. We have

$$(\mathsf{A} - \lambda_3 \mathrm{I})\mathbf{v}_3 = \begin{pmatrix} -2 & 1 & 0 \\ 1 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

which means that

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

People tend to become unstuck at this point. Certainly, we see that one equation (the last one) is redundant, as should be the case. From the first two, we get

$$\begin{aligned} 2x &= y \\ x &= 2y \end{aligned}$$

The only solution to these is $x = y = 0$. (Substitute one into the other to show this rigorously.) The common mistake at this point, for some reason, is to think that $z = 0$ also. Actually, *z can be anything*: the simultaneous equations do not restrict z in any way. So the eigenvector must be of the form

$$\mathbf{v}_3 = \begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix}$$

for any z, and hence we can take it to be

$$\mathbf{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$