Lecture 2: Linear operators

Rajat Mittal

IIT Kanpur

The mathematical formulation of Quantum computing requires vector spaces and linear operators. So, we need to be comfortable with linear algebra to study quantum computing. These notes will assume that the reader is familiar with the concept of vector space, basis, linear independence and tensor product. Strang's book, Linear Algebra and its applications, is a good source to brush up these concepts.

This exposition will focus on vectors, matrices and properties of matrices. *Dirac's notation* is used widely in quantum computing, because it simplifies the understanding of quantum mechanical concepts. We will switch between the standard vector notation and Dirac notation in these notes.

Exercise 1. Read about vector space, basis, linear independence and tensor product if you are not comfortable with these words.

One of the most fundamental concept in linear algebra is that of a *vector space*. We will mostly concern ourselves with the vector space $\mathbb{C}^n$, the vector space of dimension n over the field of complex numbers. This means that the scalars used in these vector spaces are complex numbers.

A vector is the most basic unit of a vector space. Using Dirac's notation, a column vector will be denoted by $|\psi\rangle$. Suppose $|v_1\rangle, |v_2\rangle, \cdots, |v_n\rangle$ is the basis of the vector space, then any vector $|v\rangle$ can be written as,

$$|v\rangle = a_1|v_1\rangle + \cdots + a_n|v_n\rangle.$$

For a vector space with dimension n, the standard basis is denoted by $|0\rangle, |1\rangle, \cdots, |n-1\rangle$. Here you can think of $|i\rangle$ as the vector with 1 at the $(i-1)^{th}$ position and 0 otherwise. For example, a qubit lives in a 2-dimensional space with basis $|0\rangle, |1\rangle$.

Note 1. There is a difference between vector $|0\rangle$ and vector 0, the vector with all entries 0. First one is a basis vector with the first entry 1 and rest 0.

The notation $\langle\psi|$ denotes the row vector whose entries are complex conjugate of the entries of the vector $|\psi\rangle$. The space $\mathbb{C}^n$ is equipped with the natural inner product,

$$((x_1, x_2, \cdots, x_n), (y_1, y_2, \cdots, y_n)) = \sum_{i=1}^{n} x_i^* y_i.$$

The inner product (dot product) between two vectors $|\psi\rangle, |\phi\rangle$ is denoted by $\langle\psi|\phi\rangle$. The vector for the tensor product space, $|\psi\rangle \otimes |\phi\rangle$ will be simply written as $|\psi\rangle|\phi\rangle$. In Dirac's notation, the expression $A = |\psi\rangle\langle\phi|$ is a matrix which takes $|v\rangle$ to $\langle\phi|v\rangle|\psi\rangle$. The analog of this expression in the simple vector notation would be, $A = \psi(\phi)^T$.

1 Operators

Given two vector spaces, V and W over $\mathbb{C}$, a *linear* operator $M : V \to W$ is defined as an operator satisfying the following properties.

- $M(x+y) = M(x) + M(y)$.
- $M(\alpha x) = \alpha M(x),\ \forall \alpha \in \mathbb{C}$.

These conditions imply that the *zero* of the vector space V is mapped to the *zero* of the vector space W. Also,

$$M(\alpha_1 x_1 + \cdots + \alpha_k x_k) = \alpha_1 M(x_1) + \cdots + \alpha_k M(x_k)$$

Where $x_1, \cdots, x_k$ are elements of V and α_i's are in $\mathbb{C}$. Because of this linearity, it is enough to specify the value of a linear operator on any basis of the vector space V. In other words, a linear operator is uniquely defined by the values it takes on any particular basis of V.

Let us define the addition of two linear operators as $(M+N)(u) = M(u) + N(u)$. Similarly, αM (scalar multiplication) is defined to be the operator $(\alpha M)(u) = \alpha M(u)$. The space of all linear operators from V to W (denoted $L(V, W)$) is a vector space in itself. The space of linear operators from V to V will be denoted by $L(V)$.

Exercise 2. Given the dimension of V and W, what is the dimension of the vector spaces $L(V, W)$?

1.1 Matrices as linear operators

Given two vector spaces $V = \mathbb{C}^n, W = \mathbb{C}^m$ and a matrix M of dimension $m \times n$, the operation $x \in V \to Mx \in W$ is a linear operation. So, a matrix acts as a linear operator on the corresponding vector space.

To ask the converse, can any linear operator be specified by a matrix?

Let f be a linear operator from a vector space V (dimension n) to a vector space W (dimension m). Suppose $\{e_1, e_2, \cdots, e_n\}$ is a basis for the vector space V. Denote the images of this basis under f as $\{w_1 = f(e_1), w_2 = f(e_2) \cdots, w_n = f(e_n)\}$.

