CS419 Linear Algebra

December 2019

1 What do we need to know?

By the end of this booklet, you should know all the linear algebra you need for CS419. More specifically, you'll understand:

- Vectors spaces (the important ones, at least)
- Dirac (bra-ket) notation and why it's quite nice to use
- Linear combinations, linearly independent sets, spanning sets and basis sets
- Matrices and linear transformations (they're the same thing)
- Changing basis
- Inner products and norms
- Unitary operations
- Tensor products
- Why we care about linear algebra

2 Vector Spaces

In quantum computing, the 'vectors' we'll be working with are going to be made up of complex numbers[1]. A **vector space**, V, over $\mathbb{C}$ is a set of vectors with the vital property[2] that $\alpha u + \beta v \in V$ for all $\alpha, \beta \in \mathbb{C}$ and $u, v \in V$. Intuitively, this means we can add together and scale up vectors in V, and we know the result is still in V.

Our vectors are going to be lists of n complex numbers, $v \in \mathbb{C}^n$, and $\mathbb{C}^n$ will be our most important vector space. Note we can just as easily define vector spaces over $\mathbb{R}$, the set of real numbers. Over the course of this module, we'll see the reasons[3] we use $\mathbb{C}$, but for all this linear algebra, we can stick with $\mathbb{R}$ as everyone is happier with real numbers. Rest assured for the entire module, every time you see something like "Consider a vector space V", this vector space will be $\mathbb{R}^n$ or $\mathbb{C}^n$ for some $n \in \mathbb{N}$.

We'll see soon why we care about vector spaces, once we've completely changed up the notation!

3 Dirac Notation (and why we use it)

Dirac was a physicist in the 1900s known to be very precise in the work he did, and as we will see as we delve deeper into linear algebra, the idea of a function of a vector and the idea of a vector gets very blurred. He came up with notation that tries to differentiate between these two concepts. Other physicists working in quantum mechanics liked this, so it stuck, and everything you read about in quantum

[1] The set of complex numbers being denoted $\mathbb{C}$, and the set of real numbers being $\mathbb{R}$.

[2] Note this is a simplified definition for this course, as this is really all we need to know, mathematicians will tell you how actually you have to worry about field axioms, and they'll say words like 'commutativity' and 'distributivity', but have no fear, they did all the hard work in the 1800s to win us this freedom.

[3] $\mathbb{C}$ is really nice to work with, it looks all complicated, but it's much nicer than the reals.

mechanics uses this notation, so it's well worth using and getting used to.

The notation uses the idea of the 'bra' and the 'ket'. Vectors as we know them are 'kets', and are denoted $|v\rangle$. From now on, we won't say we have a vector $v \in \mathbb{R}^n$, we'll say we have $|v\rangle \in \mathbb{R}^n$.

So, now we move on to 'bras', and more importantly the idea of a function of a vector. To start with, the simple definition is that for a vector $|v\rangle$, the bra is $\langle v| = |v\rangle^\dagger$, the conjugate transpose[4] of $|v\rangle$. The deeper meaning comes from considering a linear functional[5] $f : \mathbb{R}^2 \to \mathbb{R}$, and a vector $|v\rangle = \begin{bmatrix} v_0 \\ v_1 \end{bmatrix}$ with $f(|v\rangle) = v_0 + v_1$. This is actually equivalent to simply multiplying $|v\rangle$ by the vector $\begin{bmatrix} 1 & 1 \end{bmatrix}$ and this is our bra. Our bra notation is $\langle u|$, and so instead of writing $f(|v\rangle)$, we write $\langle u||v\rangle$, shortened to $\langle u|v\rangle$.

Fundamentally, all you need to know is that, for a vector $v \in V$, we now denote v by $|v\rangle$, and we denote the conjugate transpose of v, $v^\dagger$, by $\langle v|$.

And that's Dirac's Bra-Ket Notation. It's gonna appear everywhere in this module from now on. Whilst we're dealing with notation, from here onwards in this module, in $\mathbb{R}^n$ we say

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, |1\rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \cdots, |n-1\rangle = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

4 Linear Combinations

We've seen so far that vectors spaces are closed, meaning $\alpha|u\rangle + \beta|v\rangle \in \mathbb{R}^n$ for all $\alpha, \beta \in \mathbb{R}$ and $|u\rangle, |v\rangle \in \mathbb{R}^n$. This $\alpha|u\rangle + \beta|v\rangle$ is what we call a **linear combination**.

