Eigenvectors and eigenvalues

Jeremy Orloff

Here is a short note on eigenvalues and eigenvectors. We will look almost exclusively at 2x2 matrices. These have almost all the features of bigger square matrices and they are computationally easy.

Contents

Etymology:
Eigenvector is German for 'own vector'. I can sort-of see how this makes sense, but it is not very descriptive for me.

Definition:
For a square matrix M an *eigenvector* is a non-zero vector $\mathbf{v}$ that satisfies the equation

$$M\,\mathbf{v} = \lambda\,\mathbf{v} \text{ for some number } \lambda. \tag{1}$$

The number λ is called the *eigenvalue* corresponding to $\mathbf{v}$. We will call equation (1) the *eigenvector equation.*

Comments:
1. Using λ for the eigenvalue is a fairly common practice when looking at generic matrices. If the eigenvalue has a physical interpretation I'll often use a corresponding letter. For example, in population matrices the eigenvalues are growthrates, so I'll use the letter r.

2. Eigenvectors are not unique. That is, if $\mathbf{v}$ is an eigenvector with eigenvalue λ then so is $2\,\mathbf{v}$, $3.14\,\mathbf{v}$, indeed so is any scalar multiple of $\mathbf{v}$.

Why eigenvectors are special:

Example 1: Let $A = \begin{pmatrix} 6 & 5 \\ 1 & 2 \end{pmatrix}$.

We will explore how A transforms vectors and what makes an eigenvector special. We will see that A scales and rotates most vectors, but only scales eigenvectors. That is, eigenvectors lie on lines that are unmoved by A.

Take $\mathbf{u_1} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \Rightarrow A\mathbf{u_1} = \begin{pmatrix} 6 \\ 1 \end{pmatrix}$; Take $\mathbf{u_2} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \Rightarrow A\mathbf{u_2} = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$.

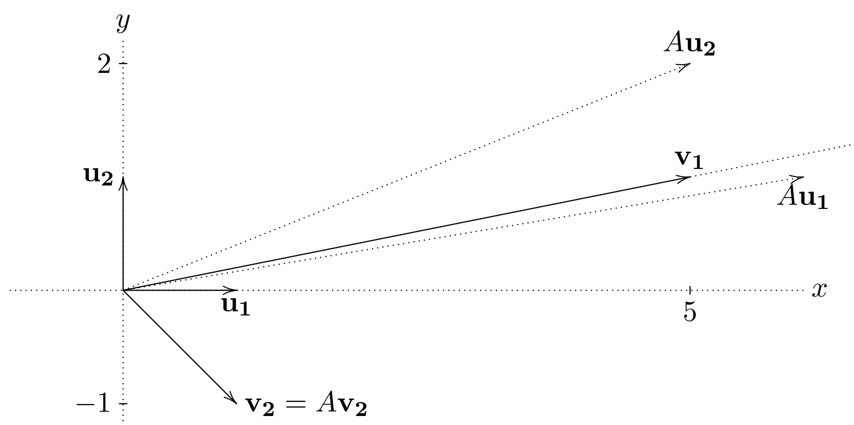
We see that A scales and turns most vectors.

Now take $\mathbf{v_1} = \begin{pmatrix} 5 \\ 1 \end{pmatrix} \Rightarrow A\mathbf{v_1} = \begin{pmatrix} 35 \\ 7 \end{pmatrix} = 7\mathbf{v_1}$.

The eigenvector is special since A just scales it by 7.

Likewise, $\mathbf{v_2} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ then $A\mathbf{v_2} = \mathbf{v_2}$.

The eigenvector $\mathbf{v_2}$ is really special, it is unmoved by A.



Action of the matrix A on vectors

Example 2: Any rotation in three dimensions is around some axis. The vector along this axis is fixed by the rotation. I.e. it is an eigenvector with eigenvalue 1.

Computational approach:

In order to find the eigenvectors and eigenvalues we use the following basic fact about matrices.

Fact: The equation $M\mathbf{v} = 0$ is satisfied by a non-zero vector $\mathbf{v}$ if and only if $\det(M) = 0$.

Now, as always, let I be the identity matrix. We manipulate the eigenvector equation (equation 1) as follows.

$$A\,\mathbf{v} = \lambda\,\mathbf{v} \;\Leftrightarrow\; A\,\mathbf{v} = \lambda\,I\mathbf{v} \;\Leftrightarrow\; A\,\mathbf{v} - \lambda\,I\mathbf{v} = 0 \;\Leftrightarrow\; (A - \lambda\,I)\mathbf{v} = 0.$$

Now we apply our 'fact' to the last equation above to conclude that λ is an eigenvalue if and only if

$$\det(A - \lambda\,I) = 0 \tag{2}$$

(Recall that an eigenvector is by definition non-zero.)

