This problem set is organized around two major questions—projections and least squares—which exemplify core ideas of this class, and which I want to use as motivation to develop concepts and tools. Two comments:

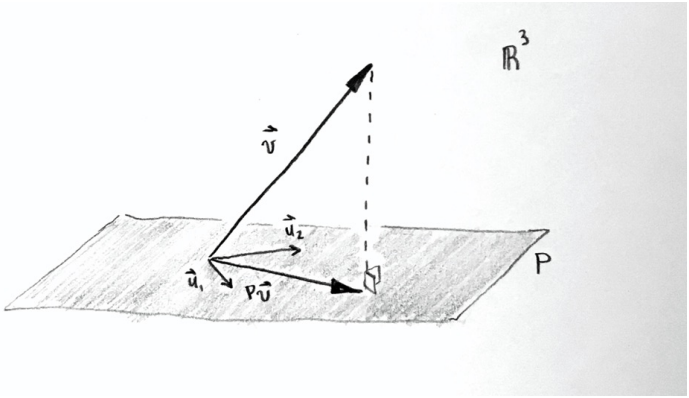
1) At least for now, I want you to try to work through these in your own way. You can look up linear algebra and trigonometry basics that you don't remember, but please don't look up solutions to these problems. My intention here is twofold: by trying these problems on your own, a) you can start to feel out what the difficulties are, and you can more directly understand what is central and/or clever about the slicker solutions we will develop later; and b) you can start to see what tools we are missing or would be nice to have, which will motivate their development next.

2) This may feel uncomfortable, because you may feel like you are needing things that you haven't learned, or don't remember, or which have not been carefully defined. Please don't worry! Right now, we are figuring out the shape of what you do and don't know already, and charting a path forward. And I am intending to start a careful logical development with the next problem set. I did not want to *start* with the logical development, because I didn't want to do that without some indication of where it was going. As we go forward, we will work forwards from the logical development, and backwards from filling in holes in our attempted solutions of these problems, and meet in the middle (and then go forward beyond these problems).

PROJECTIONS

Projections are a huge part of linear algebra, both the theory and the applications. They will recur in almost everything we do in this class. That is why I want to start with this problem.

BIG PROBLEM: PROJECTION Suppose that we are given a plane V' in $V = \mathbb{R}^3$. The plane has been specified by giving two vectors $\vec{u}_1$ and $\vec{u}_2$ in S; the plane is the set of all linear combinations $t_1\vec{u}_1 + t_2\vec{u}_2$, $t_1, t_2 \in \mathbb{R}$, of these two vectors.[1] We are also given a vector $\vec{v}$ in $\mathbb{R}^3$. The problem is to find the perpendicular projection $P\vec{v}$ of the vector $\vec{v}$ down onto the plane V'. (Part of the question will be for you to be more precise about what that means.)



What I would like you to try to do (you can go out of order, and partial progress is fine!):

[1] I am not assuming that the vectors $\vec{u}_1$ and $\vec{u}_2$ are perpendicular, or unit length, or anything. I am assuming that they are both non-zero, and they are not collinear, so that we actually get a plane from their linear combinations.

- Try to figure out a method of solving this, any way that you can.
- Try it on some particular numbers, and check that you get the right answer (question: how do you know if the answer is right??).
- Write out your solution in a general form. That is, you may have some steps in your method that you follow; do those steps for the vectors $\vec{v} = (v_1, v_2, v_3)$, $\vec{u}_1 = (u_1 1, u_1 2, u_1 3)$, and $\vec{u}_2 = (u_2 1, u_2 2, u_2 3)$, so that your answer is a formula for $P\vec{v}$ in terms of these given variables. (I'm assuming that the vectors $\vec{v}$, $\vec{u}_1$ and $\vec{u}_2$ are given by their components in some standard coordinate system.)
- Try to make your solution neater looking! It's quite probable you have something pretty messy looking: is it possible to either a) arrange it so it looks nicer (using matrices maybe?), and/or b) do a different method which will give the answer more easily?
- I would like to generalize this result to higher dimensions: given a vector $\vec{v}$ in $V = \mathbb{R}^n$, and a k-dimensional subspace V' which is the span of k vectors $\vec{u}_1, \ldots, \vec{u}_k$, I would like to find the perpendicular projection $P\vec{v}$ of the vector $\vec{v}$ onto the subspace V'. Depending on how you solved the problem in $\mathbb{R}^3$, you may find that your solution generalizes immediately; or, more likely, you may find that you need to add or change something to make things work in higher dimensions. You might need to go back and do one step a totally different way.

