SPECTRAL THEOREM FOR NORMAL OPERATORS

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: 5 Jan 2021.

We'll now look at a central theorem about normal operators, known as the *spectral theorem*.

We've seen that if a matrix M has a set v of eigenvectors that span the space, then we can diagonalize M by means of the similarity transformation

$$D_M = A^{-1} M A \tag{1}$$

where D_T is diagonal and the columns of A are the eigenvectors of M. In the general case, there's no guarantee that the eigenvectors of M are orthonormal. However, if there *is* an orthonormal basis in which M is diagonal, then M is said to be *unitarily diagonalizable*. Suppose we start with some arbitrary orthonormal basis $(e_1, \ldots, e_n)$ (we can always construct such a basis using the Gram-Schmidt procedure). Then if the set of eigenvectors of M form an orthonormal basis $(u_1, \ldots, u_n)$, there is a unitary matrix U that transforms the e_i basis into the u_i basis (since unitary operators preserve inner products):

$$u_i = U e_i = \sum_j U_{ji} e_j \tag{2}$$

Using this unitary operator, we therefore have for a unitarily diagonalizable operator M

$$D_M = U^{-1} M U = U^{\dagger} M U \tag{3}$$

The *spectral theorem* now states:

Theorem 1. *An operator M in a complex vector space has an orthonormal basis of eigenvectors (that is, it's unitarily diagonalizable) if and only if M is normal.*

Proof. Since this is an 'if and only if' theorem, we need to prove it in both directions. First, suppose that M is unitarily diagonalizable, so that 3 holds for some U. Then

$$M = UD_M U^\dagger \tag{4}$$

$$M^\dagger = UD_M^\dagger U^\dagger \tag{5}$$

The commutator is then, since $U^\dagger U = I$

$$\left[M^\dagger, M\right] = UD_M^\dagger D_M U^\dagger - UD_M D_M^\dagger U^\dagger \tag{6}$$

$$= U\left[D_M^\dagger, D_M\right]U^\dagger \tag{7}$$

$$= 0 \tag{8}$$

where the result follows because all diagonal matrices commute and, since D_M is diagonal, so is its hermitian conjugate $D_M^\dagger$. Thus M is normal, and this completes one direction of the proof.

Going the other way is a bit trickier. We need to show that for any normal matrix M with elements defined on some arbitrary orthonormal basis (that is, a basis that is not necessarily composed of eigenvectors of M), there is a unitary matrix U such that $U^\dagger MU$ is diagonal. Since we started with an orthonormal basis and U preserves inner products, the new basis is also orthonormal which will prove the theorem.

The proof uses mathematical induction, in which we first prove that the result is true for one specific dimension of vector space, say dim $V = 1$. We can then assume the result is true for some dimension $n-1$ and from that assumption, prove it is also true for the next higher dimension n.

Since any 1×1 matrix is diagonal (it consists of only one element), the result is true for dim $V = 1$. So we now assume it's true for a dimension of $n-1$ and prove it's true for a dimension of n.

We take an arbitrary orthonormal basis of the n-dimensional vector space V to be $(|1\rangle, \ldots, |n\rangle)$. In that basis, the matrix M has elements $M_{ij} = \langle i|M|j\rangle$. We know that M has at least one eigenvalue λ_1 with a normalized eigenvector $|x_1\rangle$:

$$M|x_1\rangle = \lambda_1 |x_1\rangle \tag{9}$$

and, since M is normal, the eigenvector $|x_1\rangle$ is also an eigenvector of $M^\dagger$:

$$M^\dagger |x_1\rangle = \lambda_1^* |x_1\rangle \tag{10}$$

Starting with a basis of V containing $|x_1\rangle$, we can use Gram-Schmidt to generate an orthonormal basis $(|x_1\rangle, \ldots, |x_n\rangle)$. We now define an operator U_1 as follows:

$$U_1 \equiv \sum_i |x_i\rangle \langle i| \tag{11}$$

where $|i\rangle$ is a vector in the original orthonormal basis defined above. U_1 is unitary, since

$$\begin{aligned} U_1^\dagger &= \sum_i |i\rangle \langle x_i| & (12)\\ U_1^\dagger U &= \sum_i \sum_j |i\rangle \langle x_i | x_j\rangle \langle j| & (13)\\ &= \sum_i \sum_j |i\rangle \delta_{ij} \langle j| & (14)\\ &= \sum_i |i\rangle \langle i| & (15)\\ &= I & (16) \end{aligned}$$