Definition 1. *A linear combination of a set of vectors $|v_0\rangle, ..., |v_k\rangle \in \mathbb{R}^n$ is a vector*

$$v = \alpha_0|v_0\rangle + ... + \alpha_k|v_k\rangle$$

Where $\alpha_0, ..., \alpha_k \in \mathbb{R}^n$.

Given a set of vectors $|v_0\rangle, ..., |v_k\rangle \in \mathbb{R}^n$, we can find all possible linear combinations of these vectors, and this itself will form a vector space.

Definition 2. *The **span** of a set of vectors, $S = |v_0\rangle, ..., |v_k\rangle \in \mathbb{R}^n$, $\text{Span}(S)$ is the set of all linear combinations of S.*

Example 1. *For $S = \{|0\rangle, |1\rangle, |0\rangle + |1\rangle\} \subset \mathbb{R}^3$, $\text{Span}(S)$ is the set of all vectors of the form $\begin{bmatrix} \alpha \\ \beta \\ 0 \end{bmatrix}$ for $\alpha, \beta \in \mathbb{R}$.*

Exercise 1. *Show that $\text{Span}(S)$ for an arbitrary $S = |v_0\rangle, ..., |v_k\rangle \in \mathbb{R}^n$ is a vector space. (Hint: Show that $\alpha|u\rangle + \beta|v\rangle \in \text{Span}(S)$ for all $\alpha, \beta \in \mathbb{R}$ and $|u\rangle, |v\rangle \in \text{Span}(S)$.)*

From example 1, we can see that the set $\{|0\rangle, |1\rangle\} \subset \mathbb{R}^3$ would have the same spanning set as S, because $|0\rangle + |1\rangle$ is a[6] linear combination of $|0\rangle$ and $|1\rangle$. On the other hand, $|0\rangle$ is *not* a linear combination of $|1\rangle$, and vice-versa. These two vectors are thus **linearly independent**.

[4] i.e. the vector is transposed (flipped from $v \in \mathbb{C}^{n\times 1}$ to $v^T \in \mathbb{C}^{1\times n}$ or vice-versa) and the complex parts have their sign flipped (go from $v_k = a + bi$ to $\overline{v_k} = a - bi$).

[5] We'll define this properly later, basically it just means a linear function, i.e. $f(\alpha|u\rangle + \beta|v\rangle) = \alpha f(|u\rangle) + \beta f(|v\rangle)$, with the extra property that $f : V \to \mathbb{R}$ or $\mathbb{C}$ (the underlying field).

[6] rather unimaginative

Definition 3. *A set of vectors $S = \{|v_0\rangle, ..., |v_k\rangle\} \subset \mathbb{R}^n$ is linearly independent if no vector of S is a linear combination of the others. Equivalently,*

$$\alpha_0|v_0\rangle + ... + \alpha_k|v_k\rangle = 0$$

implies

$$\alpha_0 = ... = \alpha_k = 0.$$

Now we can start with the real juicy linear algebra, and discuss basis and dimension. A **basis** of a vector space V is a linearly independent set of vectors, S, such that Span(S) $= V$.

Example 2. *Consider a vector space V. Then* Span$(V) = V$.

This isn't a basis however, as V is not linearly independent.

Exercise 2. *For $\mathbb{R}^2$, show the set $S = \{|0\rangle, |1\rangle\}$ is a basis.*

Theorem 1. *Every basis of a vector space V has the same size. We call this number the* ***dimension of V****, and denote it* Dim(V).

For $\mathbb{R}^n$, Dim$(\mathbb{R}^n)$ is n[7]. The canonical basis of $\mathbb{R}^n$ is the set $\{|i\rangle : 0 \leq i < n\}$. In quantum computing we sometimes call this the computational basis.

We've alluded to the fact that a the span of a set of vectors is also a vector space.

