LONG HW 1: **Working with Vector Spaces**

This is the 'long' assignment assigned every two weeks. You should aim to complete it within two weeks because you'll have new assignments by then, but the formal due date is two weeks + two days. Your explainer video assignment will be to present one of these problems.

Part of the challenge of the 'long' homework may be to figure out exactly what is being asked. Be sure to use the beginning of lecture (or office hours) to ask early when you are confused. Sometimes we are *intentionally* using jargon that you may not be familiar with in order to encourage you to ask and/or look things up.[1] A good mantra that was once told to the professor when he was a student: *There are no stupid questions. Only stupid students who do not ask questions when they are confused.*

Background

In this problem set we use the abstract bra-ket notation. Perhaps you are used to writing vectors as $\vec{v}$ or $\mathbf{v}$. In our abstract notation, vectors are called 'kets' and are written $|v\rangle$. This is just notation, there is no math or science involved in this.

The usual basis vectors for $\mathbb{R}^N$ are $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, etc. Sometimes we write them as $\hat{\mathbf{e}}_1$, $\hat{\mathbf{e}}_2$, etc. In other words,

$$\hat{\mathbf{x}} = \hat{\mathbf{e}}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \end{pmatrix} \qquad\qquad \hat{\mathbf{y}} = \hat{\mathbf{e}}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \end{pmatrix} , \tag{0.1}$$

and so forth. The basis kets are typically written as $|1\rangle = \hat{\mathbf{x}} = \hat{\mathbf{e}}_1$ and similarly for $|2\rangle$, $|3\rangle$, and so forth. The notation has its pros and cons—you will soon appreciate them in this class—but since it is new, we should get used to it. A vector $|v\rangle$ is expanded in basis kets as

$$|v\rangle = v^1|1\rangle + v^2|2\rangle + \cdots . \tag{0.2}$$

The components v^i can be arranged as a column of numbers if you want to recover the expression for vectors.

1 Matrices as Vectors

Here's something that sounds totally wacky. We can define a vector space of matrices. Here is a basis for 2×2 real matrices:

$$|1\rangle = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad |2\rangle = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad |3\rangle = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \qquad |4\rangle = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} . \tag{1.1}$$

---

[1]Sometimes this is just because the professor is absent minded about what students know and do not yet know.

1.1 Writing a matrix as a vector

What are the vector components v^i of the following matrix?

$$|v\rangle = \begin{pmatrix} 1 & 4 \\ 2 & 7 \end{pmatrix} . \tag{1.2}$$

COMMENT: we say *matrix*, but really this object is a 2×2 array of numbers that should be understood as a funny-looking vector space.

1.2 Writing a vector as a matrix

Write the following vector, $|w\rangle$ as a 2×2 matrix using the basis (1.1):

$$|w\rangle = \pi|1\rangle + e|2\rangle - 2|3\rangle + 2.4|4\rangle . \tag{1.3}$$

1.3 Changing Basis

Here's a new basis for the vector space of 2×2 real matrices:

$$|1'\rangle = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad |2'\rangle = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad |3'\rangle = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad |4'\rangle = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} . \tag{1.4}$$

What are the components $(v')^i$ and $(w')^i$ of the vectors $|v\rangle$ and $|w\rangle$ in this basis? These are defined by, for example:

$$|v\rangle = (v')^1|1'\rangle + (v')^2|2'\rangle + (v')^3|3'\rangle + (v')^4|4'\rangle . \tag{1.5}$$

Make sure to explain how you do this; your explanation is more valuable than the final answer.

1.4 Linear Independence

Consider the **subspace** of the vector space of 2×2 matrices that is **spanned** by only the first three basis vectors in (1.1). This is the three-dimensional vector space whose vectors can all be written in the form

$$|v\rangle = v^1|1\rangle + v^2|2\rangle + v^3|3\rangle . \tag{1.6}$$

Let's call this subspace V_3. We can analogously define a subspace V_3' formed out of linear combinations of first three basis vectors in (1.4). These are composed of vectors of the form

$$|v\rangle = (v')^1|1'\rangle + (v')^2|2'\rangle + (v')^3|3'\rangle . \tag{1.7}$$

Please write down one example of a vector from the 'full' vector space of 2×2 real matrices that live in each four regions of the following Venn diagram:
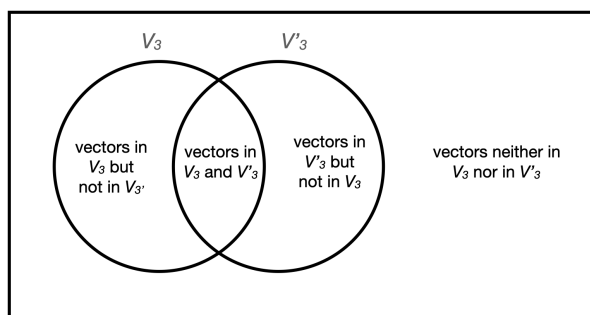
2 The Gram–Schmidt Procedure

In microbiology, a Gram-positive bacteria are those that have a peptidoglycan cell wall layer[2]. This problem has nothing to do with that. In linear algebra, the Gram–Schmidt procedure is a way to take a set of linearly independent vectors and produce an **orthonormal** basis out of them.