Exercise 3. What is the lower-bound/ upper-bound on the dimension of the vector space spanned by $\{w_1, w_2, \cdots, w_n\}$?

Define M_f to be the matrix with columns $w_1, w_2, \cdots, w_n$. Notice that M_f is a matrix of dimension $m \times n$. It is a simple exercise to verify that the action of the matrix M_f on a vector $v \in V$ is just $M_f v$. Here we assume that v is expressed in the chosen basis $\{e_1, e_2, \cdots, e_n\}$.

Exercise 4. Convince yourself that Mv is a linear combination of columns of M.

The easiest way to see the above fact is: notice that the matrix M_f and the operator f act exactly the same on the basis elements of V. Since both the operations are linear, they are exactly the same operation. This proves that any linear operation can be specified by a matrix.

The previous discussion does not depend upon the chosen basis. We can pick our favorite basis, and the linear operator can similarly be written in the new basis as a matrix (The columns of this matrix are images of the basis elements). In other words, given bases of V and W and a linear operator f, it has a unique matrix representation.

To compute the action of a linear operator, express $v \in V$ in the preferred basis and multiply it with the matrix representation. The output will be in the chosen basis of W. We will use the two terms, linear operator and matrix, interchangeably in future (the bases will be clear from the context).

For a matrix A, A^T denotes the transpose of the matrix and A^* denotes the adjoint of the matrix (take complex conjugate and then transpose).

Let us look at some simple matrices which will be used later.

- Zero matrix: The matrix with all the entries 0. It acts trivially on every element and takes them to the 0 vector.
- Identity matrix: The matrix with 1's on the diagonal and 0 otherwise. It takes $v \in V$ to v itself.
- All 1's matrix (J): All the entries of this matrix are 1.

Exercise 5. What is the action of matrix J?

1.2 Kernel, image and rank

For a linear operator/matrix (from V to W), the *kernel* is defined to be the set of vectors which map to 0.

$$ker(M) = \{x \in V : Mx = 0\}$$

Here 0 is a vector in space W.

Exercise 6. What is the kernel of the matrix J?

The *image* is the set of vectors which can be obtained through the action of the matrix on some element of the vector space V.

$$img(M) = \{x \in W : \exists y \in V,\ x = My\}$$

Exercise 7. Show that $img(M)$ and $ker(M)$ are subspaces.

Exercise 8. What is the image of J?

Notice that $ker(M)$ is a subset of V, but $img(M)$ is a subset of W. The dimension of $img(M)$ is known as the *rank* of M ($rank(M)$). The dimension of $ker(M)$ is known as the nullity of M ($nullity(M)$). For a matrix $M \in L(V, W)$, by the famous rank-nullity theorem,

$$rank(M) + nullity(M) = dim(V).$$

Here $dim(V)$ is the dimension of the vector space V.

Proof. Suppose $u_1, \cdots, u_k$ is the basis for $ker(M)$. We can extend it to the basis of V, $u_1, \cdots, u_k, v_{k+1}, \cdots, v_n$. We need to prove that the dimension of $img(M)$ is $n - k$. It can be proved by showing that the set $\{Mv_{k+1}, \cdots, Mv_n\}$ forms a basis of $img(M)$.

Exercise 9. Prove that any vector in the image of M can be expressed as linear combination of $Mv_{k+1}, \cdots, Mv_n$. Also any linear combination of $Mv_{k+1}, \cdots, Mv_n$ can't be zero vector.

□

Given a vector v and a matrix M, it is easy to see that the vector Mv is a linear combination of columns of M. To be more precise, $Mv = \sum_i M_i v_i$ where M_i is the ith column of M and v_i is the ith co-ordinate of v. This implies that any element in the image of M is a linear combination of its columns.

Exercise 10. Prove the rank of a matrix is equal to the dimension of the vector space spanned by its columns (column-space).

The dimension of the column space is sometimes referred as the *column-rank*. We can similarly define the *row-rank*, the dimension of the space spanned by the rows of the matrix. Luckily, row-rank turns out to be equal to column-rank and we will call both of them as the rank of the matrix. This can be proved easily using *Gaussian elimination*. We will give a *visual* proof of the theorem.