Definition 4. *A* ***subspace*** *U of a vector space V is a subset that is itself a vector space.*

Example 1 demonstrates this concept, with Span(S) being a subspace of $\mathbb{R}^3$. Note that within $\mathbb{R}^n$ and $\mathbb{C}^n$, all subspaces will be equivalent[8] to $\mathbb{R}^m$ or $\mathbb{C}^m$ for some $m \leq n$.

An important thing to note is that the basis we're using is very important. Most of the time we'll be using the canonical basis, however we'll see cases where other basis are more useful (and unavoidable).

Example 3. *Consider a vector $|v\rangle \in \mathbb{R}^3$, in the canonical basis (whenever a basis isn't specified, it'll be in the canonical basis).*

$$|v\rangle = \begin{bmatrix} 5 \\ 3 \\ 1 \end{bmatrix}$$

Written as this, we're actually being told 'how much' of each basis vector $|v\rangle$ is using, $|v\rangle = 5|0\rangle + 3|1\rangle + 1|2\rangle$. What if we used a different basis?

$$T = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} \right\}$$

In this basis, $|v\rangle$ wouldn't be written as above, we'd write

$$|v^{(T)}\rangle = \begin{bmatrix} 3 \\ -2 \\ 2 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - 2 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}$$

Note $|v\rangle$ and $|v^{(T)}\rangle$ are describing exactly the same point in $\mathbb{R}^3$, just one is using the canonical basis, one is using the basis T.

Changing the basis you're in is a remarkably useful thing to be able to do, and we'll see how to do it once we understand linear transformations.

[7] For $\mathbb{C}^n$, it depends: $\mathbb{C}$ can be seen either as a vector space of dimension 2 over $\mathbb{R}$ (by "forgetting" the operation of multiplication of complex numbers and just treating a complex number as $a + bi = (a, b) \in \mathbb{R}^2$) or over $\mathbb{C}$.

[8] the technical word is "isomorphic", and means that the number of scalars needed to identify a vector in each of the vector spaces is the same. This number is (not coincidentally) the dimension, so all vector spaces of the same dimension are isomorphic.

5 Linear Transformations

In section 3 we discussed how linear functionals (linear functions mapping into the underlying field) can be represented equivalently by multiplying a vector by a 'bra'. We now want to discuss linear functions that map vector spaces to vector spaces.

Definition 5. *A function (transformation) $f : V \to W$ is linear if*

$$f(\alpha|u\rangle + \beta|v\rangle) = \alpha f(|u\rangle) + \beta f(|v\rangle)$$

For all $\alpha, \beta \in \mathbb{R}$ and $|u\rangle, |v\rangle \in V$.

Linear transformations are incredibly useful, and in this module you'll see how everything in quantum computing can be done with a just a subset of linear transformations.

We can describe how a linear transformation $f : \mathbb{R}^3 \to \mathbb{R}^3$ (for example) works by simply writing out what it does to each $|v\rangle$, e.g.

$$f(|v\rangle) = f\left(\begin{bmatrix} v_0 \\ v_1 \\ v_2 \end{bmatrix}\right) = \begin{bmatrix} v_0 + v_1 + 3v_2 \\ v_1 - 4v_2 \\ v_0 + 6v_1 \end{bmatrix}$$

And we can work with this, this tells us what f does, however, what if we wanted to compute $f(f(|v\rangle))$? We want a better way to describe linear transformations.

A useful property of linear transformations is that they are uniquely determined by their action on a basis! That mouthful of a sentence tells us that a linear transformation $f : V \to W$ can be described by considering solely how it acts on the basis vectors of V, and these can then be used to determine what it is doing within W, due to the linearity of the function.

Example 4. *Take the f we described before, and consider how it acts on the canonical basis of $\mathbb{R}^3$.*

$$f(|0\rangle) = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad f(|1\rangle) = \begin{bmatrix} 1 \\ 1 \\ 6 \end{bmatrix}, \quad f(|2\rangle) = \begin{bmatrix} 3 \\ -4 \\ 0 \end{bmatrix}$$

Now if we wanted to evaluate $f(|v\rangle)$, this will be $f(v_0|0\rangle + v_1|1\rangle + v_2|2\rangle) = v_0 f(|0\rangle) + v_1 f(|1\rangle) + v_2 f(|2\rangle)$. A useful shorthand here would be to stick the three vectors together into one set of brackets, and then write our vector $|v\rangle$ next to it, as follows

$$f(|v\rangle) = v_0 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + v_1 \begin{bmatrix} 1 \\ 1 \\ 6 \end{bmatrix} + v_2 \begin{bmatrix} 3 \\ -4 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 3 \\ 0 & 1 & -4 \\ 1 & 6 & 0 \end{bmatrix} \begin{bmatrix} v_0 \\ v_1 \\ v_2 \end{bmatrix}$$

We'll call the block of numbers a ***matrix***[9]*.*

Matrices are simply useful shorthand for writing linear transformations, and the above example gives an idea as to why we multiply matrices the unusual way we do.