From its definition

$$\begin{aligned} U_1 |1\rangle &= |x_1\rangle & (17)\\ U_1^\dagger |x_1\rangle &= |1\rangle & (18) \end{aligned}$$

Now consider the matrix M_1 defined as

$$M_1 \equiv U_1^\dagger M U_1 \tag{19}$$

M_1 is also normal, as can be verified by calculating the commutator and using $\left[M^\dagger, M\right] = 0$. Further

$$\begin{aligned} M_1 |1\rangle &= U_1^\dagger M U_1 |1\rangle & (20)\\ &= U_1^\dagger M |x_1\rangle & (21)\\ &= \lambda_1 U_1^\dagger |x_1\rangle & (22)\\ &= \lambda_1 |1\rangle & (23) \end{aligned}$$

where we used 9 to get the third line. Thus $|1\rangle$ is an eigenvector of M_1 with eigenvalue λ_1.

The matrix elements in the first column of M_1 in the original basis $(|1\rangle, \ldots, |n\rangle)$ are

$$\langle j | M_1 | 1\rangle = \lambda_1 \langle j | 1\rangle = \lambda_1 \delta_{1j} \tag{24}$$

Thus all entries in the first column are zero except for the first row, where it is λ_1. How about the first row? Using 10 we have

$$
\begin{aligned}
\langle 1 | M_1 | j \rangle &= \left(\left\langle j \middle| M_1^\dagger \middle| 1 \right\rangle \right)^* && (25) \\
&= \left(\lambda_1^* \langle j | 1 \rangle \right)^* && (26) \\
&= \lambda_1 \delta_{1j} && (27)
\end{aligned}
$$

Thus all entries in the first row, except the first, are also zero. Thus in the original basis $(|1\rangle, \ldots, |n\rangle)$ we have

$$
M_1 = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & M' & \\ 0 & & & \end{bmatrix} \quad (28)
$$

where M' is an $(n-1) \times (n-1)$ matrix. We have

$$
\begin{aligned}
M_1^\dagger &= \begin{bmatrix} \lambda_1^* & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & (M')^\dagger & \\ 0 & & & \end{bmatrix} && (29) \\
M_1^\dagger M_1 &= \begin{bmatrix} |\lambda_1|^2 & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & (M')^\dagger M' & \\ 0 & & & \end{bmatrix} && (30) \\
M_1 M_1^\dagger &= \begin{bmatrix} |\lambda_1|^2 & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & M' (M')^\dagger & \\ 0 & & & \end{bmatrix} && (31)
\end{aligned}
$$

Since M_1 is normal, we must have $M_1 M_1^\dagger = M_1^\dagger M_1$, which implies

$$
\left(M'\right)^\dagger M' = M' \left(M'\right)^\dagger \quad (32)
$$

so that M' is also a normal matrix. By the induction hypotheses, since M' is an $(n-1) \times (n-1)$ normal matrix, it is unitarily diagonalizable by some unitary matrix U', that is

$$
U'^\dagger M' U' = D_{M'} \quad (33)
$$

is diagonal. We can extend U' to an $n \times n$ unitary matrix by adding a 1 to the upper left:

$$U = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & U' & \\ 0 & & & \end{bmatrix} \tag{34}$$

We can check that U is unitary by direct calculation, using $(U')^\dagger U' = I$

$$U^\dagger U = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & (U')^\dagger & \\ 0 & & & \end{bmatrix} \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & U' & \\ 0 & & & \end{bmatrix} \tag{35}$$

$$= \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & (U')^\dagger U' & \\ 0 & & & \end{bmatrix} \tag{36}$$

$$= \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & & \\ \vdots & & 1 & \\ 0 & & & 1 \end{bmatrix} = I \tag{37}$$