In the first parts of this problem, please be sure to draw your pictures carefully and use them to explain the process.

2.1 Two linearly independent vectors

Consider the two-dimensional vector space $\mathbb{R}^2$. Suppose we have the linearly independent vectors,[3]

$$|v\rangle = \begin{pmatrix} 4 \\ 1 \end{pmatrix} \qquad\qquad |w\rangle = \begin{pmatrix} 1 \\ -2 \end{pmatrix} . \tag{2.1}$$

Here we have written the vectors in the 'standard basis' so that the first component is the length in the x-direction and the second component is the length in the y-direction. Please draw these two vectors.

2.2 Creating the first orthonormal basis vector

Define $|\hat{v}\rangle$ to be the unit vector that is parallel to $|v\rangle$. What are the components of $|\hat{v}\rangle$ in the standard basis? These are the components $\hat{v}^i$ such that

$$|\hat{v}\rangle = \hat{v}^1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \hat{v}^2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} . \tag{2.2}$$

HINT 1: what is the condition for a vector to be a unit vector? (Please do not say that the vector is large.[4]) What does this condition imply on the components $\hat{v}^i$?

HINT 2: What does it mean for two vectors to be parallel? What does this condition imply on the components $\hat{v}^i$?

Please draw the vector $|\hat{v}\rangle$ along with the vectors $|v\rangle$ and $|w\rangle$. We will use $\hat{v}$ as one of our new basis vectors. Hooray.

2.3 Orthogonal and non-orthogonal pieces

Our next goal is to find $|\hat{w}_\perp\rangle$, a second basis vector that is (1) orthogonal to $|\hat{v}\rangle$ and (2) it is a unit vector. Let us do the first part. The goal is to separate $|w\rangle$ into two pieces:

$$|w\rangle = |w_\parallel\rangle + |w_\perp\rangle , \tag{2.3}$$

where $|w_\parallel\rangle$ is parallel to $|v\rangle$ and $|w_\perp\rangle$ is perpendicular to v. Find the components $w_\parallel^i$ and $w_\perp^i$ of these vectors.

HINT: The parallel component is related to the dot product of $|w\rangle$ and $|\hat{v}\rangle$, which we may write as $\langle w, \hat{v}\rangle \equiv \mathbf{w} \cdot \hat{\mathbf{v}}$. Figure out this relation and write $|w_\parallel\rangle$ with respect to $\langle w, \hat{v}\rangle$ and $|\hat{v}\rangle$. Then use (2.3) to figure out $|w_\perp\rangle$.

---

[2]https://en.wikipedia.org/wiki/Gram-positive_bacteria
[3]For familiarity we're using column notation. You could also write $|v\rangle = 4|\hat{x}\rangle + 1|\hat{y}\rangle$. They mean the same thing.
[4]https://knowyourmeme.com/memes/absolute-unit

2.4 Normalize

We now know that $|w_\perp\rangle$ is perpendicular to $|\hat{v}\rangle$. It is not yet normalized. Follow the steps in Part 2.2 to define a normalized unit vector $|\hat{w}_\perp\rangle$ that is parallel to $|w_\perp\rangle$. Draw $|\hat{w}_\perp\rangle$, $|\hat{v}\rangle$, $|v\rangle$, and $|w\rangle$ together. Confirm from the drawing that we have successfully derived an orthonormal basis: $|\hat{v}\rangle$ and $|\hat{w}_\perp\rangle$.

2.5 What if we went the other way around?

Would you end up with the same set of basis vectors if you started the Gram–Schmidt process with $|w\rangle$ first? Explain why or why not; your explanation must include a picture illustrating the result.

2.6 More dimensions

Mo' dimensions, mo' problems[5]? How would the Gram–Schmidt process change if the vector space is three-dimensional? You end up repeating the steps for each additional. Please explain what happens in Part 2.3 when you have to separate the third vector into perpendicular and parallel parts.

COMMENTARY: For the second vector, you just needed parts perpendicular and parallel to the first vector. For the third vector, you now have a choice between projections onto the first and the second vector.

