Chapter 27

Eigenvalues and Eigenvectors

27.1 Main Idea

When a matrix multiplies a vector it can do two things: (i) it rotates the vector and (ii) it changes its length. Indeed, what else can it do?, as matrix multiplication maps a vector to another vector. What is *absolutely amazing* is that for a class of matrices called Hermitian matrices, one can find exactly n different orthogonal basis vectors for an $n \times n$ matrix such that when the matrix multiplies each of those basis vectors, there is no rotation, only scaling.

Hence, eigenvalues and eigenvectors are found by choosing the basis for the map such that when the matrix M multiplies one of the basis vectors $\vec{e}_i$ it results in a number λ_i times the basis vector $\vec{e}_i$. In equations, we have

$$M\vec{e}_i = \lambda_i \vec{e}_i. \tag{27.1}$$

In words we summarize as follows:

"A matrix times a vector equals a number times that vector."

Keep this mantra in mind to help you always understand what an eigenvalue and an eigenvector are. It is probably the most misunderstood concept by students of physics until they finally get it.

Again,

"A matrix times a vector equals a number times that vector."

How do we find these solutions? We rearrange the equations to have a shifted matrix annihilating a vector as follows:

$$\sum_{j=1}^{n} M_{ij}e_j = \lambda e_i \implies (M - \lambda I)\vec{e} = 0. \tag{27.2}$$

This has a nontrivial solution only if $\det(M - \lambda I) = 0$, which is an n-degree polynomial in λ. It will have n solutions, but if we work with a real vector space only the real solutions are valid, so it may have fewer than n *real* solutions.

27.2 Some Properties from the Spectral Theorem

If M is real and symmetric, it has exactly n real eigenvalues (this is called a symmetric matrix; the results actually hold a bit more broadly to include Hermitian matrices when matrix elements are complex valued). In addition, there might be multiple roots where $\lambda_i = \lambda_j$. When this occurs we say there is a *degeneracy*.

The eigenvectors are found by solving the equation

$$M\vec{e}_\lambda = \lambda\vec{e}_\lambda \tag{27.3}$$

for the specific λ value chosen (we already found the possible λ values when we found the roots to $\det(M - \lambda I) = 0$). Note, this becomes a row-reduction problem, which we have already solved.

If M is Hermitian, it also has exactly n real solutions. (A Hermitian matrix satisfies $M^{T*} = M$, which means we take the transpose and the complex conjugate of all elements; a symmetric matrix has all real values and satisfies $M^T = M$—obviously all symmetric matrices are also Hermitian). This result is called the spectral theorem. We will not prove the spectral theorem here.

One important property is that the the eigenvectors are orthogonal to each other if they have different eigenvalues. In the derivation below, we use a λ superscript to denote the eigenvalue associated with the eigenvector $\vec{e}^{\,\lambda}$. To see that the corresponding eigenvectors are orthogonal, note that the relation $\sum_j M_{ij}\vec{e}_j^{\,\lambda} = \lambda\vec{e}_i^{\,\lambda}$ can be rewritten as $\sum_j \vec{e}_j^{\,\lambda} M_{ji} = \lambda\vec{e}_i^{\,\lambda}$ because the matrix M is symmetric ($M_{ij} = M_{ji}$). So, let's compute $\sum_{ij} \vec{e}_i^{\ \lambda} M_{ij} \vec{e}_j^{\ \lambda'}$. By summing over j first, this becomes $\lambda' \vec{e}^{\ \lambda} \cdot \vec{e}^{\ \lambda'}$. But summing over i first, we have $\lambda \vec{e}^{\ \lambda} \cdot \vec{e}^{\ \lambda'}$. Of course, these two must be equal, so we learn that

$$(\lambda - \lambda')\vec{e}^{\ \lambda} \cdot \vec{e}^{\ \lambda'} = 0. \tag{27.4}$$

Hence, if $\lambda \neq \lambda'$, then we must have $\vec{e}^{\ \lambda} \cdot \vec{e}^{\ \lambda'} = 0$, implying the two vectors are orthogonal. It turns out that we can always find orthogonal eigenvectors when the eigenvalues are degenerate as well, so you can assume the eigenvectors are always orthogonal.