The fact every matrix yields a linear transformation comes by definition, however, the converse also holds.

Theorem 2. *For every linear transformation $f : V \to W$, there exists a matrix A such that $f(|v\rangle) = A|v\rangle = |w\rangle$ for $|v\rangle \in V$, $|w\rangle \in W$.*

6 Changing Basis

Likely everything you've read so far you're already pretty happy with from first year and the like, however, for whatever reason, changing basis is always a bit of nightmare. Fortunately in day-to-day life it

[9] !!!

doesn't come up too often. Unfortunately, it comes up a lot in quantum computation[10].

First off we'll consider how to simply change between two bases of the same vector space. Let's say for a d-dimensional[11] vector space V, and pick two bases $S = \{|s_0\rangle, ..., |s_{d-1}\rangle\}$ and $T = \{|t_0\rangle, ..., |t_{d-1}\rangle\}$[12]. We want to come up with a function $f_{ST} : \mathbb{R}^d \to \mathbb{R}^d$ that takes the coordinates of a vector written on the basis S to the coordinates of *the same* vector, but written on the basis T.

Example 5. *Returning to Example 3, we saw two bases for $\mathbb{R}^3$. Let $S = \{|0\rangle, |1\rangle, |2\rangle\}$ and T be as in the example. We want to end up with a linear transformation f_{ST} such that*

$$f_{ST}\left(\begin{bmatrix}5\\3\\1\end{bmatrix}\right) = \begin{bmatrix}3\\-2\\2\end{bmatrix}$$

I.e. $f(|v^{(S)}\rangle) = |v^{(T)}\rangle$, where $|v^{(S)}\rangle$ is in basis S, and $|v^{(T)}\rangle$ is in basis T. Using the linearity tricks in Example 4, we see

$$f_{ST}\left(\begin{bmatrix}v_0^{(S)}\\v_1^{(S)}\\v_2^{(S)}\end{bmatrix}\right) = v_0^{(S)} f_{ST}(|0\rangle) + v_1^{(S)} f_{ST}(|1\rangle) + v_2^{(S)} f_{ST}(|2\rangle) = \begin{bmatrix}v_0^{(T)}\\v_1^{(T)}\\v_2^{(T)}\end{bmatrix}$$

What does this tell us? If we consider the matrix C_{ST} representing f_{ST}, we can see that the columns of C_{ST} will be the representations of the 'old' basis vectors S in the 'new' basis T. In our particular case, this means finding how to represent $|0\rangle$, $|1\rangle$ and $|2\rangle$ in T.

$$|0\rangle = 0\begin{bmatrix}1\\1\\1\end{bmatrix} + 0\begin{bmatrix}1\\0\\1\end{bmatrix} + \frac{1}{2}\begin{bmatrix}2\\0\\0\end{bmatrix} \quad \textit{i.e. } |0^{(T)}\rangle = \begin{bmatrix}0\\0\\\frac{1}{2}\end{bmatrix}$$

For $|1\rangle$ and $|2\rangle$ we get

$$|1^{(T)}\rangle = \begin{bmatrix}1\\-1\\0\end{bmatrix}, \quad |2^{(T)}\rangle = \begin{bmatrix}0\\1\\-\frac{1}{2}\end{bmatrix}$$

Hence

$$C_{ST} = \begin{bmatrix}0 & 1 & 0\\0 & -1 & 1\\\frac{1}{2} & 0 & -\frac{1}{2}\end{bmatrix}$$

To check, we can now see if this works with Example 3,

$$f_{ST}\left(\begin{bmatrix}5\\3\\1\end{bmatrix}\right) = \begin{bmatrix}0 & 1 & 0\\0 & -1 & 1\\\frac{1}{2} & 0 & -\frac{1}{2}\end{bmatrix}\begin{bmatrix}5\\3\\1\end{bmatrix} = \begin{bmatrix}3\\-2\\2\end{bmatrix}$$

Isn't linear algebra incredible?