We call equation (2) the characteristic equation. (Eigenvalues are sometimes called characteristic values.) It allows us to find the eigenvalues and eigenvectors separately in a two step process.

Notation: $\det(A) = |A|$.

Example 3: Find the eigenvalues and eigenvectors of the matrix $A = \begin{pmatrix} 6 & 5 \\ 1 & 2 \end{pmatrix}$.

Step 1, find λ (eigenvalues): $|A - \lambda I| = 0$ **(characteristic equation)**

$\Rightarrow \left| \begin{pmatrix} 6 & 5 \\ 1 & 2 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \right| = 0 \;\Rightarrow\; \begin{vmatrix} 6-\lambda & 5 \\ 1 & 2-\lambda \end{vmatrix} = 0.$

$\Rightarrow\; (6-\lambda)(2-\lambda) - 5 = \lambda^2 - 8\lambda + 7 = 0 \;\Rightarrow\; \boxed{\lambda = 7,\, 1.}$

Step 2, find $\mathbf{v}$ (eigenvectors): $(A - \lambda I)\mathbf{v} = \mathbf{0}$.

Let $\mathbf{v} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$.

$\lambda = 7$: $\begin{pmatrix} -1 & 5 \\ 1 & -5 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \;\Rightarrow\; -a_1 + 5a_2 = 0.$ Take $\mathbf{v} = \begin{pmatrix} 5 \\ 1 \end{pmatrix}$.

$\lambda = 1$: $\begin{pmatrix} 5 & 5 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \;\Rightarrow\; 5a_1 + 5a_2 = 0.$ Take $\mathbf{v} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

To repeat the comment above, any scalar multiple of these eigenvectors is also an eigenvector with the same eigenvalue.

Comment: In Matlab the function 'eig' returns the eigenvectors and eigenvalues of a matrix.

Diagonal matrices and decoupling:

In this section we will see how easy it is to work with diagonal matrices and how finding eigenvalues and eigenvectors of a matrix is like turning it into a diagonal matrix.

Example 4: Consider the diagonal matrix $B = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$

Convince yourself that $B \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x \\ 3y \end{pmatrix}$. That is B scales the x-coordinate by 2 and the y-coordinate by 3.

We also have

$$B \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad B \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 3 \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Comparing this with the eigenvector equation (equation (1)) we see that $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ are eigenvectors with eigenvalues 2 and 3 respectively. That is, for a diagonal matrix the diagonal entries are the eigenvalues and the eigenvectors all point along the coordinate axes.

In the following examples we will look at a discrete time example. This will make the appearance of eigenvectors more natural. Later in the course we will look at systems of differential equations where we will need to incorporate exponentials rather than powers into our analysis.

Before going on let's see why we call the example above a *decoupled system.* Suppose x and y are physical values that are updated at intervals according to the relations

$$\begin{aligned} x(t+1) &= 2x(t) \\ y(t+1) &= 3y(t). \end{aligned}$$

We consider x and y to be decoupled because their values have no affect on each other.

Example 5: Now consider the system $\begin{aligned} x(t+1) &= 6x(t) + 5y(t) \\ y(t+1) &= x(t) + 2y(t). \end{aligned}$

In matrix form this is

$$\begin{pmatrix} x(t+1) \\ y(t+1) \end{pmatrix} = A \begin{pmatrix} x(t) \\ y(t) \end{pmatrix},$$

where $A = \begin{pmatrix} 6 & 5 \\ 1 & 2 \end{pmatrix}$ is the matrix in examples (1) and (3) above.

In this system the variables x and y are coupled.

Thinking of x and y as populations let's see what happens to the system from various starting points.

(i) Start from an eigenvector (we scale the eigenvector in example (3) by .5):
Let

$$\begin{pmatrix} x(0) \\ y(0) \end{pmatrix} = \begin{pmatrix} 2.5 \\ 0.5 \end{pmatrix}.$$

Then

$$\begin{pmatrix} x(1) \\ y(1) \end{pmatrix} = A \begin{pmatrix} x(0) \\ y(0) \end{pmatrix} = 7 \begin{pmatrix} 2.5 \\ 0.5 \end{pmatrix}.$$

Generally

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = A^t \begin{pmatrix} x(0) \\ y(0) \end{pmatrix} = 7^t \begin{pmatrix} 2.5 \\ 0.5 \end{pmatrix}.$$

We see that our initial population distribution is a stable one that grows exponentially over time. That is, the eigenvector represents a stable distribution and the eigenvalue its growth rate.

