Chapter 2: Linear Algebra for Quantum Theory

Quantum information is written in the language of linear algebra. A qubit is represented by a vector in a complex vector space. A quantum gate is represented by a special kind of matrix. A measurement outcome is represented, in the simplest case, by a projection matrix. Later, a generalized measurement will be represented by a family of positive matrices whose sum is the identity.

This chapter builds the finite-dimensional toolkit we need before speaking carefully about Hilbert spaces, states, POVMs, and Naimark dilation. We will move slowly. The goal is not only to recognize the words, but to understand what the objects do.

Throughout most of this book, our vector spaces are finite-dimensional and complex. This is the standard setting for introductory quantum information, where systems such as qubits and finite registers are modeled using finite-dimensional complex Hilbert spaces (Nielsen and Chuang, 2010; Watrous, 2018).

2.1 Vectors and vector spaces

A vector space is a collection of objects called vectors that can be added together and multiplied by numbers called scalars.

In ordinary two-dimensional geometry, vectors may be arrows in the plane. For example,

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

can be drawn as an arrow pointing 2 units to the right and 1 unit upward. But in quantum theory, vectors are usually not physical arrows in ordinary space. They are elements of an abstract complex vector space.

The scalars in this book are usually complex numbers. A complex number has the form

$$a + bi,$$

where a,b∈ mathbbR and i²=-1. The complex conjugate of a+bi is

$$\overline{a+bi} = a - bi.$$

The most common finite-dimensional complex vector space is

$$\mathbb{C}^n = \left\{ \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix} : z_1, \ldots, z_n \in \mathbb{C} \right\}.$$

For example,

$$\begin{pmatrix} 1+i \\ 2 \\ -i \end{pmatrix} \in \mathbb{C}^3.$$

A vector space must allow two basic operations:

1. Vector addition:

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} + \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \end{pmatrix}.$$

2. Scalar multiplication:

$$(2-i)\begin{pmatrix} 1 \\ i \end{pmatrix} = \begin{pmatrix} 2-i \\ (2-i)i \end{pmatrix} = \begin{pmatrix} 2-i \\ 1+2i \end{pmatrix}.$$

These operations must satisfy familiar algebraic rules: addition is associative and commutative, there is a zero vector, every vector has an additive inverse, and scalar multiplication distributes over addition. These are the usual vector space axioms studied in linear algebra (Axler, 2015).

In quantum information, a vector often represents a possible pure state of a system, although not every nonzero vector represents a distinct physical state. For example, multiplying a quantum state vector by a nonzero complex scalar changes the vector but not necessarily the physical ray it represents. We will return to this subtlety in Chapter 4.

For now, think of vectors as the raw objects on which matrices act.

2.2 Linear combinations, span, and subspaces

A linear combination of vectors is a sum made by multiplying vectors by scalars and adding the results.

If $v_1, v_2, \ldots, v_k$ are vectors and $c_1, c_2, \ldots, c_k \in$ mathbbC, then

$$c_1 v_1 + c_2 v_2 + \cdots + c_k v_k$$

is a linear combination of $v_1, \ldots, v_k$.

For example, in mathbbC², let

$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Then

$$3v_1 + (2+i)v_2 = 3\begin{pmatrix} 1 \\ 0 \end{pmatrix} + (2+i)\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 2+i \end{pmatrix}.$$

The span of a set of vectors is the set of all possible linear combinations of those vectors. We write

$$\operatorname{span}\{v_1, \ldots, v_k\}.$$

For example,

$$\operatorname{span}\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right\} = \left\{\begin{pmatrix} c \\ 0 \end{pmatrix} : c \in \mathbb{C}\right\}.$$

This is the horizontal axis inside mathbbC², except that the coordinates are complex.

A subspace is a subset of a vector space that is itself a vector space under the same addition and scalar multiplication. A subset $W \subseteq V$ is a subspace if:

1. $0 \in W$,

2. whenever u,v∈ W, then u+v∈ W,

3. whenever v∈ W and c∈mathbbC, then cv∈ W.

