Introduction to linear algebra for quantum information

summary of ***Quantum Computation and Quantum Information****,*
by Nielsen & Chuang, §2.1,
provided by Xavier Bonet-Monroig (2019)

1 Vector spaces

1.1 Vectors

Linear algebra is the branch of mathematics that studies vector spaces and linear operations on them. A vector is an ordered tuple of numbers,

$$\vec{v} = (a_1, a_2, \ldots, a_n).$$

If the elements of a vector are real, $a_i \in \mathbb{R}$, then we say that the elements of the vector space are real vectors. Similarly, a complex vector is formed by elements $a_i \in \mathbb{C}$ that are complex numbers. In quantum mechanics and quantum information we are interested in complex vector spaces. The vector space that contains all n-tuple complex vector is $\mathbb{C}^n$. An element of such a vector space is given by

$$\vec{v} = (v_1, v_2, \ldots, v_n),\ v_i \in \mathbb{C}.$$

A set of vectors forms a vector space if it fulfills

1. **Addition**: if $\vec{v},\ \vec{w} \in \mathbb{C}^n$,
$$\vec{v} + \vec{w} = \vec{z}\ \in \mathbb{C}^n.$$
2. **Multiplication by scalar**: if $\vec{v} \in \mathbb{C}^n$ and $a \in \mathbb{C}$,
$$a\vec{v} = \begin{bmatrix} av_1 \\ \vdots \\ av_n \end{bmatrix}.$$
3. **Zero vector**: There exists a vector $\vec{0} = (0, \ldots, 0)$ such that
$$\vec{v} + \vec{0} = \vec{v}.$$

If $\vec{v}$ is a column vector then the transpose $\vec{v}^\top$ is a row vector. The complex conjugate of a vector is denoted by a $*$ (mathematicians prefer the overline, they would write $\overline{\vec{v}}$ instead of $\vec{v}^*$). The conjugate transpose of the vector $\vec{v}$ is denoted by a superscript dagger,

$$\vec{v}^\dagger = (\vec{v}^*)^\top = (\vec{v}^\top)^*.$$

1.2 Dirac notation

In the development of quantum mechanics, physicists invented a rather particular notation for the elements of vector spaces. Such a notation goes by the name of Dirac *bra-ket* notation, after the inventor Paul Dirac. Bra's and ket's are different names for row and column vectors:

- A column vector $\vec{v} \in \mathbb{C}^n$ is a *ket*

$$|v\rangle = \vec{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

- A row vector $\vec{w}^\dagger \in \mathbb{C}^n$ is a *bra*

$$\langle w| = \vec{w}^\dagger = \left[w_1^*, \ldots, w_n^*\right].$$

Notice that the move from bra to ket also includes the complex conjugation. The arrow is no longer needed so it is omitted. The zero vector is typically still denoted by $\vec{0}$, because the notations $|0\rangle$ and $|1\rangle$ are used to denote the states 0 and 1 of a qubit. So we would write $0|1\rangle = \vec{0}$ instead of $0|1\rangle = |0\rangle$.

2 Bases and linear independence

Given a vector space $\mathcal{V}$, we say that the set of vectors $\{|v_1\rangle, \ldots, |v_n\rangle\} \in \mathcal{V}$ is a *spanning* set if for any $|\tilde{v}\rangle \in \mathcal{V}$ there exists a linear combination of the spanning vectors that is equal to $|\tilde{v}\rangle$,

$$|\tilde{v}\rangle = \sum_i a_i |v_i\rangle,$$

$$\langle\tilde{v}| = \sum_i a_i^* \langle v_i|.$$

Notice how the coefficients for the bra are the complex conjugates of the coefficients for the ket.

A set of non-zero vectors $\{|v_1\rangle, \ldots, |v_n\rangle\} \in \mathcal{V}$ are said to be linearly dependent if there exists a set of complex numbers $\{a_1, \ldots, a_n\}$, not all equal to zero, such that

$$a_1|v_1\rangle + a_2|v_2\rangle + \ldots + a_n|v_n\rangle = \vec{0}.$$

On the contrary, if no such combination exists, we say that the set of vectors are linearly independent.

We define a basis as the set of vectors $\{|b_1\rangle, \ldots, |b_n\rangle\} \in \mathcal{V}$ that is spanning and linearly independent. If $\mathcal{V}$ is a vector space it is guaranteed that a basis exists. The number of elements of the set that defines a basis for the vector space is called dimension of the space.

All of this assumes a finite-dimensional vector space. In quantum information we only need to consider that case, which is a major simplification compared to the types of problems one encounters in atomic physics.