■

Notes:

- A reminder: please don't look up the answer to this problem right now. I'd like you to try to figure it out on your own, even if it means you are somewhat stuck. If there are basic linear algebra or trigonometry things you need to look up, go ahead (e. g. "dot product has a cos or a sin?"; "cos is which side over which?"); but please put off looking up fancier things for now (e. g. "what's the change of basis formula?").
- There are many ways to do this problem! Some end up working better or worse; but all the methods have something interesting about them, and end up being useful for something later on, so no time spent on this will be wasted!
- Partial progress is fine! This is a hard problem. Don't be put off if you feel stuck. Try to write down what you DO know, what you CAN figure out. Try numerical examples. Try replacing the problem with something easier.
- In some problems to follow (below, and possibly on future problem sets), I will make suggestions on how to approach this problem in different ways. You can do those problems to help you move forward, but please also keep working on this Big Problem in parallel.
- Similarly, as we go on to other things, please keep working on this problem in parallel. You can send me progress on it, or talk to me about it, any time!

Problems connected to the projection problem.

1. A good strategy for solving a difficult problem is often to try to find an easier problem that you can solve, which you can use. Some of you had this idea, and said, "let's try to solve the projection problem in 2D rather than 3D". This is an important thing, which you may or may not recall from linear algebra, so I'd

like everyone to work through this question (and once you solve it, I recommend trying to commit the solution to memory, at least so you can easily rederive it).

Suppose that we are given a vector $\vec{v}$ in $V = \mathbb{R}^2$, and we are given a line V' which is the span (set of all multiples) of a non-zero vector $\vec{u}$. I want to find the perpendicular projection $P\vec{v}$ of $\vec{v}$ onto the line V'. There are different ways to approach this problem, so feel free to follow different steps than the ones I give. (In particular, I will be showing you a different way to do this problem later that will be in some respects better.)

(a) (Solving the problem in 2D)
Use trigonometry to find the length of the projection, $\|P\vec{v}\|$, in terms of $\vec{v}$, $\vec{u}$, and the angle θ between them.

(b) Use the formula for the dot product,[2]

$$\boxed{\vec{v} \cdot \vec{w} = \|\vec{v}\| \, \|vcw\| \cos\theta}$$

to rewrite your formula for $\|P\vec{v}\|$ in terms of dot product only (no θ). (Since the dot product can also be computed by

$$\boxed{\vec{v} \cdot \vec{w} = v_1 w_1 + v_2 w_2}$$

when $\vec{v}$ and $\vec{w}$ are given in components, this means that with this new formula, we don't need to know the angle θ in order to compute $\|P\vec{v}\|$.)

(c) Now, use your knowledge of $\|P\vec{v}\|$, plus the fact that $P\vec{v}$ is a multiple of $\vec{u}$, to find a formula for the vector $P\vec{v}$. Simplify as much as you can. (When simplifying, it may be helpful to keep in mind the rule

$$\boxed{\vec{v} \cdot \vec{v} = \|\vec{v}\|^2,}$$

which you can prove assuming the rules above.)

(d) Once you have your formula, test it!

(i) Try it on some numbers: make up two vectors $\vec{v}$ and $\vec{u}$ in components, and find $P\vec{v}$ using your formula. How can you be sure that your answer is correct?!? Recall that (from the formula above),

$$\boxed{\vec{v} \perp \vec{w} \quad \text{if and only if} \vec{v} \cdot \vec{w} = 0}$$

(ii) Can you test the formula more conceptually? That is, can you tell that your formula for $P\vec{v}$ actually satisfies the conditions you wanted for the perpendicular projection of $\vec{v}$ on $\vec{u}$?

2. (Finding perpendiculars)
The most common strategy that people tried for the projection problem in 3D involved finding a vector that is perpendicular to the plane V'. The most common way people did that was by the "cross product". In this problem, I want to do two things: remind/tell you about the cross product, and suggest alternate ways of finding perpendiculars.

(a) Given two vectors $\vec{v}$ and $\vec{w}$ in $\mathbb{R}^3$, the *cross product* of $\vec{v}$ and $\vec{w}$ is the vector[3]

$$\boxed{\vec{v} \times \vec{w} = (v_2 w_3 - v_3 w_2, v_3 w_1 - v_1 w_3, v_1 w_2 - v_2 w_1)}$$

[2]This formula is a little mysterious. We will prove it, and also show where it comes from, a little later in the class!

[3]We will see later where this crazy-looking formula comes from! We will also see why the cross product and dot product in particular are singled out as ways of combining vectors in three dimensions. (The dot product works in any dimension, but the cross product only in three (and sort of in two).