For example,

$$W = \left\{ \begin{pmatrix} z \\ 0 \end{pmatrix} : z \in \mathbb{C} \right\}$$

is a subspace of mathbbC². But

$$S = \left\{ \begin{pmatrix} z \\ 1 \end{pmatrix} : z \in \mathbb{C} \right\}$$

is not a subspace, because it does not contain the zero vector.

Subspaces matter because projections, measurements, and eigenspaces are all built from them. A projective measurement asks, in effect, which mutually orthogonal subspace the state belongs to or overlaps with.

2.3 Linear independence, bases, and dimension

A list of vectors v₁,…,vₖ is linearly independent if the only way to make the zero vector as a linear combination is the trivial way:

$$c_1 v_1 + \cdots + c_k v_k = 0 \quad \Longrightarrow \quad c_1 = \cdots = c_k = 0.$$

If there is a nontrivial choice of coefficients, not all zero, producing the zero vector, then the vectors are linearly dependent.

For example, in mathbbC²,

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

are linearly independent. If

$$c_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

then

$$\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

so $c_1=c_2=0$.

But the vectors

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

are linearly dependent because

$$2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

A basis of a vector space V is a list of vectors that is both:

1. linearly independent, and
2. spans V.

A basis gives coordinates. For mathbbC^n, the standard basis is

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Every vector

$$v = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}$$

can be written uniquely as

$$v = z_1 e_1 + z_2 e_2 + \cdots + z_n e_n.$$

The number of vectors in any basis of a finite-dimensional vector space is called the dimension of the space. It is a theorem of linear algebra that every basis of a finite-dimensional vector space has the same number of vectors (Axler, 2015). Thus mathbbCⁿ has dimension n.

In quantum information, the dimension of the Hilbert space is the number of perfectly distinguishable basis states available to the system. A qubit has a two-dimensional Hilbert space. A pair of qubits has a four-dimensional Hilbert space. More generally, m qubits have dimension 2ᵐ (Nielsen and Chuang, 2010).

2.4 Inner products: geometry for complex vector spaces

Linear algebra becomes geometry when we can speak about lengths and angles. This is done using an inner product.

For vectors x,y∈mathbbCⁿ, we use the standard inner product

$$\langle x, y \rangle = x^\dagger y = \overline{x_1} y_1 + \overline{x_2} y_2 + \cdots + \overline{x_n} y_n.$$

Here x^† means the conjugate transpose of x: turn the column vector into a row vector and conjugate each entry.

For example, let

$$x = \begin{pmatrix} 1+i \\ 2 \end{pmatrix}, \qquad y = \begin{pmatrix} 3 \\ i \end{pmatrix}.$$

Then

$$\langle x, y\rangle = \overline{1+i}\cdot 3 + \overline{2}\cdot i = (1-i)3 + 2i = 3 - 3i + 2i = 3 - i.$$

In this book, the inner product is conjugate-linear in the first input and linear in the second input:

$$\langle ax_1 + bx_2, y\rangle = \overline{a}\langle x_1, y\rangle + \overline{b}\langle x_2, y\rangle,$$

while

$$\langle x, ay_1 + by_2\rangle = a\langle x, y_1\rangle + b\langle x, y_2\rangle.$$

This convention matches the common column-vector expression ⟨ x,y⟩=x^† y, widely used in quantum information (Nielsen and Chuang, 2010; Watrous, 2018).