3 Hilbert spaces

3.1 Inner product

The term Hilbert space denotes a vector space with an inner product. The inner product of two elements of a vector space $|v\rangle, |w\rangle \in \mathcal{V}$ is a function from the space to a scalar, $\mathcal{V} \to \mathbb{C}$. The inner product in Dirac notation is defined as

$$\langle w|v\rangle = a, \ a \in \mathbb{C}. \tag{1}$$

The inner product of a vector space satisfies the following properties:

1. **Linear in the *ket*:** $\langle v|w\rangle = \langle v| \sum_i \lambda_i w_i\rangle = \sum_i \lambda_i \langle v|w_i\rangle$.
2. **Conjugate symmetric:** $\langle v|w\rangle = \langle w|v\rangle^*$.
3. **Positive definite:** $\langle v|v\rangle > 0$ if $|v\rangle \neq \vec{0}$, $\langle v|v\rangle = 0$ if $|v\rangle = \vec{0}$.

It follows that the inner product is conjugate linear in the *bra*:$\langle \sum_i \lambda_i v_i | w \rangle = \sum_i \lambda_i^* \langle v_i | w \rangle$.

Explicitly, in the vector space $\mathbb{C}^n$ the inner product is the element-wise product of the bra and the ket,

$$\langle v|w\rangle = \begin{bmatrix} v_1^*, v_2^*, \ldots, v_n^* \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \sum_{i=1}^{n} v_i^* w_i. \tag{2}$$

Infinite-dimensional Hilbert spaces need an extra convergence condition, which we do not need in the finite-dimensional case.

3.2 Orthonormal bases

Two vectors $|v\rangle, |w\rangle \in \mathcal{V}$ are orthogonal if

$$\langle v|w\rangle = \langle w|v\rangle = 0. \tag{3}$$

The norm of a vector $|v\rangle$ in a vector space is defined as

$$\| |v\rangle \| = \sqrt{\langle v|v\rangle}. \tag{4}$$

A vector with norm equal to 1 is called a unit vector. Any nonzero vector of a vector space can be normalized to unity,

$$|v_{\text{norm}}\rangle = \frac{|v\rangle}{\sqrt{\langle v|v\rangle}}. \tag{5}$$

A quantum state is represented by a normalized vector. Two normalized vectors $|v\rangle$ and $e^{i\phi}|v\rangle$ that differ by a phase $\phi \in \mathbb{R}$ represent the same quantum state. The states of the form $e^{i\phi}|v\rangle$ with real ϕ are said to form a *ray* in the Hilbert space.

It is common to label the set of vectors that form an orthonormal basis as $\{|i\rangle\}$, such that for all i, j one has

$$\langle i|j\rangle = \delta_{ij}. \tag{6}$$

4 Linear operators

4.1 Matrix representation

A linear operator A between two vector spaces $\mathcal{V}$ and $\mathcal{W}$ is a function from $\mathcal{V} \to \mathcal{W}$ that is linear on its inputs,

$$A\Big(\sum_i a_i |v_i\rangle\Big) = \sum_i a_i A(|v_i\rangle). \tag{7}$$

Example: The identity operator I and the zero operator $[0]$ are linear operators from $\mathcal{V} \to \mathcal{V}$. Given $|v\rangle \in \mathcal{V}$, the action of I and $[0]$ on it is given by

$$I|v\rangle = |v\rangle$$
$$[0]|v\rangle = \vec{0}.$$

Let's suppose that A is a linear operator between vector spaces $A: \mathcal{V} \to \mathcal{W}$, and

$$\{|v_1\rangle, \ldots, |v_n\rangle\} \in \mathcal{V},$$
$$\{|w_1\rangle, \ldots, |w_n\rangle\} \in \mathcal{W},$$

are bases of such vector spaces. For each $j \in 1, \ldots, m$ there exist a set of scalars $A_{1j}, \ldots, A_{nj}$ such that

$$A|v_j\rangle = \sum_{i=1}^{n} A_{ij} |w_i\rangle. \tag{8}$$

The scalars A_{ij} form an $n \times m$ matrix representation of the linear operator A in the bases of $\mathcal{V}$ and $\mathcal{W}$.