Proof. Given an $m \times n$ matrix M, say $\{c_1, c_2, \cdots, c_k\}$ span the column space of M. Suppose, C be the $m \times k$ matrix with columns $\{c_1, c_2, \cdots, c_k\}$. Then, there exist an $k \times n$ matrix R, s.t., $CR = M$. If $\{d_1, d_2, \cdots, d_k\}$ are the columns of R, then the equation $CR = M$ can be viewed as,

$$\begin{pmatrix} \vdots & \vdots & \vdots & \vdots \\ c_1 & c_2 & \cdots & c_k \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} \begin{pmatrix} \vdots & \vdots & \vdots & \vdots \\ d_1 & d_2 & \cdots & d_k \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} = \begin{pmatrix} \vdots & \vdots & \vdots & \vdots \\ Cd_1 & Cd_2 & \cdots & Cd_k \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

Another way to view the same equation is,

$$C \begin{pmatrix} \cdots & r_1 & \cdots \\ \cdots & r_2 & \cdots \\ \cdots & \cdot & \cdots \\ \cdots & \cdot & \cdots \\ \cdots & r_k & \cdots \end{pmatrix} = \begin{pmatrix} \cdots & \sum_i C_{1i} r_i & \cdots \\ \cdots & \sum_i C_{2i} r_i & \cdots \\ \cdots & \cdot & \cdots \\ \cdots & \cdot & \cdots \\ \cdots & \sum_i C_{ki} r_i & \cdots \end{pmatrix}$$

This shows that the k columns of R span the row-space of M. Hence, column-rank is smaller than the row-rank.

Exercise 11. Show that row-rank is less than column-rank by a similar argument.

□

Note 2. The column-rank is equal to row-rank. It does not mean that the row-space is same as the column-space.

Using these characterizations of rank, it can be proved easily that $rank(A) = rank(A^*)$ and $rank(A + B) \leq rank(A) + rank(B)$.

Exercise 12. Show that $|\psi\rangle\langle\phi|$ is a linear operator with rank 1.

1.3 Operations on matrices

Lets look at some of the basic operations on these matrices.

- Trace: The *trace* of a matrix is the sum of all the diagonal elements.

$$tr(A) = \sum_i A[i, i]$$

- Entry-wise multiplication: The entry-wise multiplication of two matrices is known as *Hadamard product* and only makes sense when both of them have same number of rows and columns. The Hadamard product of two matrices A, B is

$$(A \circ B)[i, j] = A[i, j]B[i, j].$$

  The related operation is when you add up the entries of this Hadamard product.

$$(A \bullet B) = \sum_{i,j} A[i, j]B[i, j]$$

  Notice that $A \bullet B$ is a scalar and not a matrix.

  Exercise 13. Given a matrix, express $\bullet$ operation in terms of multiplication and trace operation.

- Inverse: Inverse of a matrix M is the matrix M^{-1}, s.t., $MM^{-1} = M^{-1}M = I$. The inverse only exists if the matrix has full rank (the columns of M span the whole space).

  Exercise 14. What is the inverse of matrix J (all 1's matrix).

  Exercise 15. Suppose M, N are two square matrices, show that $MN = I \Rightarrow NM = I$.

  Exercise 16. Show that the inverse of a matrix exists iff it has full rank.

2 Eigenvalues and eigenvectors

A matrix $M \in L(V, W)$ is square if $dim(V) = dim(W)$. In particular, a matrix $M \in L(V)$ is always square. Consider a matrix $M \in L(V)$, any vector $v \in V$ satisfying,

$$Mv = \lambda v \text{ for some } \lambda \in \mathbb{C},$$

is called the *eigenvector* of matrix M with *eigenvalue* λ.

Exercise 17. Given two eigenvectors v, w, when is their linear combination an eigenvector itself?

The previous exercise can be used to show that all the eigenvectors corresponding to a particular eigenvalue form a subspace. This subspace is called the *eigenspace* of the corresponding eigenvalue.

An eigenvalue λ of an $n \times n$ matrix M satisfies the equation

$$Det(\lambda I - M) = 0,$$

where $Det(M)$ denotes the determinant of the matrix M. The polynomial $Det(\lambda I - M) = 0$, in λ, is called the *characteristic polynomial* of M. The characteristic polynomial has degree n and will have n roots in the field of complex numbers. Though, these roots might not be real.

Exercise 18. Give an example of a matrix with no real eigenvalue.

Theorem 1. *Given a matrix P of full rank, matrix M and matrix $P^{-1}MP$ have the same set of eigenvalues.*

Proof. Suppose λ is an eigenvalue of $P^{-1}MP$, we need to show that it is an eigenvalue for M too. Say λ is an eigenvalue with eigenvector v. Then,

$$P^{-1}MPv = \lambda v \Rightarrow M(Pv) = \lambda Pv.$$

Hence Pv is an eigenvector with eigenvalue λ.

The opposite direction follows similarly. Given an eigenvector v of M, it can be shown that $P^{-1}v$ is an eigenvector of $P^{-1}MP$.

$$P^{-1}MP(P^{-1}v) = P^{-1}Mv = \lambda P^{-1}v$$

Hence proved. □

Exercise 19. Where did we use the fact that P is a full rank matrix?

3 Spectral decomposition

Exercise 20. Let v_1, v_2 be two eigenvectors of a matrix M with distinct eigenvalues. Show that these two eigenvectors are linearly independent.

