Final Project **Due on the Day you take the Final**

In this project you will learn about one of the most important (and beautiful) applications of Linear Algebra - an application of "dot product" to the "least squares" fitting.

This project has 3 parts.

1. You MUST complete Part 1 of this project.
2. If you only complete Part 1, then this project will count for 15% of your final exam score.
3. If you complete Parts 1 and 2, then this project will count for 25% of your final exam score, unless your final exam score is higher, in which case the project will count for 10% (this will help boost your class grade)
4. If you complete Parts 1, 2, and 3 then this project will count for 30% of your final exam score, unless your final exam score is higher, in which case the project will count for 10% (this will help boost your class grade)

Example: Suppose you do Part 1 and get a 95% and on the in-class final you earn a 70%. Then your final exam score will be

$$(.15)(.95) + (.85)(.70) = .7375$$

If you completed Part 2 as well (and earned a 95%) then your 70% in-class score would be increased to

$$(.20)(.95) + (.80)(.70) = .7625$$

You may ask me for help with this project, and you may also work with your peers. However, I ask that everyone write up their solutions individually. Some of the questions are quite challenging, but that's ok - you can do it!

What do I have to do?

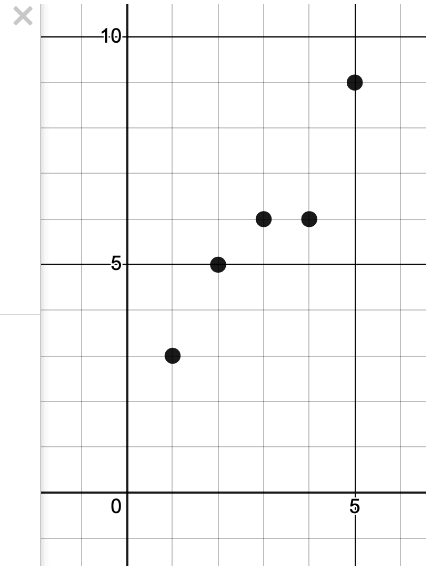
- Below, I coach you through the computational steps required to find a "least squares" linear approximation to a data set.
- Please read through the instructions and work everything out.
- There may be parts that are confusing at first, but hopefully by the end, you will have a solid grasp on what we're doing.
- In particular you should hopefully have a grasp on how by choosing a point "closest" to a plane, we are minimizing something.
- After you are done working through everything on scratch paper, please write up a report summarizing what you did, showing your work and explaining your steps.
- For instance, at the beginning you will probably refer to a set of points - make sure your report includes a picture, or else the reader won't know what you're talking about.
- This doesn't need to be long, but it does need to be well-written and presented clearly. I think a full report might be 1-2 pages long (plus some pictures I will ask you to include.)
- It's OK if you just work through the steps on the bulleted points below and explain the connections.
- This might just feel like a long homework assignment with some calculuations. I recommend doing your calculations very carefully. If you're stuck please ask me for help!
- Parts of this project use formulas that we won't get a chance to prove in our class, but if you want to explore why they are true, that's what Parts 2 and 3 are for.
- Note that Part 1 is a written report. Parts 2 and 3 require you present your solution either in my office or in a video.

Part 1:

Consider the following problem. You are trying to find the equation of the line that passes through the data points in the picture. Now you're probably thinking - there's no line that passes through all of these points.

- Justify this statement. Pay careful attention to your writing and your explanation. Don't overthink this, but make sure your explanation is clear and convincing.

x_1	y_1
1	3
2	5
3	6
4	6
5	9



However, you have perhaps heard of something called a **best fit line** that roughly approximates these data points. There are lots of different ways we could do this, but the way we will focus on in this problem is something called the **least-squares fit.** Go on Desmos and after watching this short video: `https://help.desmos.com/hc/en-us/articles/4406972958733-Regressions`