The matrix element A_{ij} can be calculated by taking the inner product of $|w_i\rangle$ with $|v_j\rangle$,

$$A_{ij} = \langle w_i|A|v_j\rangle. \tag{9}$$

Equivalently, $\{|j\rangle\} \in \mathcal{V}$ and $\{|i\rangle\} \in \mathcal{W}$ are orthonormal bases of $\mathcal{V}$ and $\mathcal{W}$ then

$$A_{ij} = \langle i|A|j\rangle. \tag{10}$$

Notice how compactly the bra-ket notation expresses this matrix element. Keep in mind that the matrix representation of a linear operator is basis-dependent. Many properties of operators are basis-independent, so one should be able to derive them without committing to particular basis.

The Pauli operators I, X, Y, Z play a central role in quantum information. Their vector space is $\mathbb{C}^2$. In the basis $|0\rangle = \binom{1}{0}$ and $|1\rangle = \binom{0}{1}$ they have the matrix representation

$$I = \sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad X = \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y = \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad Z = \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

4.2 Outer product

We have defined the inner product between two vectors of a vector space as a map from vectors to scalars. In a similar fashion, vectors can be multiplied to form linear operators that act on the vector space. Such a construction is known as the *outer* product.

Given $|v\rangle \in \mathcal{V}$ and $|w\rangle \in \mathcal{W}$ we can define a linear map from $\mathcal{V}$ to $\mathcal{W}$ as

$$|w\rangle\langle v| : \mathcal{V} \to \mathcal{W}.$$

The action of this linear operator on a vector $|v'\rangle \in \mathcal{V}$ is given by

$$(|w\rangle\langle v|)(|v'\rangle) \equiv |w\rangle(\langle v|v'\rangle) = \langle v|v'\rangle|w\rangle. \tag{11}$$

Remember that $\langle v|v'\rangle$ is a scalar, hence this equation has two interpretations:

1. The action of a linear operator on a vector $|v'\rangle$,
2. The product of a scalar with the vector $|w\rangle$.

4.3 Completeness relation

An important result from the outer product is the so-called *completeness relation*. Let $\{|i\rangle\}$ be any orthonormal basis of a vector space and $|v\rangle$ any vector of such space, $|v\rangle = \sum_i v_i|i\rangle$. The action of the application of the linear operator formed by the basis vectors on the vector $|v\rangle$ is

$$\Big(\sum_j |j\rangle\langle j|\Big)|v\rangle = \sum_j |j\rangle\langle j| \sum_i v_i|i\rangle = \sum_{ij} v_i|j\rangle\langle j|i\rangle = \sum_{ij} v_i|j\rangle\delta_{ij} = \sum_i v_i|i\rangle = |v\rangle, \tag{12}$$

hence this returns the vector itself and therefore this operator is the identity operator,

$$\sum_i |i\rangle\langle i| = I. \tag{13}$$

This formula is known as the completeness relation, or also as the "resolution of the identity".

4.4 Hermitian and unitary operators

If the linear operator A has matrix representation A_{ij} then the conjugate transpose A^*_{ji} represents the adjoint operator $A^\dagger$, also called the Hermitian conjugate of A. (Warning: mathematicians denote the adjoint by $*$, so a mathematician would write $A^*_{ij} = \overline{A}_{ji}$, while a physicist writes $A^\dagger_{ij} = A^*_{ji}$.) A operator A such that $A = A^\dagger$ is called self-adjoint or Hermitian.

Check that, if $|u\rangle = A^\dagger|w\rangle$, then

$$\langle w|A|v\rangle = \langle u|v\rangle.$$

Also check that

$$(|w\rangle\langle v|)^\dagger = |v\rangle\langle w|.$$

If the operator A is invertible, it means that there exists an operator A^{-1} such that

$$AA^{-1} = A^{-1}A = I.$$

If $A^{-1} = A^\dagger$ then A is called a unitary operator. Check that a unitary operator preserves the inner product: If $|w'\rangle = U|w\rangle$ and $|v'\rangle = U|v\rangle$ and U is unitary, then

$$\langle w|v\rangle = \langle w'|v'\rangle.$$

5 Tensor product vector space

It is possible to construct larger vector spaces from smaller ones using the tensor product.

Let $\mathcal{V}$ and $\mathcal{W}$ be two vector spaces of dimension m and n respectively. The tensor product vector space $\mathcal{T} = \mathcal{V} \otimes \mathcal{W}$ is a new vector space with dimension mn. The elements of the tensor vector space are of the form

$$|t\rangle = |v\rangle \otimes |w\rangle, \tag{14}$$

for $|v\rangle \in \mathcal{V}$ and $|w\rangle \in \mathcal{W}$. Moreover, if $\{|i\rangle\} \in \mathcal{V}$ and $\{|j\rangle\} \in \mathcal{W}$ are orthonormal bases of their respective spaces, then

$$\{|k\rangle = |i\rangle \otimes |j\rangle\} \tag{15}$$

is also an orthonormal basis of $\mathcal{T}$.

