Spectral theorems for Hermitian and unitary operators

A. Eremenko

March 18, 2023

1. An *Hermitian product* on a complex vector space V is an assignment of a complex number (x, y) to each pair of vectors x, y, which has the following properties for all vectors x, y, z and for all numbers α, β:

$$(x, y) = \overline{(y, x)},$$

$$(x, \alpha y + \beta z) = \alpha(x, y) + \beta(x, z),$$

$$(x, x) \geq 0,$$

with equality only for $x = 0$.

Example. $(x, y) = \overline{x_1} y_1 + \ldots + \overline{x_n} y_n$. This example is called the *standard Hermitian product* on $\mathbf{C}^n$.

It follows from the first two properties that $(\alpha x, y) = \overline{\alpha}(x, y)$. They say that (x, y) is linear with respect to the second argument and *anti-linear* with respect to the first one.

An *Hermitian transposition* is the combination of two operations: ordinary transposition and complex conjugation. It is denoted by star, $A^* = \overline{A}^T$, where the bar is the complex conjugation. So the standard Hermitian product can we written as $(x, y) = x^* y$.

Two vectors are called *orthogonal* if $(x, y) = 0$. Vectors orthogonal to some given set of vectors form a subspace. If V' is a subspace of V then its *orthogonal complement* consists of all vectors orthogonal to each vector of V. Two subspaces are called orthogonal if each vector of one of them is orthogonal to each vector of another one.

A square matrix A is called *Hermitian* if

$$A^* = A.$$

A real matrix is Hermitian if and only if it is symmetric. Hermitian matrices are characterized by the property

$$(Ax, y) = (x, Ay), \quad \text{for all} \quad x, y \quad \text{in} \quad V, \tag{1}$$

where $(.,.)$ is the standard Hermitian product. Indeed, $A^* = A$ is equivalent to

$$(Ax, y) = (Ax)^* y = x^* A y = (x, Ay), \quad \text{for all} \quad x, y \quad \text{in} \quad V.$$

A square matrix U is called *unitary* if

$$U^* U = I,$$

which is the same as $U^* = U^{-1}$. In other words, a unitary matrix is such that its columns are orthonormal. Unitary matrices are characterized by the property

$$(Ux, Uy) = (x, y) \quad \text{for all} \quad x, y \quad \text{in} \quad V. \tag{2}$$

Indeed,

$$(Ux, Uy) = (Ux)^* Uy = x^* U^* U y = x^* y = (x, y).$$

A real matrix is unitary if and only if it is orthogonal.

We recall that each $n \times n$ matrix defines a linear operator on $\mathbf{C}^n$ acting by the rule $L(x) = Ax$. And conversely, each linear operator in a finite-dimensional vector space is described by a matrix. This correspodence between matrices and linear operators depends on the choice of a basis.

2. **Spectral theorem for Hermitian matrices.** *For an Hermitian matrix:*

a) all eigenvalues are real,

b) eigenvectors corresponding to distinct eigenvalues are orthogonal,

c) there exists an orthogonal basis of the whole space, consisting of eigenvectors.

Thus all Hermitian matrices are diagonalizable. Moreover, for every Hermitian matrix A, there exists a unitary matrix U such that

$$AU = U\Lambda,$$

where Λ is a real diagonal matrix. The diagonal entries of Λ are the eigenvalues of A, and columns of U are eigenvectors of A.

Proof of Theorem 2. a). Let λ be an eigenvalue, then

$$Ax = \lambda x, \quad x \neq 0$$

for some vector x. Multiply both sides on x:

$$(Ax, x) = (\lambda x, x) = \overline{\lambda}(x, x).$$

Property (1) shows that (Ax, x) equals

$$(x, Ax) = (x, \lambda x) = \lambda(x, x).$$

As $(x, x) \neq 0$, we conclude that $\lambda = \overline{\lambda}$, that is λ is real. This proves a).

Proof of b). Suppose we have two distinct eigenvalues $\lambda \neq \mu$. Then

$$Ax = \lambda x, \quad Ay = \mu y, \tag{3}$$

where x, y are eigenvectors. Multiply the first equation on y, use (1) and the fact that λ is real which was just established.

$$\lambda(x, y) = (\lambda x, y) = (Ax, y) = (x, Ay) = (x, \mu y) = \mu(x, y).$$

As $\lambda \neq \mu$, we conclude that $(x, y) = 0$, which proves b).