Given an $n \times n$ matrix M, it need not have n linearly independent eigenvectors. The matrix M is called *diagonalizable* iff the set of eigenvectors of M span the complete space $\mathbb{C}^n$.

For a diagonalizable matrix, the basis of eigenvectors need not be an orthogonal basis. We will be interested in matrices which have an orthonormal basis of eigenvectors.

A *normal* matrix is defined to be a matrix M, s.t., $MM^* = M^*M$. Spectral theorem shows that we can form an orthonormal basis of $\mathbb{C}^n$ using the eigenvectors of a normal matrix.

Theorem 2. *Spectral theorem: For a normal matrix $M \in L(V)$, there exists an orthonormal basis $|x_1\rangle, \cdots, |x_k\rangle$ of V, s.t.,*

$$M = \sum_{i=1}^{n} \lambda_i |x_i\rangle\langle x_i|.$$

Here $\forall i: \ \lambda_i \in \mathbb{C}$.

Note 3. It means that any normal matrix $M = U^*DU$ for some unitary U and diagonal matrix D.

Exercise 21. Show that $|x_i\rangle$ is an eigenvector of M with eigenvalue λ_i.

Note 4. $\langle y|x\rangle$ is a scalar, but $|y\rangle\langle x|$ is a matrix.

Note 5. The λ_i's need not be different. If we collect all the $|x_i\rangle$'s corresponding to a particular eigenvalue λ, the space spanned by those $|x_i\rangle$'s is the eigenspace of λ.

Proof idea. The proof of spectral theorem essentially hinges on the following lemma.

Lemma 1. *Given an eigenspace S (of eigenvalue λ) of matrix M, the matrix M acts on the space S and $S^\perp$ separately. In other words, $M|v\rangle \in S$ if $|v\rangle \in S$ and $M|v\rangle \in S^\perp$ if $|v\rangle \in S^\perp$.*

Proof of lemma. Since S is an eigenspace, $M|v\rangle \in S$ if $|v\rangle \in S$. For a vector $|v\rangle \in S$,

$$MM^*|v\rangle = M^*M|v\rangle = \lambda M^*|v\rangle.$$

This shows that M^* preserves the subspace S. Suppose $|v_1\rangle \in S^\perp,\ |v_2\rangle \in S$, then $M^*|v_2\rangle \in S$. So,

$$0 = \langle v_1|M^*|v_2\rangle = \langle Mv_1|v_2\rangle.$$

Hence $M|v_1\rangle \in S^\perp$. Hence, matrix M acts separately on S and $S^\perp$. □

The lemma implies that M is a linear operator on $S^\perp$, i.e., it moves every element of $S^\perp$ to an element in $S^\perp$ linearly. It can be easily shown that this linear operator (the action of M on $S^\perp$) is also normal. The proof of spectral theorem follows by using induction and is given below.

From the fundamental theorem of Algebra, there is at least one root λ_0 of $det(\lambda I - M) = 0$. Start with the eigenspace of the eigenvalue λ_0. Using Lem. 1, we can restrict the matrix to orthogonal subspace (which is of smaller dimension). We can divide the entire space into orthogonal eigenspaces by induction.

Exercise 22. Show that if we take the orthonormal basis of all these eigenspaces, then we get the required decomposition.

Exercise 23. Given the spectral decomposition of M, what is the spectral decomposition of M^*?

□

Exercise 24. If M is normal, prove that the rank of M is the sum of the dimension of the non-zero eigenspaces.

It is easy to show that any matrix with orthonormal set of eigenvectors is a normal matrix. Hence, spectral decomposition provides another characterization of normal matrices.

Clearly the spectral decomposition is not unique (essentially because of the multiplicity of eigenvalues). But the eigenspaces corresponding to each eigenvalue are fixed. So there is a unique decomposition in terms of eigenspaces and then any orthonormal basis of these eigenspaces can be chosen.

Note 6. It is also true that if an eigenvalue is a root of characteristic polynomial with multiplicity k, then its eigenspace is of dimension k.

The eigenvalues and eigenvectors have more structure if we look at specific classes of normal matrices. We will take a look at these special classes of normal matrices below.

3.1 Hermitian matrix

A matrix M is said to be *Hermitian* if $M = M^*$. It is easy to check that any Hermitian matrix is normal. You can also show that all the eigenvalues of a Hermitian matrix are real (given as an exercise).

Conversely if all the eigenvalues are real for a normal matrix then the matrix is Hermitian (from Spectral theorem). For any matrix B, a matrix of the form B^*B or $B + B^*$ is always Hermitian.

The sum of two Hermitian matrices is Hermitian, but the multiplication of two Hermitian matrices need not be Hermitian.

Exercise 25. Give an example of two Hermitian matrices whose multiplication is not Hermitian.