27.3 Examples

We begin with a simple example.

$$M = \begin{pmatrix} 10 & 6 \\ 6 & -10 \end{pmatrix} \tag{27.5}$$

M is real and symmetric, so it has 2 real roots (and hence, two eigenvectors). We find the eigenvalues by forming the shifted matrix

$$M - \lambda I = \begin{pmatrix} 10 - \lambda & 6 \\ 6 & -10 - \lambda \end{pmatrix} \tag{27.6}$$

and setting the determinant of this matrix to zero [$\det(M - \lambda I) = 0$]. Hence,

$$0 = (10 - \lambda)(-10 - \lambda) - 36 \tag{27.7}$$

$$0 = \lambda^2 - 100 - 36 = \lambda^2 - 136 \tag{27.8}$$

$$\lambda = \pm\sqrt{136} = \pm 2\sqrt{34}. \tag{27.9}$$

Now we find the eigenvectors using the procedure described above. First, we work out the eigenvectors for $\lambda = 2\sqrt{34}$:

$$\begin{pmatrix} 10 - 2\sqrt{34} & 6 \\ 6 & -10 - 2\sqrt{34} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = 0. \tag{27.10}$$

So, working with the top row equation, we have

$$(10 - 2\sqrt{34})a + 6b = 0, \tag{27.11}$$

which gives us

$$b = -\frac{1}{6}(10 - 2\sqrt{34})a. \tag{27.12}$$

And, as you see here, we only need to solve *one* not *two* equations, because the second equation is also immediately solved (this follows because the determinant was zero); check this directly if you are unsure. To determine both coefficients we must *normalize* the eigenvector. So we want $a^2 + b^2 = 1$, and then

$$a^2\left(1 + \frac{1}{36}(100 - 40\sqrt{34} + 136)\right) = 1. \tag{27.13}$$

Now, carrying out the rest of the algebra can get messy:

$$a^2 = \frac{36}{272 - 40\sqrt{34}} = \frac{18}{136 - 20\sqrt{34}} = \frac{9}{68 - 10\sqrt{34}} \tag{27.14}$$

$$a = \frac{3}{\sqrt{68 - 10\sqrt{34}}} \tag{27.15}$$

$$b = -\frac{1}{2}\frac{10 - 2\sqrt{34}}{\sqrt{68 - 10\sqrt{34}}}. \tag{27.16}$$

We change the sign of the square root term for the other eigenvector. You can then check that $\vec{e_1} \cdot \vec{e_2} = 0$. Summarizing, we have

$$\vec{e}_1 = \begin{pmatrix} 3 \\ -5 + \sqrt{34} \end{pmatrix} \frac{1}{\sqrt{68 - 10\sqrt{34}}} \tag{27.17}$$

and

$$\vec{e}_2 = \begin{bmatrix} 3 \\ -5 - \sqrt{34} \end{bmatrix} \frac{1}{\sqrt{68 - 10\sqrt{34}}}, \tag{27.18}$$

with $\lambda_1 = 2\sqrt{34}$ and $\lambda_2 = -2\sqrt{34}$. Note that one can essentially guess the second eigenvector by simply ensuring it is perpendicular to the first. Since we are working in a two-dimensional space, this is sufficient to determine the other eigenvector. One can also go through all the algebra again as well. If in doubt, you can do this, or simply check that $\vec{e}_1 \cdot \vec{e}_2 = 0$.