Once we have the change of basis matrix (C_{ST}) for one direction, it's easy to find the change of basis matrix for the other direction, C_{TS}.

Theorem 3. *For two bases $S = \{|s_0\rangle, ..., |s_{d-1}\rangle\}$ and $T = \{|t_0\rangle, ..., |t_{d-1}\rangle\}$ of a vector space V, and a change of basis matrix $C_{ST} \in \mathbb{R}^{d\times d}$, then the change of basis matrix $C_{TS} = C_{ST}^{-1}$.*

As a quick summary, for those who just want to know how to do it:

- We have two bases, $S = \{|s_0\rangle, ..., |s_{d-1}\rangle\}$ and $T = \{|t_0\rangle, ..., |t_{d-1}\rangle\}$.
- The change of basis matrix from S to T, C_{ST} is the $d \times d$ matrix with columns $|{C_{ST}}_0^{\downarrow}\rangle$,...,$|{C_{ST}}_{d-1}^{\downarrow}\rangle$ with $|{C_{ST}}_i^{\downarrow}\rangle = |s_i\rangle$ written in the basis T.

[10] As a little taster, we need this in order to make sense of measurements in quantum computers, before long you'll be changing basis in order to test Elitzur-Vaidman Bombs, and basically do anything in this module.

[11] Remember in our setting, this means we're dealing with $\mathbb{R}^d$ or $\mathbb{C}^d$.

[12] See how we're using Dirac's notation here? If we weren't we wouldn't know what was going on, if I wrote $S = \{s_0, ..., s_{d-1}\}$ is s_0 the $0th$ element of a vector s? It'd be chaos. Thanks Dirac.

- The change of basis matrix from T to S is C_{ST}^{-1}.

Now, if we have a linear transformation $f : V \to W$, f is going to take a vector in V, in some chosen basis S_V (likely the canonical basis), and will transform it into a vector in W, in some chosen basis S_W. Note in example 4 we used the canonical basis for S_V and S_W.

There will be times when we don't want to always be in the canonical basis, and we start in V using the canonical basis, but shift to W in a different basis. Luckily, we can fairly naturally use our change of basis matrices.

Example 6. *Say we have a linear transformation $f : V \to W$ with $\text{Dim}(V) = n$ and $\text{Dim}(W) = m$. Consider four bases*

$$G = \{|g_0\rangle, ..., |g_{n-1}\rangle\}$$
$$H = \{|h_0\rangle, ..., |h_{n-1}\rangle\}$$
$$S = \{|s_0\rangle, ..., |s_{m-1}\rangle\}$$
$$T = \{|t_0\rangle, ..., |t_{m-1}\rangle\}$$

Where G and H are bases of V, and S and T are bases of W. We can, as in Example 5 find two matrices C_{GH} and C_{ST} that are the change of basis matrices for the two respective vector spaces.

Remembering Section 5 using basis G for vector space V and basis S for vector space W, we can create a matrix[13] A that represents $f : V \to W$ in these basis sets. Similarly, we could choose basis H for V, and T for W, and this would give us a different matrix for f, B.

We can find the the matrix B for our function if we have access to A, C_{GH} and C_{ST}. Consider the following diagram

$$\begin{array}{ccc} \textit{Vector Space} & f & \textit{Vector Space} \\ V & \to & W \\ & & \\ & A & \\ \textit{Basis } G & \to & \textit{Basis } S \\ & & \\ C_{GH} \downarrow \quad \uparrow C_{GH}^{-1} & & C_{ST} \downarrow \quad \uparrow C_{ST}^{-1} \\ & B & \\ \textit{Basis } H & \to & \textit{Basis } T \end{array}$$

If we have a vector $|v^{(G)}\rangle \in V$ in basis G, we can get $|w^{(S)}\rangle \in W$ in basis S by computing $A|v^{(G)}\rangle = |w^{(S)}\rangle$. If we have a vector $|v^{(H)}\rangle \in V$ in basis H, then we can find $w^{(S)}$ by following the arrows from H to S, and performing those operations, i.e. $C_{ST}^{-1}B|v^{(H)}\rangle = |w^{(S)}\rangle$. If we want to find what B, we can follow the arrows to see what composition of operations it's the same as, which will be

$$B = C_{GH}^{-1}AC_{ST}$$

7 Inner Product and Norms

Remember when we spoke about Dirac notation, and I was writing about bras, and $\langle u|v\rangle$? Well now it's time for that to become relevant!

