Eigenvalues and eigenvectors

Let's say that we want to calculate the following quantity:

$$A^{10}x = y \tag{1}$$

where A is a 2×2 matrix and x is a 2×1 matrix. For example if $A = \begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix}$ and $x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, the quantity we want to calculate is the following:

$$\begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix}^{10} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{2}$$

Certainly it would not be easy to calculate this directly as the matrix multiplications are complicated. Is there a simpler way? The answer is yes - using the concept of eigenvectors and eigenvalues, which is the topic of this article (it should be noted however that this is not merely a tool for calculation; it has wide applications in science and engineering).

Now, let's go back to our original problem. Suppose we could find two 2×1 matrices e_1 and e_2 that satisfy the following conditions:

$$\begin{aligned} Ae_1 &= \lambda_1 e_1 \\ Ae_2 &= \lambda_2 e_2 \end{aligned} \tag{3}$$

where λ_1 and λ_2 are simply numbers.

We obtain the following:

$$\begin{aligned} A^{10}e_1 &= A^9(Ae_1) = A^9\lambda_1 e_1 = \lambda_1 A^9 e_1 \\ &= \lambda_1 A^8(Ae_1) = \lambda^1 A^8 \lambda_1 e_1 = (\lambda_1)^2 A^8 e_1 \end{aligned} \tag{4}$$

Continuing in this way, we obtain:

$$A^{10}e_1 = \lambda_1^{10} e_1 \tag{5}$$

And similarly:

$$A^{10}e_2 = \lambda_2^{10} e_2 \tag{6}$$

Therefore, if our x in formula (1) happens to be like e_1 or e_2, our job is very much simplified. We just take the corresponding λ, calculate its tenth power, and multiply x by it. What is less obvious is that we can use a similar trick even if x itself doesn't happen to be like e_1 or e_2. Let's say that we can express x as a linear combination of e_1 and e_2 as follows:

$$x = c_1 e_1 + c_2 e_2 \tag{7}$$

where c_1 and c_2 are simply numbers.

We obtain:

$$A^{10}x = A^{10}(c_1e_1+c_2e_2) = A^{10}c_1e_1+A^{10}c_2e_2 = c_1(\lambda_1)^{10}e_1+c_2(\lambda_2)^{10}e_2 \tag{8}$$

which completes the calculation.

Let me remind you what we had to do. To calculate (1), we had to find an e_1 and e_2 that satisfy (3). Then we had to find a c_1 and c_2 that satisfy (7). This was our procedure. The point is that the whole calculation gets simplified if we can find e_1 and e_2 that satisfy (3), and then write x in terms of these. e_1 and e_2 are called eigenvectors and λ_1 and λ_2 are called (the corresponding) eigenvalues.

All of this seems abstract, so let's work an example. Let's go back to our example (2) from the beginning of the article. First, let's find the eigenvectors and the eigenvalues. Let $e_i = \begin{pmatrix} v_i \\ w_i \end{pmatrix}$ where i is 1 or 2. We have:

$$\begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} v_i \\ w_i \end{pmatrix} = \lambda_i \begin{pmatrix} v_i \\ w_i \end{pmatrix} \tag{9}$$

Componentwise, this is:

$$\begin{aligned} v_i + 2w_i &= \lambda_i v_i \\ -v_i + 4w_i &= \lambda_i w_i \end{aligned}$$

This implies:

$$(1-\lambda_i)v_i = -2w_i \tag{10}$$

$$-v_i = (\lambda_i - 4)w_i \tag{11}$$

which implies:

$$(1-\lambda_i)(\lambda_i-4) = -2 \times -1$$

Solving this equation we get:

$$\lambda_i = 2, 3 \tag{12}$$

For convenience, and without loss of generality, let's choose $\lambda_1 = 2$, $\lambda_2 = 3$ (the other choice would be $\lambda_1 = 3$, $\lambda_2 = 2$). Plugging this to the formula (10), we get:

$$e_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{13}$$

Of course, we could have chosen e_1 to be $(4,2)$ or $(6,3)$ instead of $(2,1)$, and e_2 to be $(-1,-1)$ or $(2,2)$ instead of $(1,1)$, but this choice won't change the result of the computation, as long as we are consistent.