Here is another example. Find the eigenvalues and eigenvectors of the following matrix:

$$M = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 0 & -1 \\ 2 & -1 & -1 \end{pmatrix}. \tag{27.19}$$

We start by shifting and taking the determinant

$$\det(M - \lambda I) = 0, \tag{27.20}$$

which becomes

$$\det \begin{pmatrix} 1-\lambda & 0 & 2 \\ 0 & -\lambda & -1 \\ 2 & -1 & -1-\lambda \end{pmatrix} = 0. \tag{27.21}$$

There are many ways to evaluate the determinant. I like to do it the same way one calculates cross products, but many like to do it by minors. However you prefer to do it, you will end up with the result

$$\lambda - \lambda^3 + 4\lambda - 1 + \lambda = 0. \tag{27.22}$$

Simplifying, we have

$$\lambda^3 - 6\lambda + 1 = 0. \tag{27.23}$$

From here, we must go numerical to solve for the three different λ's and for the eigenvectors. This is often the case for these kinds of problems. You might have thought, wait this is a cubic, we can find the roots *analytically.* While true, often the formulas for the roots are so complicated that computing them numerically is better.

How do we make the transformation to "diagonalize" the matrix? It is a similarity transformation (which means multiply by one matrix on the left and by its inverse on the right). As before, we form P via

$$P = \begin{pmatrix} \cdots & — & e_1 & — & \cdots \\ \cdots & — & e_2 & — & \cdots \\ \vdots & & \vdots & & \vdots \\ \cdots & — & e_n & — & \cdots \end{pmatrix} \tag{27.24}$$

$$P^{-1} = \begin{pmatrix} | & | & \vdots & | \\ e_1 & e_2 & ... & e_n \\ | & | & \vdots & | \end{pmatrix} = P^T. \tag{27.25}$$

Then multiplying out and recalling that the eigenvectors are an orthonormal set yields

$$PMP^{-1} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ .0 & \lambda_2 & & \cdots \\ \vdots & & \vdots & \vdots \\ 0 & ... & 0 & \lambda_n \end{pmatrix}. \tag{27.26}$$

Knowing this allows us to form some interesting matrices.

For example,

$$e^M = P^{-1} \begin{pmatrix} e^{\lambda_1} & 0 & ... & 0 \\ .0 & e^{\lambda_2} & ... & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & e^{\lambda_{n-1}} & 0 \\ 0 & \cdots & 0 & e^{\lambda_n} \end{pmatrix} P \tag{27.27}$$

because

$$e^M = 1 + M + \frac{1}{2}M^2 + \frac{1}{6}M^3 + ... \tag{27.28}$$

$$= P^{-1}P\left(1 + M + \frac{1}{2}M^2 + \frac{1}{6}M^3 + ...\right)P^{-1}P \tag{27.29}$$

$$= P^{-1}\left(1 + PMP^{-1} + \frac{1}{2}(PMP^{-1})^2 + ...\right)P \tag{27.30}$$

$$= P^{-1} \begin{pmatrix} e^{\lambda_1} & 0 & ... & 0 \\ .0 & e^{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & e^{\lambda_{n-1}} & 0 \\ 0 & \cdots & 0 & e^{\lambda_n} \end{pmatrix} P \tag{27.31}$$

. Similarly, we can compute the absolute value of a matrix (or, if you like, the square root of the square of a matrix). Here, we choose the unique root that has all nonnegative eigenvalues, which is also called the absolute value of a matrix. It is given by

$$|M| = P^{-1} \begin{pmatrix} |\lambda_1| & 0 & \cdots & 0 \\ .0 & |\lambda_2| & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & |\lambda_{n-1}| & 0 \\ 0 & \cdots & 0 & |\lambda_n| \end{pmatrix} P. \tag{27.32}$$

Note that, in general after multiplying by P^{-1} and P, we have that the ij element of $|M|_{ij}$ is often not equal to $|M_{ij}|$.

This approach can obviously be extended to many different functional forms, and so one can think about constructing matrix-valued functions in this fashion. Usually the exponential and the absolute value (and the square root) are the three most important ones to consider.