You'll all know what the inner product is, in old notation, we'd have u, v, and we want to find

$$u^\dagger v = \sum_{i=0}^{\text{Dim}(V)} \overline{u_i} v_i$$

Where $u^\dagger$ is the conjugate transpose[14] of u. Remember in our new notation, $u^\dagger$ is $\langle u|$, so the inner product of $|u\rangle$ and $|v\rangle$ will be written as $\langle u|v\rangle$. You'll see this a lot.

[13] Note $A \in \mathbb{R}^{m\times n}$ (or $A \in \mathbb{C}^{m\times n}$, depending on the underlying field of V and W)
[14] Or just u transpose when we have $V = \mathbb{R}^n$.

But what does it mean? You'll see soon what it means in terms of quantum computing[15]. In terms of linear algebra, the inner product allows us to geometrically interpret vectors, with an inner product we can define the length of a vector, and the angle between two vectors.

The concept of the length of a vector is quite involved, and mathematicians will spend much of their second year[16] learning about it. For us, we'll simply define a thing called a 'norm', denoted $||\cdot||$ and the norm of a vector will tell us it's length.

$$\big|\big||v\rangle\big|\big| = \sqrt{\langle v|v\rangle}$$

That is, the square root of the inner product of the vector with itself. For the spaces we're worried about, this works out nicely, in the reals this gives us the 'euclidean distance', which is effectively the distance we'd get if we used a ruler[17].

An important type of vector is a **unit vector** or **normalised vector**, a vector of length 1, i.e. $\sqrt{\langle v|v\rangle} = \langle v|v\rangle = 1$. Unit vectors are tremendously important in quantum computing[18]. Given an arbitrary vector $|v\rangle$, we can **normalise** the vector by finding

$$|v'\rangle = \frac{|v\rangle}{\sqrt{\langle v|v\rangle}}$$

I.e. scaling the vector down by its length, so its length is one[19].

Now we'll briefly deal with how to think about inner product and angles. Given two vectors, $|u\rangle$ and $|v\rangle$, the angle between these two vectors is

$$\theta = \cos^{-1}\left(\frac{\langle u|v\rangle}{\sqrt{\langle u|u\rangle\langle v|v\rangle}}\right)$$

You don't need to explicitly worry about this complicated looking formula too much, but note that if $\theta = 90°$,[20] i.e. we have a right angle, then we call these vectors **orthogonal**, and you can see that $\langle u|v\rangle = 0$. If the two vectors are also unit vectors, we call them **orthonormal**, and if our basis is made up of orthonormal vectors, we call it an **orthonormal basis**[21]. We like orthonormal bases, they've got tonnes of useful properties, which will become apparent when we talk about unitary operations. The canonical basis is an example of an orthonormal basis.

8 Unitary Operations

Now we're ready to see the main reason we need all the preceding linear algebra, **unitary operators**. These are the key tool for every operation we'll do in quantum computing. For us, the key feature of a unitary operation is that it preserves distances, i.e. if we have a unitary operator[22] $U \in \mathbb{R}^{n\times n}$ and $|v\rangle \in \mathbb{R}^n$, then

$$\big|\big||v\rangle\big|\big| = \big|\big|U|v\rangle\big|\big|$$

As it happens, this property is equivalent to many different statements.