Now plugging into (7) for our example of (2), we get :

$$x = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = c_1 \begin{pmatrix} 2 \\ 1 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{14}$$

$$\begin{aligned} 1 &= 2c_1 + c_2 \\ 0 &= c_1 + c_2 \end{aligned}$$

Therefore, we get $c_1 = 1$, $c_2 = -1$.

Now we can complete our calculation:

$$A^{10}x = A^{10}(e_1+e_2) = 2^{10}e_1+3^{10}e_2 = \begin{pmatrix} 2^{10} \times 2 - 3^{10} \times 1 \\ 2^{10} \times 1 - 3^{10} \times 1 \end{pmatrix} = \begin{pmatrix} -57001 \\ -58025 \end{pmatrix}$$

It turns out that an eigenvector with magnitude (called also "norm") 1 is often useful in physics. In our case, e_1 and e_2 don't satisfy such a condition. Nevertheless, from any eigenvectors, one can always construct eigenvectors with magnitude 1 by a process called "normalization." It is so simple. Just divide the eigenvector by the magnitude of the original eigenvector. Then, the new eigenvector will have magnitude 1 and it will still be an eigenvector with same eigenvalue. For example, in our case, e_1 has magnitude $\sqrt{2^2 + 1^2} = \sqrt{5}$. Then, define the new eigenvector as follows:

$$e_1' = \frac{e_1}{\sqrt{5}} = \begin{pmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix} \tag{15}$$

The norm is clearly 1, since $(2/\sqrt{5})^2 + (1/\sqrt{5})^2 = 1$, and it is still an eigenvector with eigenvalue 3 since

$$Ae_1' = A(e_1/\sqrt{5}) = (Ae_1)/\sqrt{5} = 3e_1/\sqrt{5} = 3(e_1/\sqrt{5}) = 3e_1' \tag{16}$$

Similarly, for the second normalized eigenvector, we have:

$$e_2' = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{17}$$

In general, an $n \times n$ matrix has n eigenvalues and n eigenvectors. For example, given a 4×4 matrix A, we can find 4 λ's (i.e. $\lambda_1, \lambda_2, \lambda_3, \lambda_4$) and 4 e's (i.e. e_1, e_2, e_3, e_4) that satisfy the following equations:

$$\begin{aligned} Ae_1 &= \lambda_1 e_1 \\ Ae_2 &= \lambda_2 e_2 \\ Ae_3 &= \lambda_3 e_3 \\ Ae_4 &= \lambda_4 e_4 \end{aligned}$$

The eigenvalues and the coefficients of the eigenvectors can be sometimes complex numbers even when the coefficients of the concerned matrix is real. Also, in general, these n vectors form the basis of n-dimensional vector space so that an n-dimensional vector can be uniquely expressed in terms of the linear combination of the eigenvectors.

We will briefly consider how the problem of finding eigenvalues and eigenvectors in cases when the dimension considered is more than two in our later article "Finding eigenvalues and eigenvectors." There, we will also learn why an $n \times n$ matrix has generally n eigenvalues and n eigenvectors.

Problem 1. What are the eigenvalues of an identity matrix? Show that any vector is an eigenvector. This fact will be useful when we explain in our later article "Neutrino oscillation, clarified" why the presence of neutrino oscillation implies that not all the three masses of neutrinos are same

Problem 2. Explain why eigenvalues and eigenvectors can be only defined for square matrices (i.e., matrices with the same number of rows and columns.)

Problem 3. Let's denote a square matrix A as follows:

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix} \tag{18}$$

Show that if at least one of the eigenvalues of the above matrix is 0, then the following n vectors

$$\begin{bmatrix} A_{11} \\ A_{21} \\ \cdots \\ A_{n1} \end{bmatrix}, \begin{bmatrix} A_{21} \\ A_{22} \\ \cdots \\ A_{n2} \end{bmatrix}, \cdots, \begin{bmatrix} A_{1n} \\ A_{2n} \\ \cdots \\ A_{nn} \end{bmatrix} \tag{19}$$

are linearly dependent. (Hint[1])

[1]If at least one of the eigenvalues of the matrix A is 0, there exists a non-zero $\vec{v}$ that satisfies $A\vec{v} = 0$.