(ii) Start from another eigenvector:
Let

$$\begin{pmatrix} x(0) \\ y(0) \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Then

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

Generally

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = A^t \begin{pmatrix} x(0) \\ y(0) \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

We see that our initial population distribution is a stable one that stays constant over time. There is of course one problem, the initial vector has a negative value and is thus not a realistic population vector.

(iii) Start from a non-eigenvector:
Let

$$\begin{pmatrix} x(0) \\ y(0) \end{pmatrix} = \begin{pmatrix} 9 \\ 3 \end{pmatrix}.$$

Then

$$\begin{pmatrix} x(1) \\ y(1) \end{pmatrix} = A \begin{pmatrix} x(0) \\ y(0) \end{pmatrix} = \begin{pmatrix} 69 \\ 15 \end{pmatrix}.$$

Working like this it is difficult to see how the population will evolve. Here's where eigenvalues do their thing. Note

$$\begin{pmatrix} 9 \\ 3 \end{pmatrix} = 2 \begin{pmatrix} 5 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

That is, we can decompose our initial vector into a sum of eigenvectors. Now it's easy to apply A any number of times

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = A^t \begin{pmatrix} x(0) \\ y(0) \end{pmatrix} = 7^t\, 2 \begin{pmatrix} 5 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Notice that over time the $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ piece of the population vector becomes insignificant when compared with the first piece. This is a standard phenomenon, the largest eigenvalue is the dominant one, the others are insignificant in the long run.

The above hints at diagonalization and decoupling. In order to make it explicit we introduce two new variables u and v with the relation

$$\begin{pmatrix} x \\ y \end{pmatrix} = u \begin{pmatrix} 5 \\ 1 \end{pmatrix} + v \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

So, u and v are the coefficients in the decomposition of $\begin{pmatrix} x \\ y \end{pmatrix}$ into eigenvectors. For the moment we will set aside any physical interpretation of u and v and see how they act as a system.

We have

$$\begin{pmatrix} x(t+1) \\ y(t+1) \end{pmatrix} = A \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = A \left(u(t) \begin{pmatrix} 5 \\ 1 \end{pmatrix} + v(t) \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right).$$

Unwinding this we respect to u and v we have the two equations

$$\begin{pmatrix} x(t+1) \\ y(t+1) \end{pmatrix} = u(t+1) \begin{pmatrix} 5 \\ 1 \end{pmatrix} + v(t+1) \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

and

$$\begin{pmatrix} x(t+1) \\ y(t+1) \end{pmatrix} = 7\, u(t) \begin{pmatrix} 5 \\ 1 \end{pmatrix} + v(t) \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

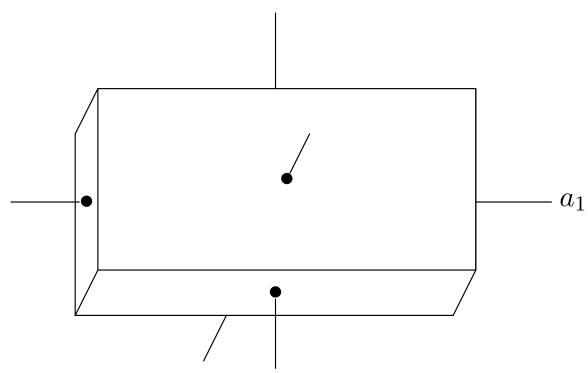
That is we have the decoupled system

$$
\begin{aligned}
u(t+1) &= 7u(t) \\
v(t+1) &= v(t)
\end{aligned}
$$

Note that the eigenvalues of A are precisely the coefficients of the decoupled system.

Intutitive meaning and modes:
Physically we think of each of the eigenvectors as representing pure modes.

Example 6: A book has three axes of symmetry. It will spin nicely about these axes. (Although the one labeled a_1 is so unstable that it is difficult to actually spin the book around it.) These are the pure modes. Spinning in any other way is a combination of these pure modes. The axes are eigenvectors of the coefficient matrix for the system of differential equations modeling this system.



Population matrices:
Here we still think of each of the eigenvectors as pure modes. As in example (5) not all the pure modes are physically realistic. Some eigenvectors may have negative entries and some eigenvalues may be negative or complex. However, any population vector is a combination of these pure modes.

Again, as in example (5), there is one realistic mode and it is dominant in the sense that the other modes become insignificant in the long run.

End of Eigenvectors and Eigenvalues.