Theorem 4. *Any of the following four conditions on a matrix U are equivalent:*

1. *U preserves distances, that is, $||U|v\rangle|| = |||v\rangle||$ for all $|v\rangle \in V$.*
2. *$U^\dagger = U^{-1}$, so $U^\dagger U = I$, the identity matrix*[23]*.*

[15] Spoiler alert: Measurements.
[16] and if they choose to, their career
[17] A potentially multi-dimensional ruler.
[18] Spoiler alert: Qubits.
[19] Why not check that the length of $|v'\rangle$ is indeed one, that'll be a treat.
[20] Degrees, radians, it doesn't matter, use what you want.
[21] because they're orthonormal, and a basis.
[22] Operator, function, matrix, for the linear algebra and maths we're doing, they're equivalent. I like to just think of every function we're using as just meaning some matrix.
[23] The matrix representing an operator that doesn't do anything to the input, so $Iv = v$, it's a matrix with ones along the diagonal, and zeros everywhere else.

3. *The columns of U form an orthonormal basis.*

4. *U preserves inner products, that is, $\langle u|A^\dagger A|v\rangle = \langle u|v\rangle$ for all $|u\rangle, |v\rangle \in V$.*

The last statement, with $\langle u|A^\dagger A|v\rangle$, looks pretty complicated, but it's just our beloved bra-ket notation[24]. As mentioned, we like unitary operations as they preserve distance, so when we do unitary operations on unit vectors, they stay unit vectors.

Example 7. *Here are some orthogonal matrices in $\mathbb{R}^2$:*

$$\begin{bmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{bmatrix}$$ *This is a rotation of θ around the origin*

$$\begin{bmatrix}\cos(2\theta) & \sin(2\theta)\\ \sin(2\theta) & -\cos(2\theta)\end{bmatrix}$$ *This is a reflection along a line at angle θ above the horizontal axis*

We can check they are orthogonal by multiplying each matrix by its transpose, and using statement 1 of theorem 4.

There isn't too much to say on unitary matrices yet, just remember they're distance preserving, and $U^\dagger = U^{-1}$ and you'll go far.

9 Tensor Products

You know how to multiply two numbers, you know how to multiply two vectors, you know how to multiply two matrices, but imagine, just imagine, being able to multiply two vector spaces! The **tensor product** of two vector spaces is a vector space formed by 'multiplying' two vector spaces together.

To start with, we introduce the **outer product**, which turns two vectors into a matrix. $|v\rangle\langle w|$ is the outer product[25] of $|v\rangle \in V$ and $|w\rangle \in W$, which is

$$|v\rangle\langle w| = \begin{bmatrix} v_0w_0 & v_0w_1 & \cdots & v_0w_{n-1}\\ v_1w_0 & v_1w_1 & \cdots & v_1w_{n-1}\\ \vdots & \vdots & \ddots & \vdots\\ v_{m-1}w_0 & v_{m-1}w_1 & \cdots & v_{m-1}w_{n-1}\end{bmatrix}$$

Effectively, we take two vector spaces V and W, with basis $\{|v_0\rangle, ..., |v_{m-1}\rangle\}$ and $\{|w_0\rangle, ..., |w_{n-1}\rangle\}$ respectively, and the tensor product of V and W, denoted $V \otimes W$, is the vector space with the basis

$$\{|v_0\rangle\langle w_0|, |v_0\rangle\langle w_1|, ..., |v_{m-1}\rangle\langle w_{n-1}|\}$$

So, now we've got some questions to answer.

What is the dimension of this new vector space? It's gonna be nm, as there are m vectors in the basis of V, and n in the basis of W.

What is this a vector space of? This is a vector space of matrices, which probably seems pretty scary, but it's not too bad, remember a vector space is just a collection of things satisfying $\alpha u + \beta v \in V$ for all $\alpha, \beta \in \mathbb{C}$ and $u, v \in V$, essentially[26].

So now we've got a vector space of matrices, what happens if I did the tensor product between two vector spaces of matrices? Well, lets say we have two vectors spaces, V and W, with bases made up of matrices, $\{V_0, ..., V_{m-1}\}$ and $\{W_0, ..., W_{n-1}\}$ respectively. In order to define the basis of $V \otimes W$ this time, we need something more than the outer product, we use something called the **Kronecker Product**.

[24] $\langle u|A^\dagger|A|v\rangle = \left\langle \langle u|A^\dagger \middle| A|v\rangle \right\rangle$, i.e the inner product of the bra $\left\langle \langle u|A^\dagger \right|$ and the ket $\left| A|v\rangle \right\rangle$. We just remove the outer angle brackets for shorthand.

