EIGENVALUES AND EIGENVECTORS

Link to: physicspages home page.

To leave a comment or report an error, please use the auxiliary blog and include the title or URL of this post in your comment.

Post date: 3 Jan 2021.

In quantum mechanics, a physical state is represented by a vector in a vector space. Physically measurable quantities are represented by linear operators that operate on the state vector. If the state represents a system with a specific value of the physical quantity, applying the linear operator to the state results in that state being multiplied by the value of the quantity. Mathematically, such a state is called an *eigenvector* of the operator, and the numerical value that results is called the *eigenvalue*. The word 'eigen' is German for 'own', so an eigenvalue is a value 'owned' by the operator. Here, we'll examine eigenvalues and eigenvectors from a purely mathematical viewpoint, as it's useful to have an underlying understanding of the mathematics when applying it to quantum theory.

We start with a vector space V and an operator T. Suppose there is a one-dimensional subspace U of V which has the property that for any vector $u \in U$, $Tu = \lambda u$. That is, the operator T maps any vector u back into another vector in the same subspace U. In that case, U is said to be an *invariant subspace* under the operator T.

You can think of this in geometric terms. Suppose we have some n-dimensional vector space V, and a one-dimensional subspace U consisting of all vectors parallel to some straight line within V. Let the operator T acting on any vector u parallel to that line produce another vector which is also parallel to the same line. In other words, T multiplies a vector u in U by some number λ, which results in another vector λu parallel to u. Of course we can't push the geometric illustration too far, since in general V and U can be complex vector spaces, so the result of acting on u with T might give you some complex number λ multiplied by u.

The equation

$$Tu = \lambda u \tag{1}$$

is called an eigenvalue equation, and the number $\lambda \in \mathbb{F}$ is called the eigenvalue. The vector u itself is called the eigenvector corresponding to the eigenvalue λ. Since we can multiply both sides of this equation by any number c, any multiple of u is also an eigenvector corresponding to λ, so any vector 'parallel' to u is also an eigenvector. (I've put 'parallel' in quotes, since we're allowing for multiplication of u by complex as well as real numbers.)

It can happen that, for a particular value of λ, there are two or more linearly independent (that is, non-parallel) eigenvectors. In that case, the subspace spanned by the eigenvectors is two- or higher-dimensional.

Another way of writing 1 is by introducing the identity operator I:

$$(T-\lambda I)\,u=0 \tag{2}$$

If this equation has a solution other than $u=0$, then the operator $T-\lambda I$ has a non-trivial null space, which in turn means that $T-\lambda I$ is not injective (not one-to-one) and therefore not invertible. Also, the eigenvectors of T with eigenvalue λ are those vectors u in the null space of $T-\lambda I$.

An important result is

Theorem 1. *Suppose $\lambda_1,\ldots,\lambda_m$ are distinct eigenvalues of T and $v_1,\ldots,v_m$ are the corresponding non-zero eigenvectors. Then the set $v_1,\ldots,v_m$ is linearly independent.*

Proof. Suppose to the contrary that $v_1,\ldots,v_m$ is linearly dependent. Then there must be some subset that is linearly independent. Suppose that k is the smallest positive integer such that v_k can be written in terms of $v_1,\ldots,v_{k-1}$. That is, the set $v_1,\ldots,v_{k-1}$ is a linearly independent subset of $v_1,\ldots,v_m$. In that case, there are numbers $a_1,\ldots,a_{k-1}\in\mathbb{F}$ such that

$$v_k=\sum_{i=1}^{k-1}a_iv_i \tag{3}$$

If we apply the operator T to both sides and use the eigenvalue equation, we have

$$\begin{aligned} Tv_k &= \lambda_kv_k \qquad (4)\\ &= \sum_{i=1}^{k-1}a_iTv_i \qquad (5)\\ &= \sum_{i=1}^{k-1}a_i\lambda_iv_i \qquad (6)\end{aligned}$$

That is

$$\lambda_kv_k=\sum_{i=1}^{k-1}a_i\lambda_iv_i \tag{7}$$

We can multiply both sides of 3 by λ_k and subtract from 7 to get

$$(\lambda_k - \lambda_k) v_k = \sum_{i=1}^{k-1} a_i (\lambda_i - \lambda_k) v_i \tag{8}$$
$$= 0 \tag{9}$$

Since the set of vectors $v_1, \ldots, v_{k-1}$ is linearly independent, and $\lambda_k \neq \lambda_i$ for $i = 1, \ldots, k-1$, the only solution of this equation is $a_i = 0$ for $i = 1, \ldots, k-1$. But (from 3) this would make $v_k = 0$, contrary to our assumption that v_k is a non-zero eigenvector of T. Therefore the set $v_1, \ldots, v_m$ is linearly independent. □

It turns out that there are some operators on real vector spaces that don't have any eigenvalues. A simple example is the 2-dimensional vector space consisting of the xy plane. The rotation operator which rotates any vector about the origin (by some angle other than 2π) doesn't leave any vector parallel to itself and thus has no eigenvalues or eigenvectors.

However, in a complex vector space, things are a bit neater. This leads to the following theorem:

Theorem 2. *Every operator on a finite-dimensional, nonzero, complex vector space has at least one eigenvalue.*

Proof. Suppose V is a complex vector space with dimension $n > 0$. For some vector $v \in V$ we can write the $n+1$ vectors

$$v, Tv, T^2 v, \ldots, T^n v \tag{10}$$

Because we have $n+1$ vectors in an n-dimensional vector space, these vectors must be linearly dependent, which means we can find complex numbers $a_0, \ldots, a_n \in \mathbb{C}$, not all zero, such that

$$0 = a_0 v + a_1 T v + \ldots + a_n T^n v \tag{11}$$

We can consider a polynomial in z with the a_i as coefficients:

$$p(z) = a_0 + a_1 z + \ldots + a_n z^n \tag{12}$$

The Fundamental Theorem of Algebra states that any polynomial of degree n can be factored into n linear factors. In our case, the actual degree of $p(z)$ is $m \leq n$ since a_n could be zero. So we can factor $p(z)$ as follows:

$$p(z) = c(z - \lambda_1) \ldots (z - \lambda_m) \tag{13}$$

where $c \neq 0$.

Comparing this to 11, we can write that equation as

$$0 = a_0 v + a_1 T v + \ldots + a_n T^n v \tag{14}$$
$$= \left(a_0 I + a_1 T + \ldots + a_n T^n\right) v \tag{15}$$
$$= c\left(T - \lambda_1 I\right)\ldots\left(T - \lambda_m I\right) v \tag{16}$$

All the $T - \lambda_i I$ operators in the last line commute with each other since I commutes with everything and T commutes with itself, so in order for the last line to be zero, there has to be at least one λ_i such that $(T - \lambda_i I)\, v = 0$. That is, there is at least one λ_i such that $T - \lambda_i I$ has a nonzero null space, which means λ_i is an eigenvalue. □

References

(1) Axler, Sheldon (2015), *Linear Algebra Done Right*, 3rd edition, Springer. Chapter 5.
(2) Liesen, Jörg & Mehrmann, Volker (2015), *Linear Algebra,* Chapter 8.
(3) Landi, Giovanni & Zampini, Alessandro (2018), *Linear Algebra and Analytic Geometry for Physical Sciences*, Chapter 9.
(4) Shankar, R. (1994), *Principles of Quantum Mechanics*, Plenum Press, Chapter 1.

Pingbacks

Pingback: Hermitian operators
Pingback: Spectral theorem for normal operators
Pingback: Momentum eigenfunctions and eigenvalues