2.7 Orthonormal matrix space

Recall the vector space of 2×2 real matrices and the unusual basis in (1.4). Define a dot product (or inner product or metric... these are all the same thing) between vectors in this space as follows:

$$\langle v, w \rangle = \mathbf{v} \cdot \mathbf{w} = \mathrm{Tr}(\mathbf{vw}) . \tag{2.4}$$

In other words: let $|v\rangle$ and $|w\rangle$ be vectors that are 2×2 matrices. The inner product is the trace of the product of these matrices. For example:

$$\left\langle \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} 2 & 1 \\ 0 & -1 \end{pmatrix} \right\rangle = \mathrm{Tr}\left[\begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 0 & -1 \end{pmatrix} \right] = \mathrm{Tr}\begin{pmatrix} 2 & -2 \\ 4 & 1 \end{pmatrix} = 3 . \tag{2.5}$$

Take a moment to appreciate why this could be a good definition for an inner product. Confirm that $\langle v, w \rangle = \langle w, v \rangle$. Also confirm that $\langle v + w, t \rangle = \langle v, t \rangle + \langle w, t \rangle$.

Apply the Gram–Schmidt procedure to the basis in (1.4) to create an orthonormal basis.

HINT: Doing a four-dimensional Gram–Schmidt is kind of annoying. However, the basis in (1.4) naturally splits into two pairs that are mutually orthogonal. Start by showing that the diagonal

---

[5]The Notorious B.I.G. taught us *I don't know what they want from me / It's like the more money we come across / The more problems we see.* Often in physics the more dimensions there are the more subtle the system becomes. There is at least one exception in quantum field theory where physicists often go to $d + \varepsilon$ dimensions... this is like saying we go from 4 dimensions to 4.0001 dimensions. Crazy, right? This is called dimensional regularization and it is used as a trick to tame some unruly integrals. By the way, 4 dimensions = 3 spatial dimensions and 1 time dimension. There are some cases where we work with spacetimes with two time dimensions, but physical space is always some subspace with only one time dimension. A famous example is Anti de-Sitter space, which is at the core of what string theorists call the hologrpahic principle and is the kind of thing illustrated in some of M.C. Escher's most famous prints.

matrices and the off-diagonal matrices are orthogonal to each other with respect to the inner product (2.4).

3 Rotations

3.1 Rotations and inverse rotations in Euclidean space

In $\mathbb{R}^2$, one can perform a clockwise rotation by angle θ on a vector by multiplying the vector by the rotation matrix:

$$R(\theta) = \begin{pmatrix} R^1{}_1 & R^1{}_2 \\ R^2{}_1 & R^2{}_2 \end{pmatrix} \qquad R^1{}_1 = R^1{}_1 = \cos\theta \qquad R^1{}_2 = -R^2{}_1 = \sin\theta . \qquad (3.1)$$

In column vector notation, write out the rotation by angle θ of a vector $\mathbf{v}$ with components

$$\mathbf{v} = \begin{pmatrix} v^1 \\ v^2 \end{pmatrix} . \qquad (3.2)$$

3.2 Passive transformations

There's a subtle point that always pops up with rotations. Are we rotating the *vector* (as we did above) or are we rotating the *coordinate system* (basis vectors)? Let $\hat{\mathbf{e}}_1$ and $\hat{\mathbf{e}}_2$ be the standard unit vectors in the x and y-directions,

$$\hat{\mathbf{e}}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \hat{\mathbf{e}}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} . \qquad (3.3)$$

Define a new pair of basis vectors by rotating these basis vectors:

$$\hat{\mathbf{e}}'_1 = R(\theta)\hat{\mathbf{e}}_1 \qquad \hat{\mathbf{e}}'_2 = R(\theta)\hat{\mathbf{e}}_2 . \qquad (3.4)$$

It should be obvious[6] that the new basis is orthonormal. Write out new basis vectors as column vectors.

3.3 Active versus passive

Rotating the vector $\mathbf{v} \to R(\theta)\mathbf{v}$ is called an **active transformation**. Rotating the basis vectors $\hat{\mathbf{e}}_i \to R(\theta)\hat{\mathbf{e}}_i$ is called a **passive transformation**. Show that an active transformation on $\mathbf{v}$ by angle θ is the same as a passive transformation on $\mathbf{v}$ by an angle $-\theta$.

HINT: use $\sin(-\theta) = -\sin\theta$.

3.4 Inverse Rotation

What are the components of $R^{-1}(\theta)$, the inverse transformation of the rotation by angle θ? Show that this satisfies $R^{-1}(\theta)R(\theta) = R(\theta)R^{-1}(\theta) = 1$, the unit matrix.