3.2 Unitary matrix

A matrix M is unitary if $MM^* = M^*M = I$. In other words, the columns of M form an orthonormal basis of the whole space. Unitary matrices need not be Hermitian, so their eigenvalues can be complex. For a unitary matrix, $M^{-1} = M^*$.

Exercise 26. Give an example of a unitary matrix which is not Hermitian.

Unitary matrices can be viewed as matrices which implement a change of basis. Hence they preserve the angle (inner product) between the vectors. So for unitary M,

$$\langle u|v\rangle = \langle Mu|Mv\rangle.$$

Exercise 27. Prove the above equation.

If two matrices A, B are related by $A = M^{-1}BM$, where M is unitary, then they are unitarily equivalent. If two matrices are unitarily equivalent then they are similar.

Spectral theorem can be stated as the fact that normal matrices are unitarily equivalent to a diagonal matrix. The diagonal of a diagonal matrix contains its eigenvalues.

Exercise 28. What is the rank of a unitary matrix?

3.3 Positive semidefinite matrix

A matrix M is positive semidefinite if it is Hermitian and all its eigenvalues are non-negative. If all eigenvalues are strictly positive then it is called a positive definite matrix.

Theorem 3. *For a Hermitian $n \times n$ matrix $M \in L(V)$, following are equivalent.*

1. $\langle v|M|v\rangle \geq 0$ *for all* $|v\rangle \in V$.
2. *All the eigenvalues are non-negative.*
3. *There exists a matrix B, s.t., $B^*B = M$.*

Proof. $1 \Rightarrow 2$: Say λ is an eigenvalue of M. Then there exist an eigenvector $|v\rangle \in V$, s.t., $M|v\rangle = \lambda|v\rangle$. So $0 \leq \langle v|M|v\rangle = \lambda\langle v|v\rangle$. Since $\langle v|v\rangle$ is positive for all $|v\rangle$, implies λ is non-negative.

$2 \Rightarrow 3$: Since the matrix M is Hermitian, it has a spectral decomposition.

$$M = \sum_i \lambda_i |x_i\rangle\langle x_i|$$

Define $|y_i\rangle = \sqrt{\lambda_i}|x_i\rangle$. This definition is possible because λ_i's are non-negative. Then,

$$M = \sum_i |y_i\rangle\langle y_i|.$$

Define B^* to be the matrix whose columns are y_i. Then it is clear that $B^*B = M$. From this construction, B's columns are orthogonal.

Note 7. In general, any matrix of the form B^*B is positive semi-definite. The matrix B need not have orthogonal columns (it can even be rectangular).

But this representation is not unique and there always exists a matrix B with orthogonal columns for M, s.t., $B^*B = M$. This decomposition is unique if B is positive semidefinite. The positive semidefinite B, s.t., $B^*B = M$, is called the square root of M.

Exercise 29. Prove that the square root of a matrix is unique.

Hint: Use the spectral decomposition to find one of the square root. Suppose A is any square root of M. Then use the spectral decomposition of A and show the square root is unique (remember the decomposition to eigenspaces is unique) .

$3 \Rightarrow 1$: We are given a matrix B, s.t., $B^*B = M$. Then,

$$\langle v|M|v\rangle = \langle Bv|Bv\rangle \geq 0.$$

Exercise 30. Prove $2 \Rightarrow 1$ directly.

□

Note 8. A matrix M of the form $M = \sum_i |x_i\rangle\langle x_i|$ is positive semidefinite (Exercise: Prove it), even if x_i's are not orthogonal to each other.

Note 9. A matrix of the form $|y\rangle\langle x|$ is a rank one matrix. It is rank one because all columns are scalar multiples of $|y\rangle$. Similarly, all rank one matrices can be expressed in this form.

4 Singular value decomposition

Singular value decomposition is one of the most important factorizations of a matrix. The statement says,

Theorem 4. *Given a linear operator M in $L(V, W)$. There exists a decomposition of the form:*

$$M = \sum_{i=1}^{r} s_i |y_i\rangle\langle x_i|$$

Where $|x_1\rangle, \cdots, |x_r\rangle$ (called right singular vectors) and $|y_1\rangle, \cdots, |y_r\rangle$ (called left singular vectors) are orthonormal basis of V and W respectively. The numbers $s_1, \cdots, s_r$ (called singular values) are positive real numbers and r itself is the rank of the matrix M.

The statement of the theorem can also be written as $M = A\Delta B^*$, where $A \in L(W), B \in L(V)$ are unitary matrices and Δ is the diagonal matrix of singular values. With this interpretation, any linear operation can be viewed as rotation in subspace V then scaling the standard basis and then another rotation in W subspace.