Proof of c). Let λ_1 be an eigenvalue, and x_1 an eigenvector corresponding to λ_1 (every square matrix has an eigenvalue and an eigenvector). Let V_1 be the set of all vectors orthogonal to x_1. Then A maps V_1 into itself: for every $x \in V_1$ we also have $Ax \in V_1$. Indeed, $x \in V_1$ means that $(x_1, x) = 0$, then we have using (1):

$$(x_1, Ax) = (Ax_1, x) = \lambda_1(x_1, x) = 0,$$

so $x \in V_1$. Now the linear operator $L(x) = Ax$ when restricted to V_1 is also Hermitian, and it has an eigenvalue λ_2 and an eigenvector $x_2 \in V_1$. By definition of V_1, x_2 is orthogonal to x_1. Let V_2 be the orthogonal complement of the span of x_1, x_2. Then A also maps V_2 into itself, as before. Continuing this way, we find a sequence λ_k, x_k and subspaces V_k containing x_k such that V_k is orthogonal to $x_1, \ldots, x_{k-1}$. The sequence must terminate on the n-th step because $\dim V_k = n - k$: on every step dimension decreases by 1. This completes the proof.

3. **Spectral theorem for unitary matrices.** *For a unitary matrix:*

a) all eigenvalues have absolute value 1.

b) eigenvectors corresponding to distinct eigenvalues are orthogonal,

c) there is an orthogonal basis of the whole space, consisting of eigenvectors.

Thus unitary matrices are diagonalizable. Moreover, for each unitary matrix A there exists a unitary matrix U such that

$$AU = U\Lambda$$

where U is a diagonal matrix whose diagonal entries have absolute value 1. The columns of U are eigenvectors of A.

Proof of Theorem 2. a) Let λ be an eigenvalue. Then

$$Ax = \lambda x, \quad x \neq 0.$$

Using (2) we obtain

$$(x, x) = (Ax, Ax) = \overline{\lambda}\lambda(x, x).$$

As $(x, x) \neq 0$, we conclude that $\overline{\lambda}\lambda = |\lambda|^2 = 1$, which proves a).

Proof of b). Begin with (2), (3), and obtain

$$(x, y) = (Ax, Ay) = \overline{\lambda}\mu(x, x).$$

As $|\lambda| = 1$, we conclude that $\overline{\lambda} = \lambda^{-1}$, so the multiple in the RHS is $\mu/\lambda \neq 1$ by our assumption that $\mu \neq \lambda$. So $(x, y) = 0$, which proves b).

Proof of c). Let λ_1 be an eigenvalue, and x_1 an eigenvector corresponding to this eigenvalue, Let V_1 be the set of all vectors orthogonal to x_1. As in the proof in section 2, we show that $x \in V_1$ implies that $Ax \in V_1$. Indeed

$$(Ax, x_1) = (x, A^*x_1) = (x, A^{-1}x_1) = \lambda^{-1}(x, x_1) = 0,$$

where we used (2) which is equivalent to $A^* = A^{-1}$. The proof is now completed in exactly the same way as in the previous section.

4. **Exponentials of Hermitian matrices.** *Let A be an Hermitian matrix. Then e^{iA} is unitary, and conversely, every unitary matrix has the form e^{iA} for some Hermitian matrix A.*

Let B be a real matrix, and $A = iB$. Then A is Hermitian if and only if B is skew symmetric ($B^T = -B$):

$$A^* = (-i)B^T = iB = A.$$

So we obtain a

Corollary: For a real matrix B, e^B is orthogonal if and only if B is skew-symmetric.

Proof. Let $U = e^{iA}$, where A is Hermitian. Then

$$UU^* = e^{iA}e^{-iA^*} = e^{iA}e^{-iA} = I.$$

Conversely, let U be a unitary matrix. Then, by the Spectral Theorem for unitary matrices (section 3), there is another unitary matrix B such that $U = B\Lambda B^{-1}$, and $\Lambda = \mathrm{diag}\,(\lambda_1, \ldots, \lambda_n)$. As all $|\lambda_k| = 1$, we write them as $\lambda_k = e^{i\theta_k}$, where θ_k are real numbers. Then set

$$A = B\,\mathrm{diag}\,(\theta_1, \ldots, \theta_n)B^{-1} = B\Lambda_1 B^{-1}.$$

Then A is Hermitian:

$$A^* = (B^{-1})^*\Lambda_1 B^* = B\Lambda_1 B^{-1} = A,$$

and evidently $\exp(iA) = U$.

5. These three theorems can be generalized to infinite-dimensional spaces. Unlike the Jordan form theorem. One can say that we understand well Hermitian and unitary operators, but not arbitrary linear operators.

These three theorems and their infinite-dimensional generalizations make the mathematical basis of the most fundamental theory about the real world that we possess, namely quantum mechanics.