[25] This is using Dirac's notation. In old notation it's $vw^\dagger = v\overline{w}^T$, and often we'd write this as $v \otimes w$, but this gets confusing when we introduce tensor and Kronecker product (which also use $\otimes$).

[26] Just don't look at https://en.wikipedia.org/wiki/Vector_space#Definition

For A, an $m \times n$ matrix, and B, a $p \times q$ matrix, the Kronecker product of A and B is

$$A \otimes B = \begin{bmatrix} A_{1,1}B & \cdots & A_{1,n-1}B \\ \vdots & \ddots & \vdots \\ A_{m-1,1}B & \cdots & A_{1,n-1}B \end{bmatrix}$$

So $A \otimes B$ is a $mp \times nq$ matrix.

So, now we're able to define our basis for the tensor product of two vector spaces of matrices. Using V and W with bases made up of matrices $\{V_0, ..., V_{m-1}\}$ and $\{W_0, ..., W_{n-1}\}$ as before, the basis for $V \otimes W$ is

$$\{V_0 \otimes W_0, V_0 \otimes W_1, ..., V_{m-1} \otimes W_{n-1}\rangle\}$$

So what does this vector space look like? Well the basis is made up of matrices of size $nm \times nm$, so it's simply a vector space of $nm \times nm$ matrices.

So we have outer products on vectors, Kronecker products on matrices, and tensor products on vector spaces? Yeah, basically, but the term tensor product gets thrown around a lot, and sometimes you'll see the Kronecker product referred to as the tensor product, but don't worry about it[27]. The symbol is the same for Kronecker and tensor product, which is annoying[28]. It's okay as long as you remember the Kronecker product is done on matrices, and the tensor product is done on vector spaces, and make sure you just keep track of what's going on.

So hopefully, you've understood everything written above, and now you can happily delve into the real[29] world of quantum computation, and soon when you see all the crazy hadamard gates, qubits, superpositions, and joint states, you'll be able to say 'ooooh, I understand all this, a hadamard gate is a useful example of a unitary operator, a qubit is just an unit vector in $\mathbb{C}^2$, superpositions are the same as linear combinations, and that joint state is just the Kronecker product of a bunch of qubits!'.

10 Quantum Computers and Linear Algebra

To wrap things up, we'll have a brief discussion of why vectors and vector spaces are just so great. By framing classical computation in this language as well, we can get a better grasp of how, exactly, "going quantum" changes things.

In classical computing, we usually have some input of n bits, and then we do something that transforms those n bits into a different arrangement of n bits. Think of these bits as vectors in a 2-dimensional vector space with basis $\{|0\rangle, |1\rangle\}$. Say we're dealing with two bits, we can say this is a 4-dimensional vector space with basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ (which is actually the tensor product of the original basis $\{|0\rangle, |1\rangle\}$ space with itself). Continuing up to n bits, we get a space of dimension 2^n, whose basis consists of all possible n-bit strings (i.e., the computational basis).

What do our classical algorithms do? They do any linear transformation that takes a basis vector (an n-bit string) to another basis vector (a different n-bit string). Quantum computers let us do any *unitary* transformation, so we can start with a basis vector, and go to any other unit vector that is a linear combination of basis vectors[30]. This looks like a clear win for quantum, but we should be careful: the restriction to unitary transformations is quite significant, and prevents quantum computers from performing some remarkably simple computations (consider the algorithm that ignores its input and always outputs $|0^n\rangle$); thus, we shouldn't think of quantum computers as beating classical computers outright.

[27] Tensor products are actually so much more complicated than explained here, to quote the wikipedia article on tensor products 'This article may be too technical for most readers to understand'.

[28] If only Dirac was here.

[29] complex

[30] This is called a 'superposition'.

11 Exercises

1. Show that $|+\rangle = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$ and $|-\rangle = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$ form a basis for $\mathbb{C}^2$.
2. Show that the matrices $H = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$, $X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and $Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ are unitary.
3. Solve exercises 3 and 4 of Ryan O'Donnel's second problem sheet (referred to as 2.3 and 2.4, from here on).
4. Solve exercise 1.3 of Ronald de Wolf's lecture notes.
5. * Solve Ryan's exercises 1.5, 2.2 and 2.6.
6. * Prove Theorem 4.