An inner product must satisfy three essential properties:

1. Conjugate symmetry:

$$\langle x, y\rangle = \overline{\langle y, x\rangle}.$$

2. Linearity in the second input:

$$\langle x, ay + bz\rangle = a\langle x, y\rangle + b\langle x, z\rangle.$$

3. Positive definiteness:

$$\langle x, x\rangle \geq 0,$$

and

$$\langle x, x\rangle = 0 \quad \text{if and only if} \quad x = 0.$$

The norm or length of a vector is

$$\|x\| = \sqrt{\langle x, x\rangle}.$$

For example,

$$x = \begin{pmatrix} 1+i \\ 2 \end{pmatrix}$$

has length

$$\|x\| = \sqrt{|1+i|^2 + |2|^2} = \sqrt{2+4} = \sqrt{6}.$$

A vector x is called a unit vector if

$$\|x\| = 1.$$

Quantum pure states are usually represented by unit vectors, because probabilities computed from them must sum to 1.

2.5 Orthogonality and orthonormal bases

Two vectors x and y are orthogonal if

$$\langle x, y \rangle = 0.$$

Orthogonality means that the vectors are perpendicular in the geometry determined by the inner product.

For example, in mathbbC²,

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

are orthogonal because

$$\langle e_1, e_2 \rangle = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0.$$

A list of vectors $u_1,\ldots,u_n$ is orthonormal if each vector has length 1 and different vectors are orthogonal:

$$\langle u_i, u_j \rangle = \delta_{ij},$$

where

$$\delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

An orthonormal basis is a basis that is also orthonormal.

Orthonormal bases are especially useful because coordinates become inner products. If $u_1,\ldots,u_n$ is an orthonormal basis of V, then every vector $v \in V$ can be written as

$$v = \sum_{j=1}^{n} \langle u_j, v \rangle u_j.$$

For example, in $\mathbb{C}^2$, with the standard basis e_1, e_2,

$$v = \begin{pmatrix} 3 \\ 2+i \end{pmatrix} = \langle e_1, v \rangle e_1 + \langle e_2, v \rangle e_2 = 3e_1 + (2+i)e_2.$$

The formula is simple because the basis vectors do not overlap with each other.

In quantum mechanics, an orthonormal basis often represents a set of perfectly distinguishable alternatives. For a qubit, the computational basis is usually written as

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The notation $|0\rangle$ and $|1\rangle$ is called Dirac notation or bra-ket notation. A column vector $|\psi\rangle$ is called a ket. Its conjugate transpose $\langle \psi|$ is called a bra. Thus

$$\langle \psi | \phi \rangle$$

is the inner product of |ψ⟩ and |φ⟩. Dirac notation is standard in quantum theory and quantum information (Nielsen and Chuang, 2010).

2.6 Linear maps and matrices

A linear map is a function between vector spaces that respects addition and scalar multiplication.

If T:V→ W, then T is linear if

$$T(av + bw) = aT(v) + bT(w)$$

for all v,w∈ V and all scalars a,b∈mathbbC.

For example, define T:mathbbC²→mathbbC² by

$$T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + y \\ 2y \end{pmatrix}.$$

Then T is linear.

Every linear map between finite-dimensional spaces can be represented by a matrix once bases are chosen. For the map above,

$$T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

So the matrix of T in the standard basis is

$$\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}.$$

A matrix is not merely a table of numbers. It represents an action on vectors. In quantum theory, such actions may represent time evolution, observables, measurement effects, noise processes, or changes of basis, depending on their special properties.

If A is an m× n matrix, it maps vectors in mathbbCⁿ to vectors in mathbbCᵐ:

$$A : \mathbb{C}^n \to \mathbb{C}^m.$$

The identity matrix I is the matrix that leaves every vector unchanged:

$$Iv = v.$$

For example,

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The zero matrix sends every vector to the zero vector.

2.7 Matrix multiplication as composition

If A:V→ W and B:W→ X are linear maps, then their composition is

$$B \circ A : V \to X,$$

defined by

$$(B \circ A)(v) = B(A(v)).$$

Matrix multiplication represents composition of linear maps.