The statement of singular value decomposition is easy to prove if we don't need any condition on $|y_i\rangle$'s. Any basis of V will be sufficient to construct such a decomposition (why?). We can even choose all singular values to be 1 in that case. But it turns out that with the singular values we can make the y_i's to be orthonormal.

The proof of singular value decomposition follows by applying spectral decomposition on matrices MM^* and M^*M.

Exercise 31. Prove that M^*M and MM^* have the same set of eigenvalues.

Suppose the eigenvectors of M^*M are $|x_i\rangle$'s for eigenvalues λ_i, then eigenvectors of MM^* are $\frac{Mx_i}{\|Mx_i\|} = y_i$'s.

Exercise 32. Prove the above statement.

From the assignment $rank(A) = rank(A^*A)$ and hence it is enough to specify the action of M on x_i's. So,

$$M = \sum_{i=1}^{r} \|Mx_i\| |y_i\rangle\langle x_i|$$

Exercise 33. Prove that $\|Mx_i\| = \sqrt{\lambda_i}$.

This implies the singular value decomposition.

$$M = \sum_{i=1}^{r} \sqrt{\lambda_i} |y_i\rangle\langle x_i| = \sum_{i=1}^{r} s_i |y_i\rangle\langle x_i|.$$

The eigenvectors of MM^* are left singular vectors and eigenvectors of M^*M are right singular vectors of M. The eigenvalues of MM^* or M^*M are the squares of the singular values of M.

We will restrict our attention to normal matrices for the rest of this lecture note.

5 Simultaneously diagonalizable matrices

We saw that a matrix can be diagonalized (by an orthonormal basis) if it is normal. When can two Hermitian matrices be diagonalized by the same orthonormal basis? In other words, when are two Hermitian matrices A, B simultaneously diagonalizable? It turns out that there is a simple criteria to check it.

Theorem 5. *Two normal matrices A, B are simultaneously diagonalizable iff $AB = BA$.*

Proof. If two matrices are simultaneously diagonalizable then they commute. This is given as an assignment.

Let us look at the other direction. Suppose A, B commute. Like the case of spectral decomposition, we will show that the eigenspace of A is preserved by B. More precisely, if V_a is the eigenspace of A corresponding to eigenvalue a. Then operator B takes V_a to itself. Let $|v\rangle \in V_a$,

$$AB|v\rangle = BA|v\rangle = aB|v\rangle \Rightarrow B|v\rangle \in V_a.$$

So we can consider that restriction of B on V_a.

Exercise 34. Show that this matrix is normal.

Using spectral decomposition of the restriction of B, divide V_a into eigenspaces $V_{a,b}$ for eigenvalue b of B. The bases of $V_{a,b}$ for all a, b is the orthonormal basis which will diagonalize both matrices A and B.

□

6 Operator functions

The first notion is of applying a function on a linear operator. We will assume that the linear operators given to us belong to the set of normal operators or some subset of it.

Suppose we have a function, $f : \mathbb{C} \to \mathbb{C}$, from complex numbers to complex numbers. It can naturally extended to be a function on a normal linear operator in $L(\mathbb{C}^n)$. By definition of operator function, we apply the function on all the eigenvalues of the operator. So, if

$$A = \lambda_1 x_1 x_1^* + \cdots + \lambda_n x_n x_n^*.$$

then

$$f(A) = f(\lambda_1) x_1 x_1^* + \cdots + f(\lambda_n) x_n x_n^*.$$

In particular, we can now define the square-root, exponential and logarithm of an operator.

Exercise 35. Check that this definition of square-root agrees with the definition of square root defined in the previous lecture notes.

Note 10. This means that we define square root only for positive semi-definite operators.

Pauli matrices are used widely in quantum computing. They are defined as,

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \; Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \; Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

X is the quantum NOT gate and Z is known as the phase gate.

Exercise 36. Show that the Pauli matrices are Hermitian as well as Unitary by calculating their eigenvalue.

Exercise 37. Show that the Pauli matrices (with identity) form a basis of all Hermitian 2×2 operators.

Exercise 38. Find e^{iX}, e^{iY}, e^{iZ}.

Another very important function on operators introduced before is *trace*. We defined trace to be $Tr(A) = \sum_i A_{ii}$. At this point, it is a function on matrices and not linear operators.

Exercise 39. What is the problem?

For a linear operator, trace might be different for different bases. In other words, there is no guarantee that it is independent of the basis (from the definition given above).

Exercise 40. Show that the trace is cyclic, i.e., $tr(AB) = tr(BA)$.

This exercise implies that $tr(U^*AU) = tr(A)$. Hence, trace is independent of the representation.
We also know that $\langle v|A|w\rangle = \sum_{ij} A_{ij} v_i^* w_j$.