We then have, using 33

$$U^\dagger M_1 U = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & (U')^\dagger & \\ 0 & & & \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & M' & \\ 0 & & & \end{bmatrix} \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & U' & \\ 0 & & & \end{bmatrix} \tag{38}$$

$$= \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & U'^\dagger M' U' & \\ 0 & & & \end{bmatrix} \tag{39}$$

$$= \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & D_{M'} & \\ 0 & & & \end{bmatrix} \tag{40}$$

That is, $U^\dagger M_1 U$ is diagonal. From the definition 19 of M_1, we now have

$$U^{\dagger}M_1U = U^{\dagger}U_1^{\dagger}MU_1U \tag{41}$$
$$= (U_1U)^{\dagger}M(U_1U) \tag{42}$$

Since the product of two unitary matrices is unitary, we have found a unitary operator U_1U that diagonalizes M, which proves the result. □

Notice that the proof didn't assume that the eigenvalues are nondegenerate, so that even if there are several linearly independent eigenvectors corresponding to one eigenvalue, it is still possible to find an orthonormal basis consisting of the eigenvectors. In other words, for any hermitian or unitary operator, it is always possible to find an orthonormal basis of the vector space consisting of eigenvectors of the operator.

In the general case, a normal matrix M in an n-dimensional vector space can have m distinct eigenvalues, where $1 \le m \le n$. If $n = m$, there is no degeneracy and each eigenvalue has a unique (up to a scalar multiple) eigenvector. If $m < n$, then one or more of the eigenvalues occurs more than once, and the eigenvector subspace corresponding to a degenerate eigenvalue has a dimension larger than 1. However, the spectral theorem guarantees that it is possible to choose an orthonormal basis within each subspace, and that each subspace is orthogonal to all other subspaces.

More precisely, the vector space V can be decomposed into m subspaces U_k for $k = 1,\ldots,m$, with the dimension d_k of subspace U_k equal to the degeneracy of eigenvalue λ_k. The full space V is the direct sum of these subspaces

$$V = U_1 \oplus U_2 \oplus \ldots \oplus U_m \tag{43}$$
$$n = \sum_{k=1}^{m} d_k \tag{44}$$

It's usually most convenient to order the eigenvectors as follows:

$$\left(u_1^{(1)},\ldots,u_{d_1}^{(1)},u_1^{(2)},\ldots,u_{d_2}^{(2)},\ldots,u_1^{(m)},\ldots,u_{d_m}^{(m)}\right) \tag{45}$$

The notation $u_j^{(k)}$ means the jth eigenvector belonging to eigenvalue k.

In practice, there is a lot of freedom in choosing orthonormal eigenvectors for degenerate eigenvalues, since we can pick any d_k mutually orthogonal vectors within the subspace of dimension d_k. For example, in 3-d space we usually choose the x,y,z unit vectors as the orthonormal set, but we can pivot these three vectors about the origin, or even reflect them in a plane passing through the origin, and still get an orthonormal set of 3 vectors.

The diagonal form of the normal matrix M in this orthonormal basis is

$$D_M=\begin{bmatrix}\lambda_1 & & & & & & \\ & \ddots & & & & & \\ & & \lambda_1 & & & & \\ & & & \ddots & & & \\ & & & & \lambda_m & & \\ & & & & & \ddots & \\ & & & & & & \lambda_m\end{bmatrix} \tag{46}$$

Here, eigenvalue λ_k occurs d_k times along the diagonal.

REFERENCES

(1) Axler, Sheldon (2015), *Linear Algebra Done Right*, 3rd edition, Springer. Chapter 7.
(2) Zwiebach, Barton, Online course *Mastering Quantum Mechanics Part 1: Wave Mechanics*. Archive available here.
(3) Landi, Giovanni & Zampini, Alessandro (2018), *Linear Algebra and Analytic Geometry for Physical Sciences*, Chapter 12.
(4) Shankar, R. (1994), *Principles of Quantum Mechanics*, Plenum Press, Chapter 1.

PINGBACKS

Pingback: Functions of hermitian operators
Pingback: Functions of hermitian operators
Pingback: Determinant and trace of normal operators
Pingback: Hermitian matrices - example with 4 matrices
Pingback: Simultaneous diagonalization of hermitian matrices
Pingback: Coupled masses on springs - a solution using matrix diagonalization
Pingback: Unitarity theorem for groups
Pingback: Schur's lemma