For example, let

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 1 \\ 3 & 0 \end{pmatrix}.$$

Then

$$BA = \begin{pmatrix} 0 & 1 \\ 3 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 3 & 3 \end{pmatrix}.$$

This means: first apply A, then apply B.

Order matters. Usually,

$$AB \neq BA.$$

For example,

$$AB = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 3 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 6 & 0 \end{pmatrix},$$

which is different from BA.

This noncommutativity is not a technical annoyance. It is central in quantum theory. Different measurement operations or transformations may not be interchangeable.

2.8 Adjoints and conjugate transposes

The adjoint of a linear map is the operation that moves the map from one side of an inner product to the other.

For a matrix A, the adjoint is written A^†. It is obtained by transposing the matrix and taking complex conjugates:

$$(A^\dagger)_{ij} = \overline{A_{ji}}.$$

For example,

$$A = \begin{pmatrix} 1 & i \\ 2 & 3 - i \end{pmatrix}$$

has adjoint

$$A^\dagger = \begin{pmatrix} 1 & 2 \\ -i & 3+i \end{pmatrix}.$$

The defining property of the adjoint is

$$\langle Ax, y \rangle = \langle x, A^\dagger y \rangle.$$

Let us verify this in a simple case. Take

$$A = \begin{pmatrix} 1 & i \\ 0 & 2 \end{pmatrix}, \qquad x = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad y = \begin{pmatrix} i \\ 2 \end{pmatrix}.$$

Then

$$Ax = \begin{pmatrix} 1+i \\ 2 \end{pmatrix}.$$

So

$$\langle Ax, y \rangle = \overline{1+i}\, i + \overline{2}\, 2 = (1-i)i + 4 = 1 + i + 4 = 5 + i.$$

Also,

$$A^\dagger = \begin{pmatrix} 1 & 0 \\ -i & 2 \end{pmatrix},$$

so

$$A^\dagger y = \begin{pmatrix} i \\ -i \cdot i + 4 \end{pmatrix} = \begin{pmatrix} i \\ 5 \end{pmatrix}.$$

Thus

$$\langle x, A^\dagger y \rangle = \overline{1}\, i + \overline{1}\, 5 = 5 + i.$$

So indeed,

$$\langle Ax, y\rangle = \langle x, A^\dagger y\rangle.$$

Adjoints are everywhere in quantum information. If a matrix U describes a reversible quantum operation, then U^† describes the inverse operation. If E is a POVM element, positivity will be expressed using inner products such as

$$\langle \psi, E\psi\rangle \geq 0.$$

2.9 Self-adjoint and unitary matrices

A matrix A is self-adjoint, also called Hermitian, if

$$A = A^\dagger.$$

For example,

$$A = \begin{pmatrix} 2 & i \\ -i & 3 \end{pmatrix}$$

is self-adjoint because

$$A^\dagger = \begin{pmatrix} 2 & i \\ -i & 3 \end{pmatrix}.$$

Self-adjoint matrices are important because their eigenvalues are real and they can be diagonalized by an orthonormal basis. This is part of the finite-dimensional spectral theorem (Axler, 2015). In quantum theory, observables and projective measurements are built from self-adjoint operators and their spectral decompositions (Nielsen and Chuang, 2010).

A matrix U is unitary if

$$U^\dagger U = I$$

and, for square matrices, equivalently,

$$UU^\dagger = I.$$

A unitary matrix preserves inner products:

$$\langle Ux, Uy \rangle = \langle x, y \rangle.$$

Indeed,

$$\langle Ux, Uy \rangle = \langle x, U^\dagger U y \rangle = \langle x, Iy \rangle = \langle x, y \rangle.$$

Therefore unitary matrices preserve lengths:

$$\|Ux\| = \|x\|.$$

A standard example is the qubit Hadamard matrix

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Since H^†=H and

$$H^\dagger H = H^2 = I,$$

H is unitary.