Exercise 41. Show that $A_{ij} = \langle i|A|j\rangle$, where matrix A is represented in the standard basis $|1\rangle, \cdots, |n\rangle$.

From the previous exercise, $tr(A) = \sum_i \langle i|A|i\rangle$. In fact, for any orthonormal basis $v_1, \cdots, v_n$,

$$tr(A) = \sum_i \langle v_i|A|v_i\rangle.$$

If we take v_i to be the eigenvectors,

$$tr(A) = \sum_i \lambda_i.$$

Here, λ_i are the eigenvalues of the operator A.

7 Tensor product

We have talked about the concept of tensor product before, it was used to combine the two systems.

Suppose there is a ball which can be colored blue or red. The state of a "quantum" ball is a vector in two dimensions,

$$|v\rangle = \alpha|r\rangle + \beta|b\rangle.$$

Where $|r\rangle, |b\rangle$ represent the classical states, the ball being red or blue, the coefficients α, β follow the usual law.

How about if there are two different balls. The classical states possible are $|rr\rangle, |rb\rangle, |br\rangle, |bb\rangle$, i.e., we take the set multiplication of possible states of individual system.

What are the possible states if this system is quantum?

$$|v\rangle = \alpha|rr\rangle + \beta|rb\rangle + \gamma|br\rangle + \delta|bb\rangle.$$

This idea motivates the definition of tensor product. Given two vector spaces V, W equipped with an inner product and spanned by the orthonormal basis $v_1, v_2, \cdots, v_n$ and $w_1, w_2, \cdots, w_m$, the tensor product $V \otimes W$ is the space spanned by the mn vectors $(v_1 \otimes w_1), \cdots, (v_1 \otimes w_n), (v_2 \otimes w_1), \cdots, (v_n \otimes w_m)$.

Exercise 42. What is the dimension of space $V \otimes W$?

The tensor product should satisfy the following conditions.

– Scalar multiplication: for $\alpha \in \mathbb{C}, v \in V$ and $w \in W$,

$$\alpha(v \otimes w) = (\alpha v) \otimes w = v \otimes (\alpha w).$$

– Linearity in the first component: for $v_1, v_2 \in V$ and $w \in W$,

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

– Linearity in the second component: for $v \in V$ and $w_1, w_2 \in W$,

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

We can define the tensor product of two vectors in a canonical way for the vector spaces $\mathbb{C}^n$ and $\mathbb{C}^m$. The tensor product of two vectors $a = (a_1, \cdots, a_n) \in V$ and $b = (b_1, \cdots, b_m)$ is the vector $a \otimes b \in V \otimes W$,

$$a \otimes b = \begin{pmatrix} a_1 b_1 \\ a_1 b_2 \\ \vdots \\ a_n b_m \end{pmatrix}$$

In Dirac's notation, we will simply write $|vw\rangle$ instead of $|v\rangle \otimes |w\rangle$.

Exercise 43. Show that $(v + w) \otimes (a + b) = v \otimes a + v \otimes b + w \otimes a + w \otimes b$.

We can define the inner product on the tensor product space in the natural way,

$$\langle a \otimes b | c \otimes d \rangle = \langle a|c\rangle \langle b|d\rangle.$$

Given two linear operators $A \in L(V)$ and $B \in L(W)$, their tensor product is in space $L(V \otimes W)$. The action is specified by,

$$(A \otimes B)(a \otimes b) = Aa \otimes Bb.$$

Extend this by linearity to define the action on the complete space $V \otimes W$.

Exercise 44. Given the matrix representation of A, B, come up with matrix representation of $A \otimes B$.

Exercise 45. Write out the matrix representation of $H^{\otimes 2} = H \otimes H$ where H is the Hadamard Matrix.

$A \otimes B$ are linear operators in $L(V \otimes W)$. Can there be other linear operators?

The sum of two linear operators is a linear operator. So any operator of the form $\sum_i c_i (A_i \otimes B_i)$ is also a linear operator.

Are there any more linear operators? It turns out that these are the only linear operators in $L(V \otimes W)$, you will prove this in the assignment.

7.1 Partial trace

Partial trace is a linear operator on the tensor product of the operators. Given two vector spaces, V and W, let $A \in L(V)$ and $B \in L(W)$ be two linear operators. The partial trace Tr_W is a linear operator from $L(V \otimes W)$ to $L(V)$, s.t.,

$$Tr_W(A \otimes B) = ATr(B).$$

We can similarly define Tr_V.

Exercise 46. Show that the partial trace defined above is a linear operator.

Partial trace is mostly used in quantum computing to understand the state/operation on a part of a composite system.

8 Direct sum

A closely related concept is called direct sum. Notice the difference in the motivation and definition.