By construction, the tensor product vector space fulfills the properties of a vector space:

- Multiplication by scalar:

$$z|t\rangle = z(|v\rangle \otimes |w\rangle) = (z|v\rangle) \otimes |w\rangle = |v\rangle \otimes (z|w\rangle).$$

- For $|v_1\rangle, |v_2\rangle \in \mathcal{V}$ and $|w\rangle \in \mathcal{W}$,

$$(|v_1\rangle + |v_2\rangle) \otimes |w\rangle = |v_1\rangle \otimes |w\rangle + |v_2\rangle \otimes |w\rangle.$$

- For $|w_1\rangle, |w_2\rangle \in \mathcal{W}$ and $|v\rangle \in \mathcal{V}$,

$$|v\rangle \otimes (|w_1\rangle + |w_2\rangle) = |v\rangle \otimes |w_1\rangle + |v\rangle \otimes |w_2\rangle.$$

The inner product of the tensor product vector space can be obtained from the inner product of its components. Let $|t\rangle, |t'\rangle \in \mathcal{T}$ be two vectors of the tensor product vector space such that:

$$|t\rangle = \sum_i a_i |v_i\rangle \otimes |w_i\rangle$$
$$|t'\rangle = \sum_j b_j |v'_j\rangle \otimes |w'_j\rangle.$$

The inner product

$$\begin{aligned}\langle t|t'\rangle = \sum_i a_i^*(\langle v_i| \otimes \langle w_i|) \sum_j b_j(|v'_j\rangle \otimes |w'_j\rangle) = \\ \sum_{ij} a_i^* b_j(\langle v_i|v'_j\rangle \otimes \langle w_i|w'_j\rangle) = \sum_{ij} a_i^* b_j \langle v_i|v'_j\rangle\langle w_i|w'_j\rangle,\end{aligned} \tag{16}$$

maps vectors on $\mathcal{T}$ to scalars.

One can also define linear operators that act on the tensor product vector space. Let A and B be two linear operators acting on $\mathcal{V}$ and $\mathcal{W}$ vector spaces respectively. A linear operator C of the tensor product vector space $\mathcal{T} = \mathcal{V} \otimes \mathcal{W}$ is defined as

$$C = A \otimes B, \tag{17}$$

and its action on a vector $|t\rangle \in \mathcal{T}$ is given by

$$C|t\rangle = (A \otimes B)(|v\rangle \otimes |w\rangle) = A|v\rangle \otimes B|w\rangle. \tag{18}$$

6 Eigenvectors and eigenvalues

An eigenvector of a linear operator A on a vector space is a non-zero vector $|v\rangle$ such that

$$A|v\rangle = v|v\rangle,$$

where v is a complex number known as the eigenvalue of A corresponding to the eigenvector $|v\rangle$.

The diagonal representation (or orthonormal decomposition) of an operator A on a vector space with orthonormal basis $\{|i\rangle\}$ is the representation

$$A = \sum_i \lambda_i |i\rangle\langle i|.$$

An operator is said to be diagonalizable if it has a diagonal representation. An eigenvalue is said to be degenerate if there are two or more independent eigenvectors with the same eigenvalue.

Hermitian and unitary operators are both diagonalizable. The eigenvalues of a Hermitian operator are real numbers, the eigenvalues of a unitary operator are complex numbers of unit absolute value. Their orthonormal decomposition can be written as

$$A = UDU^\dagger,$$

with U a unitary matrix and D a diagonal matrix with the eigenvalues on the diagonal. The columns of U are the eigenvectors of A.

Some operators are not diagonalizable, but all linear operators from $\mathcal{V}$ to $\mathcal{V}$, corresponding to square matrices, have a socalled singular value decomposition

$$A = UDV^\dagger,$$

with U, V unitary operators and D a diagonal matrix with $D_{ii} \geq 0$. The diagonal elements $d_1, d_2, \ldots d_n$ of D are called the singular values of A. The columns of U are called the left eigenvectors, the columns of V are the right eigenvectors. Note that the squared singular values d_i^2 are the eigenvalues of the Hermitian operator $AA^\dagger$.

The trace $\mathrm{Tr}\, A$ of an operator is defined as the sum of the diagonal elements of its matrix representation,

$$\mathrm{Tr}\, A = \sum_i A_{ii}.$$

The trace of a diagonalizable operator is the sum of the eigenvalues. The determinant is the product of the eigenvalues.