It has the property that

- $\vec{v} \times \vec{w}$ is perpendicular to both $\vec{v}$ and $\vec{w}$;
- the direction of $\vec{v} \times \vec{w}$ is determined by the "right hand rule": put your fingers along $\vec{v}$, curl them towards $\vec{w}$, and your (unless you are unnaturally flexible), your thumb is pointing in the direction of $\vec{v} \times \vec{w}$;
- the length of $\vec{v} \times \vec{w}$ is given by

$$\boxed{\|\vec{v} \times \vec{w}\| = \|\vec{v}\| \, \|\vec{w}\| \sin\theta}$$

We will see later how to prove all of these properties. Right now, I just wanted to remind you of / tell you about the cross product, so that you can use it in the projection problem if you like. I will ask you to prove one property now:

(i) Based on the defining formula for cross products in coordinates given above, prove that $\vec{v} \times \vec{w}$ is perpendicular to both $\vec{v}$ and $\vec{w}$.

We'll prove the rest later!

(b) One disadvantage of the cross product is that it only works in three dimensions.[4] However, we don't need a cross product to find vectors perpendicular to given ones. Recall the fact that I wrote above (which we'll prove later):

$$\boxed{\vec{v} \perp \vec{w} \quad \text{if and only if} \vec{v} \cdot \vec{w} = 0}$$

(i) Suppose that we are in the situation of the original projection problem: we are in $V = \mathbb{R}^3$, we are given two vectors $\vec{u}_1$ and $\vec{u}_2$, and we want to cook up a vector $\vec{u}_3$ which is perpendicular to both given vectors $\vec{u}_1$ and $\vec{u}_2$.

(A) Make up two numerical vectors $\vec{u}_1$ and $\vec{u}_2$ in components in $\mathbb{R}^3$. Write down the conditions for the unknown vector $\vec{u}_3$ to be perpendicular to $\vec{u}_1$ and $\vec{u}_2$. How can you solve this system? (This should start to be feeling familiar, yes?) How many solutions would you expect? Does this make sense geometrically? Solve it for your numbers, and check that your answer is correct.

(B) Do the previous problem, but leave $\vec{u}_1$ and $\vec{u}_2$ unknown. Write out your answer in terms of the unknown components of $\vec{u}_1$ and $\vec{u}_2$. Check your answer.

(C) Based on your prior knowledge of linear algebra, how would this work out in higher dimensions? If we tried to find all vectors perpendicular to $\vec{u}_1$ and $\vec{u}_2$ in $\mathbb{R}^4$, for example, what would be the dimension of the solution set? How many linearly independent vectors will we be able to find that are perpendicular to both $\vec{u}_1$ and $\vec{u}_2$? If we try to find all vectors perpendicular to $\vec{u}_1, \ldots, \vec{u}_k$ in $V = \mathbb{R}^n$, how many linearly independent solutions can we find? What concepts/theorems from linear algebra are relevant here (feel free to check your old notes!!).

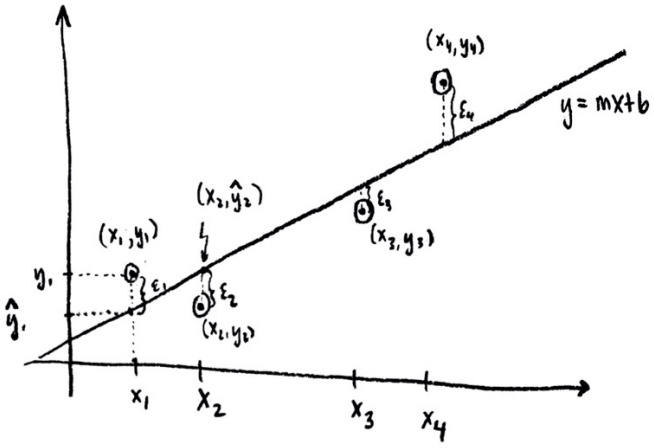
(D) Can you use these ideas in your solution to the projection problem??

More on the projection problem to come!

[4] There is a generalization of cross products to higher dimensions, called *differential forms* and *wedge products*, that we will say later on in this course. They are very interesting, and play a huge role in physics and geometry. But they turn out not to be that useful for the projection problem at hand, so I'm leaving them till later.

Least Squares

BIG PROBLEM: LEAST SQUARES Suppose that we have some data points, $(x_1, y_1), \ldots, (x_n, y_n)$. We want to find a "best fit line" $y = mx + b$ to try to fit this data.[5] That is, we want to choose the unknown parameters m and b such that, in some sense, the "total errors" of the theoretical model $y = mx + b$ in predicting the data $(x_1, y_1), \ldots, (x_n, y_n)$ are as small as possible.