Unitary matrices represent reversible time evolution in closed finite-dimensional quantum systems (Nielsen and Chuang, 2010). Later, when we implement a POVM by coupling a system to an ancilla, the combined system will evolve unitarily before an ordinary projective measurement is performed.

2.10 Eigenvalues and eigenvectors

An eigenvector of a linear operator A is a nonzero vector v whose direction is unchanged by A. That means

$$Av = \lambda v$$

for some scalar λ. The scalar λ is called the eigenvalue corresponding to v.

For example, let

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}.$$

Then

$$A \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

so e_1 is an eigenvector with eigenvalue 2. Similarly,

$$A \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \end{pmatrix} = 3 \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

so e_2 is an eigenvector with eigenvalue 3.

Eigenvectors are useful because they show directions in which a linear operator acts simply. Instead of mixing the vector into a complicated new direction, the operator only stretches, shrinks, rotates by a phase, or changes sign.

For self-adjoint matrices, eigenvalues are real. For unitary matrices, eigenvalues have complex absolute value 1. These facts follow from the definitions.

If A=A^† and Av=λ v with v≠ 0, then

$$\lambda\langle v, v\rangle = \langle v, \lambda v\rangle = \langle v, Av\rangle.$$

But since A is self-adjoint,

$$\langle v, Av\rangle = \langle Av, v\rangle = \langle \lambda v, v\rangle = \overline{\lambda}\langle v, v\rangle.$$

Because ⟨ v,v⟩>0, we get

$$\lambda = \overline{\lambda},$$

so λ is real.

If U is unitary and Uv=λ v, then

$$\|v\| = \|Uv\| = \|\lambda v\| = |\lambda|\|v\|.$$

Since v≠ 0, we get

$$|\lambda| = 1.$$

These simple arguments already show why self-adjoint and unitary operators play different roles. Self-adjoint operators have real spectral values, suitable for measurement labels. Unitary operators preserve total probability amplitude, suitable for reversible evolution.

2.11 The spectral theorem in finite dimensions

The spectral theorem is one of the central bridges between linear algebra and quantum theory. In finite dimensions, it says that every self-adjoint matrix can be diagonalized using an orthonormal basis of eigenvectors (Axler, 2015).

Concretely, if A=A^† on an n-dimensional complex inner product space, then there is an orthonormal basis

$$u_1, \ldots, u_n$$

and real numbers

$$\lambda_1, \ldots, \lambda_n$$

such that

$$Au_j = \lambda_j u_j$$

for each j.

In matrix language, this means

$$A = UDU^{\dagger},$$

where U is unitary and D is diagonal with real entries.

There is another form that will be especially important for measurement theory. Define

$$P_j = |u_j\rangle\langle u_j|.$$

This is the rank-one projection onto the line spanned by u_j. Then

$$A = \sum_{j=1}^{n} \lambda_j P_j.$$

This is called a spectral decomposition.

Let us see an example. Take

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 5 \end{pmatrix}.$$

The eigenvectors are e_1,e_2, with eigenvalues 2,5. The projections are

$$P_1 = |e_1\rangle\langle e_1| = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

and

$$P_2 = |e_2\rangle\langle e_2| = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Then

$$A = 2P_1 + 5P_2.$$

In Chapter 5, this decomposition will become the mathematical basis of projective measurement. The projectors P_j represent the possible measurement alternatives, and the real numbers λ_j can be interpreted as the numerical values assigned to those alternatives.

2.12 Projections

A projection is an operator that leaves some part of a vector unchanged and removes another part.

Algebraically, an operator P is a projection if

$$P^2 = P.$$

This equation means that applying P twice is the same as applying it once. Once a vector has already been projected, projecting again does nothing new.

In quantum theory, the most important projections are orthogonal projections. An operator P is an orthogonal projection if

$$P^2 = P \quad \text{and} \quad P^\dagger = P.$$

The first condition says P is a projection. The second says it is self-adjoint. Orthogonal projections correspond exactly to projections onto subspaces along their orthogonal complements in finite-dimensional inner product spaces (Axler, 2015).