---

[6]In this class, 'obvious' means: if it is not immediately clear to you, ask about this right away because there is a certain way of seeing things that you want to be able to appreciate. We will always use 'obvious' this way, so don't be intimidated if something 'obvious' is not clear. It only means that you should ask about how to make this idea clear.

3.5 Exponentiation of a matrix

Recall that the exponential of a number x is defined by its Taylor expansion,

$$e^x = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \cdots . \tag{3.5}$$

One can define the exponentiation of a matrix A in the same way:

$$e^A = 1 + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots , \tag{3.6}$$

where here A^n is understood to mean matrix multiplication. Show the following relation:

$$e^{i\theta T} = R(\theta) \qquad\qquad T = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} . \tag{3.7}$$

Note that θ is just some real number. [**Flip**: Correction 4/14: the form of T was originally off by a factor of i. Or maybe it is was off by $-i$. It may still be off by a factor of -1, but that boils down to a convention on the sign of θ.]

HINT: work out the first few terms explicitly. You should notice a pattern between the even powers and the odd powers. Remember what the Taylor expansions for sine and cosine look like.

COMMENT: The matrix T is special, it is called the **generator** of rotations. The theory of continuous symmetries (Lie groups) is built on the idea of generators for symmetry transformations. This idea finds a killer application in quantum mechanics.

A Extra Credit: Complex numbers as vector spaces

It turns out that complex numbers are a two-dimensional vector space. The complex numbers $\mathbb{C}$ differs from $\mathbb{R}^2$ because it has an added rule called a **complex structure**. This is basically the rule that $i^2 = -1$. The rotation matrix from Part 3.5 comes in handy, as well as the vector space of matrices in Problem 1. This problem shows the nice connection between complex numbers and rotations.[7]

A.1 Complex Arithmetic

A complex number is usually written $z = a + ib$. It has a real component a and an imaginary component b, where both a and b are real numbers. Another way to write this is as a vector $|z\rangle = a|1\rangle + b|i\rangle$, where the basis elements are

$$|1\rangle = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad\qquad |i\rangle = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} . \tag{A.1}$$

Show that the rules of complex arithmetic hold. If $z' = c + id$, show that $zz' = (ac - bd) + i(ad + bc)$ holds. That is, show that:

$$\left[a \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right] \left[ca \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + d \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right] = \left[(ac - bd) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + (ad + bc) \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right] , \tag{A.2}$$

so that $|z\rangle|z'\rangle = (ac - bd)|1\rangle + (ad + bc)|i\rangle$.

---

[7]The reason why these two are connected lies in the following cryptic statement: both are representations of the Abelian group, $U(1)$.

A.2   Complex Norm

Use the 'trace metric' from (2.4) to show that

$$\langle z, z \rangle = a^2 + b^2 \ . \tag{A.3}$$

In other words, you get the usual complex norm squared. Conveniently, this is also the Euclidean $\mathbb{R}^2$ norm squared.

A.3   Quaternions

Many physicists often wonder about the generalization of complex numbers to higher complex dimensions. The first step in this generalization is something called **quaternions**. These are complex numbers with *three* complex directions, i, j and k such that

$$i^2 = j^2 = k^2 = -1 \ . \tag{A.4}$$

It turns out that in order for this to all hang together, $ij \neq ji$ and similarly with k. The easiest way to get a feel for quaternions is to consider a space spanned by a particular set of 2×2 complex matrices, where here 'complex' refers to ordinary complex numbers. The basis vectors are:

$$|1\rangle = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad |i\rangle = \frac{i}{2}\sigma_x \qquad |j\rangle = \frac{i}{2}\sigma_y \qquad |k\rangle = \frac{i}{2}\sigma_z \ . \tag{A.5}$$

In physics the σ_i matrices are the **Pauli matrices**,

$$\sigma_x \equiv \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad \sigma_y \equiv \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} \qquad \sigma_z \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \ , \tag{A.6}$$

they are kind of a big deal.

Show that (A.4) holds with these matrices. Observe that the Pauli matrices do not *commute*: $\sigma_x \sigma_y \neq \sigma_y \sigma_x$. This is why the i, j, k complex directions do not commute under quaternionic multiplication. Work out the **commutator** of the Pauli matrices:

$$[\sigma_a, \sigma_b] \equiv \sigma_a \sigma_b - \sigma_b \sigma_a =? \tag{A.7}$$

Where a and b are in $\{x, y, z\}$.

COMMENT: The quaternions are the simplest mathematical representation of quantum mechanical spin. This lies at the heart of qubits in quantum computing.

COMMENT: Just like T was a generator for rotations, the Pauli matrices are generators of a generalization of rotations. The generalization is called the special unitary group of rank 2, or $SU(2)$ for short. This is the underlying symmetry group that governs the weak nuclear force in elementary particle physics.