6. Normal operators. According to part c) of our spectral theorems, if A is either Hermitian or unitary then there is an orthonormal basis consisting of eigenvectors. Let us describe all operators with this property. If there is an orthonormal basis of eigenvectors of A then

$$A = U\Lambda U^{-1} = U\Lambda U^*, \tag{4}$$

where columns of U are eigenvectors of our basis, and the second equation holds because U is unitary, $U^{-1} = U^*$. From (4) we conclude that

$$A^* = U\Lambda^* U^* = U\Lambda^* U^{-1}. \tag{5}$$

Notice that all pairs of diagonal matrices commute $\Lambda_1\Lambda_2 = \Lambda_2\Lambda_1$, and we conclude from (4) and (5) that

$$AA^* = A^*A.$$

Operators and matrices with this property are called *normal*. We just proved that existence of a basis of eigenvectors implies normality. Now we prove the converse.

For each normal operator A, there exists an orthonormal basis of the space consisting of eigenvectors.

The proof is similar to the proof of c) for Hermitian and unitary operators.

Let λ_1 be some eigenvalue, and V_1 the corresponding eigenspace. By definition, V_1 consists of all vectors x such that $Ax = \lambda_1 x$. Let U_1 be the orthogonal complement of V_1. By definition, U_1 consists of all vectors y such that

$$(x, y) = 0 \quad \text{for all} \quad x \in V_1. \tag{6}$$

Let us prove that A^* maps V_1 into itself. Suppose $y \in V_1$ we want to prove that $A^* y \in V_1$. We have

$$A(A^*y) = A^*Ay = A^*(\lambda y) = \lambda(A^*y),$$

thus $Ay \in V_1$ as advertised.

Now we prove that A maps U_1 to itself. That is that (6) implies

$$(x, Ay) = 0 \quad \text{for all} \quad x \in V_1.$$

We have

$$(x, Ay) = (A^*x, y) = 0,$$

because $x \in V_1$ implies $A^*x \in V_1$ as we have seen before.

Now we show that A^* also maps U_1 into itself. Indeed, if $(y, x) = 0$ for all $x \in V_1$, then for all $x \in V_1$ we have'

$$(A^*y, x) = (y, Ax) = \lambda_1(y, x) = 0,$$

So $A^*y \in U_1$.

So the restriction of A on U_1 is also normal, and the proof ends with an induction as in the proof of c) in previous theorems.

7. Orthogonal projections In general, a *projector* is an operator P with the property

$$P^2 = P. \tag{7}$$

Let V be the column space and U be the null-space. Equation (7) means that P acts as the identity on V. Now (7) also implies that $U \cap V = \{0\}$.

Indeed, if $x \in U \cap V$ then we have $Px = 0$ and $x = Py$ for some y. Then $0 = Px = P^2 y = Py = x$ by (7), so $x = 0$.

So the whole space is the direct sum of U and V, which means that every vector x has a *unique* representation

$$x = u + v, \quad \text{where} \quad u \in U \quad \text{and} \quad v \in V. \tag{8}$$

Operator P collapses U to $\{0\}$ and acts as the identity on V. In other words, for an x as in (8), $Px = v$.

A projector is called an *orthoprojector* ("orthogonal projector") if in addition to (7) it is Hermitian,

$$P^* = P. \tag{9}$$

A projector is Hermitian if and only if U is orthogonal to V, which together with (8) implies that U and V are orthogonal complements of each other. Indeed, let $x \in U$ and $y \in V$. Then $y = Pz$ for some z and $Px = 0$ by definition of U and V. So

$$(x, y) = (x, Pz) = (Px, y) = 0.$$

Exercise. Previously we derived a formula for the orthoprojector onto the column space of a (rectangular) matrix A with linearly independent columns:

$$P = A(A^* A)^{-1} A^*.$$

Show that this P has properties (1) and (9).

Exercise. Let P_1 and P_2 be two orthoprojectors. Show that $P_1 P_2 = P_2 P_1 = 0$ if and only if the subspaces V_1, V_2 on which they project are orthogonal.

Exercise. Show that every normal operator A can be written in the form

$$A = \lambda_1 P_1 + \ldots + \lambda_k P_k,$$

where λ_1, λ_k are all eigenvalues, and P_j is the orthoprojector onto the eigenspace corresponding to λ_j. Moreover, these orthoprojectors P_j satisfy

$$\sum_{j=1}^{k} P_j = I, \quad \text{and} \quad P_i P_j = 0 \quad \text{for all} \quad i, j.$$

This representation of a normal operator A is called the *spectral decomposition.* The operator A is Hermitian when all λ_j are real, and unitary when all λ_j have absolute value 1.