For example,

$$P = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

is an orthogonal projection. It sends

$$\begin{pmatrix} x \\ y \end{pmatrix}$$

to

$$\begin{pmatrix} x \\ 0 \end{pmatrix}.$$

It keeps the e_1-component and removes the e_2-component.

Check:

$$P^2 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = P,$$

and

$$P^\dagger = P.$$

If u is a unit vector, then

$$P = |u\rangle\langle u|$$

is the orthogonal projection onto the one-dimensional subspace spanned by u. For any vector v,

$$Pv = |u\rangle\langle u|v\rangle = \langle u, v\rangle u.$$

This formula says: take the component of v in the direction u, then return that multiple of u.

For example, let

$$u = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Then

$$P = |u\rangle\langle u| = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

This projection keeps the part of a vector lying along the diagonal line spanned by $(1,1)^{T}$.

Orthogonal projections will become the building blocks of projective measurements. A projective measurement with outcomes $1,\ldots,k$ is represented by projections $P_1,\ldots,P_k$ satisfying

$$P_i P_j = 0 \quad \text{for } i \neq j,$$

and

$$P_1 + \cdots + P_k = I.$$

This means the projections describe mutually exclusive alternatives that together cover the whole space.

2.13 Positive semidefinite operators

A self-adjoint operator A is called positive semidefinite, or simply positive, if

$$\langle v, Av \rangle \geq 0$$

for every vector v.

We write

$$A \geq 0$$

to mean that A is positive semidefinite.

The expression

$$\langle v, Av \rangle$$

is a complex number in general if A is arbitrary. But if A is self-adjoint, then $\langle v,Av \rangle$ is always real. Positivity says it is never negative.

For example,

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$$

is positive because

$$\left\langle \begin{pmatrix} x \\ y \end{pmatrix}, A \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle = 2|x|^2 + 3|y|^2 \geq 0.$$

The matrix

$$B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

is self-adjoint but not positive, because

$$\left\langle \begin{pmatrix} 0 \\ 1 \end{pmatrix}, B \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\rangle = -1.$$

Positive operators are central to generalized measurements. A POVM element is not necessarily a projection, but it must be positive. This guarantees that the Born rule produces nonnegative probabilities.

There is a useful spectral characterization: a self-adjoint matrix is positive semidefinite if and only if all its eigenvalues are nonnegative (Axler, 2015; Watrous, 2018).

Why? Suppose

$$A = \sum_{j=1}^{n} \lambda_j |u_j\rangle\langle u_j|$$

is the spectral decomposition of A, where $u_1,\ldots,u_n$ is an orthonormal basis. If

$$v = \sum_{j=1}^{n} c_j u_j,$$

then

$$\langle v, Av \rangle = \sum_{j=1}^{n} \lambda_j |c_j|^2.$$

If every $\lambda_j \geq 0$, this is always nonnegative. Conversely, if $A \geq 0$, then choosing $v = u_j$ gives

$$\langle u_j, Au_j \rangle = \lambda_j \geq 0.$$

So positivity is exactly the condition that the spectrum is nonnegative.

2.14 Positive square roots

Every positive semidefinite matrix A has a unique positive semidefinite square root. That is, there exists a unique positive semidefinite matrix B such that

$$B^2 = A.$$

We write

$$B = \sqrt{A}$$

or

$$B = A^{1/2}.$$

This fact follows from the spectral theorem

Document information

Chapter 2: Linear Algebra for Quantum Theory

Project	Naimark Dilation from First Principles
Document	Document 1.6
Author	mujirin
Verifier	Not verified
Downloaded	July 04, 2026 22:22 KST
Status	Working
Document link	https://www.theorytrace.com/projects/naimark-dilation-from-first-principles/documents/-chapter-2-linear-algebra-for-quantum-theory/