For each value of x in our data set, x_i, our model is predicting a value for y by

$$\hat{y}_i = mx_i + b.$$

The error in that prediction is how far off our model is from the actual data value:

$$\epsilon_i = \hat{y}_i - y_i$$

It will turn out to be useful to think of the error as an n-dimensional vector,

$$\vec{\epsilon} = (\epsilon_1, \ldots, \epsilon_n).$$

There are different ways of summing up the "total error"; a sensible way would be

$$\text{"total" error} = |\epsilon_1| + \cdots + |\epsilon_n|.$$

However, for reasons we will get into later, it will be mathematically nicer to take as the "total" error the *root mean square* of the errors:

$$\text{RMS}(\vec{\epsilon}) = \sqrt{(\epsilon_1)^2 + \cdots + (\epsilon_n)^2}.$$

The "least squares" problem is, given the data points $(x_1, y_1), \ldots, (x_n, y_n)$, to find the values of m and b such that $\text{RMS}(\vec{\epsilon})$ is as small as possible.

■

This will be an ongoing problem, and I don't want you to try very hard to solve it on your own yet, though it might be worthwhile to think about a little.[6] With the problems below, and in the coming lectures and problem sets, I want to explain

[5] We will see later that it doesn't have to be a line! If we look for a best fit parabola, or exponential curve, or sine curve, etc., we can do some very similar things.

[6] It is possible to solve this as a multivariable calculus problem! You are minimizing a function, $\text{RMS}(\vec{\epsilon})$, which depends on two variables. Since I'm not assuming you have done MV calculus, I won't spend much time on this in class; but if you have taken MV, I recommend trying it this way! We'll see later how that solution connects to the linear algebra one.

the connection of this problem to the projection problem (which I started to do very briefly at the end of the Friday February 19 class).

Problems leading up to the least squares problem.

3. It will be useful to think of the problem in a linear algebra way as follows: suppose that we are given the data points $(x_1, y_1), \ldots, (x_n, y_n)$. Think of the x values of the data points, $x_1, \ldots, x_n$, as being fixed for the remainder of the discussion. Think of the y values of the data points as an n-dimensional data vector, $\vec{y} = (y_1, \ldots, y_n)$, sitting inside an n-dimensional vector space $V = \mathbb{R}^n$. Since we have fixed all the $x_1, \ldots, x_n$, then a given value of m and b in the model $y = mx + b$ will automatically define the n predicted values $\hat{y}_i = mx_i + b$, $i = 1, \ldots, n$, which I also want to think of as an n-dimensional vector $\vec{\hat{y}} = (\hat{y}_1, \ldots, \hat{y}_n)$.

   Let V' be the set of all the possible n-dimensional vector $\vec{\hat{y}} = (\hat{y}_1, \ldots, \hat{y}_n)$, where m and b take on any possible values (each value of m and b produce a different $\vec{\hat{y}}$). I want to characterize some things about the set V'. (It might help to imagine this at first by assuming we have $n = 3$ data points, so that you can picture V' as some subset of $V = \mathbb{R}^3$.)
   (a) Is the zero vector in the set V'? Why or why not?
   (b) Suppose some vector $\vec{\hat{y}}$ is in the set V'. Will it be true that $2\vec{\hat{y}}$ will also be in the set V'? (That is, there must have been some value of m and b that produced the predicted data vector $\vec{\hat{y}}$. Is there then a value of m and b that will produce the predicted data vector $2\vec{\hat{y}}$??) Why or why not? How about $k\vec{\hat{y}}$, where k is any real number?
   (c) Suppose some two vectors $\vec{\hat{y}}_1$ and $\vec{\hat{y}}_2$ are in the set V'. Is it going to be always true that their sum is in V'? Why or why not?
   (d) Imagine that $n = 3$. If you think about what you just determined in the previous questions: what does that tell you about the geometrical shape of V'?

4. (Minimizing distance) This is actually a problem about projections. But it is particularly relevant to the least squares problem, as we will soon see. It is also a tough problem, so consider it a semi-optional challenge problem for now. It is very instructive to try it, but don't get too frustrated if you get stuck.

   Here's the problem: suppose that $V = \mathbb{R}^3$, that V' is a plane in V, that $\vec{v}$ is a vector in V, and that $P\vec{v}$ is the perpendicular projection of $\vec{v}$ onto V'. Suppose that $\vec{w}$ is any other vector in V'. Prove that, (identifying vectors with their endpoints), $P\vec{v}$ is the point on V' at a minimum distance from the point $\vec{v}$. That is, for any $\vec{w}$ in V',

$$\|\vec{w} - \vec{v}\| \geq \|P\vec{v} - \vec{v}\| .$$