Suppose you have a ball which can be colored red or blue. Then the "quantum state" of that ball could be any superposition of these classical states,

$$|v\rangle = \alpha|r\rangle + \beta|b\rangle.$$

Where $|r\rangle, |b\rangle$ represent the classical states, the ball being red or blue, the coefficients α, β follow the usual law. One day you discover that ball can also have a yellow color. Then the basis states become $|r\rangle, |b\rangle, |y\rangle$ and the state of the ball can be,

$$|v\rangle = \alpha|r\rangle + \beta|b\rangle + \gamma|y\rangle.$$

This idea gives rise to the concept of direct sum of two vector spaces. Given two vector spaces V, W spanned by basis $v_1, v_2, \cdots, v_n$ and $w_1, w_2, \cdots, w_m$, the direct sum $V \oplus W$ is the space spanned by the vectors $(v_1, 0), \cdots, (v_n, 0), (0, w_1), \cdots, (0, w_m)$. Here the zeroes in the first n entries is the vector 0 of dimension m. In the last m entries it is the 0 vector with dimension n. Similar to tensor product, direct sum of two vector spaces should satisfy the following two conditions,

- Scalar multiplication:
$$\alpha(v, w) = (\alpha v, \alpha w).$$
- Linearity:
$$(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2).$$

Exercise 47. What is the dimension of this vector space $dim(V) \oplus dim(W)$.

Given matrices $A \in L(V)$ and $B \in L(W)$, their direct sum is in $L(V \oplus W)$.

$$A \oplus B = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

Exercise 48. Why did we define direct sum of matrices this way?

The description of both direct sum as well as tensor product is given in a very simplified manner in terms of basis vectors, sufficient for use in our course. Readers are encouraged to check out the formal definitions.

9 Assignment

Exercise 49. Show that the matrix M and M^* have the same singular values.

Exercise 50. Prove that the eigenvalues of a Hermitian matrix are real.

Exercise 51. Prove that the absolute value of the eigenvalues of an unitary matrix is 1. Is the converse true. What condition do we need to get the converse?

Exercise 52. Show that a matrix M is positive semi-definite if it is the Gram matrix of vectors $|u_1\rangle, \cdots, |u_n\rangle$. That is,

$$M_{ij} = \langle u_i | u_j \rangle.$$

Exercise 53. Read about polar decomposition and how to get singular decomposition using polar decomposition.

Exercise 54. Prove that a matrix M is Hermitian iff $\langle v|M|v\rangle$ is real for all $|v\rangle$.

Exercise 55. Show that if two matrices are simultaneously diagonalizable then they commute.

Exercise 56. Show that the set of Hermitian matrices of a fixed dimension form a vector space (over which field?). What is the dimension of this vector space?

Exercise 57. Read about polar decomposition and prove it using singular value decomposition.

Exercise 58. Prove that $rank(AB) \leq rank(A)$.

Exercise 59. Prove that $rank(A) = rank(A^*A)$ without using singular or spectral decomposition.

Hint: $rank(A) \geq rank(A^*A)$ is easy. For the other direction, reduce A to its reduced row echelon form.

Exercise 60. What are the eigenvalues of $A \otimes B$, where A, B are normal matrices?

Exercise 61. What are the eigenvalues of $A \oplus B$, where A, B are normal matrices?

Exercise 62. Give a characterization of the linear operators over $V \otimes W$ in terms of linear operators over V and W. Remember that they form a vector space.

Exercise 63. Show that $\langle v|A|w\rangle = \sum_{ij} A_{ij} v_i^* w_j$.

Exercise 64. Let $\sigma = \alpha_1 X + \alpha_2 Y + \alpha_3 Z$, where $|\alpha_1|^2 + |\alpha_2|^2 + |\alpha_3|^2 = 1$. Show that,

$$e^{i\theta\sigma} = \cos(\theta) I + i \sin(\theta)\sigma.$$

Exercise 65. Show that $tr(A|v\rangle\langle v|) = \langle v|A|v\rangle$.

Exercise 66. Prove that if H is Hermitian then e^{iH} is a unitary matrix.

References

1. M. A. Nielsen and I. L. Chuang. Quantum computation and quantum information. *Cambridge*, 2010.
2. S. Arora and B. Barak. Computational Complexity: A modern approach. *Cambridge*, 2009.
3. R. Lidl and H. Niederreiter. Finite Fields. *Cambridge University Press*, 1997.
4. B. Kleinberg. Course notes: Introduction to algorithms. *http://www.cs.cornell.edu/courses/cs4820/2010sp/handouts/MillerRabin.pd* 2010.
5. D. R. Simon. On the power of quantum computation. *Foundations of Computer Science, 1994 Proceedings., 35th Annual Symposium on: 116123*, 1994.
6. A. Childs. Course notes: Quantum algorithms. *https://cs.umd.edu/ amchilds/teaching/w13/qic823.html*, 2013.