- With the data set of our problem, use Desmos to get the best fit line, like they did in the video. You'll want to use the ~ symbol as opposed to the = sign. You should get the line $y = 1.3x + 1.9$. How many of your original data points does this line go through? Draw (or screenshot) a LARGE picture of this line and you data points. You are going to do some calculations. On a separate LARGE picture (I'm thinking full page for each picture) I want you to draw a line that goes through the leftmost and rightmost points. You should get a different line. Find the equation of this line. On your pictures, for each line, I want you to draw in vertical lines, measuring how far each line is from the data points. For instance, for the line $1.3x + 1.9$ at the point where $x = 2$ your first line gives the point $(2, 4.5)$ and the data point is higher at $(2, 5)$. Draw a vertical line going UP from this line to the data point and carefully label this distance. (In this case the distance is 0.5). When you are done you should have a total of 10 distances, 5 for each graph. Some of the distances might be zero.

> By the way, on this assignment I will be paying careful attention to your writeup and your presentation. This assignment doesn't have to be long. Apart from the pictures, I think everything might be possible to fit on a single page or two. But I will be looking that your have things presented in a clear and readable way.

- Next for each line I want you to calculate the "sum of squares" of these distances. Of course you should use a calculator, but I ask you to show your work so I can verify you are doing the calculation correctly.

> The "least squares line" is the line that minimizes this "sum of squares" of the distance between the line and the data points. This approximation is very useful and we are going to explore how to find it. The one that the computer found is the least-squares line. Among all the possible lines you could every try. This one will have the smallest "sum of squares of the errors". You should have seen that the second line you drew had a high sum of squares. That line was not as good of an approximation.

What is our goal anyways? Ideally, we would want to find m and b so that $y = mx + b$ were true for all of our data points. Let's practice with a different data set to see what happens with an **exact solution** is possible.

Suppose we have the points $(1,5), (2,7), (5,13), (10,23)$. If you want to solve for m and b, you could plug each point (x, y) into $y = mx + b$ and you would get, for instance if we plugged in $x = 2, y = 7$ we'd get

$$7 = 2m + b.$$

Normally, we put the constant on the right hand side, so we'll write this as

$$2m + b = 7.$$

- Do the same for the other 3 points and you should get a system of 4 equations in 2 unknowns. Write this system clearly and neatly, in the form of

$$?m + ?b = ?$$

  like above. Using an augmented matrix, solve this system to see what m and b must be. (This will be consistent, because there is an **exact** solution.) Write the system you just solved as a matrix equation. You should have a matrix times the vector $\begin{bmatrix} m \\ b \end{bmatrix}$ equal to a vector. Using your solution, write an equation using the columns of your matrix to demonstrate that the right hand vector is in the span of the columns of your matrix.

- Returning to our original data set (the one on the front page), write down what the system of equations would be. Write it as a matrix equation:

$$A\mathbf{blah} = \mathbf{foo}$$

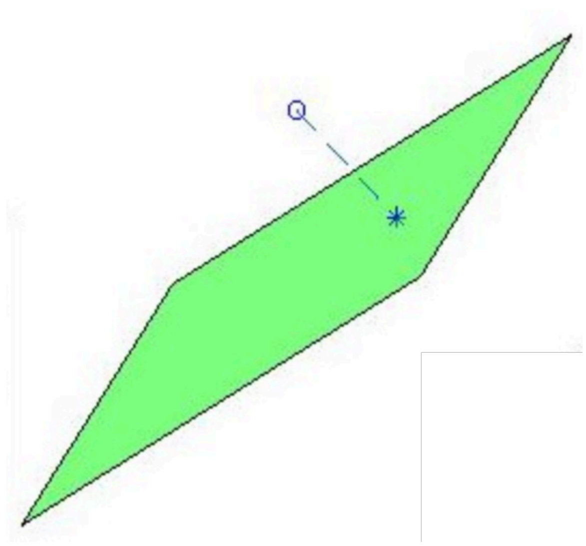
  In this setup what is **blah**? **foo**? One of these will be a vector of numbers from your data set, the other will be a vector with the **variables** you are hoping to solve for. You don't have to write blah or foo on your project, but make sure you clearly indicate what your equation is.

- Ok, so in this problem you have already justified above that there is **NOT** a line that goes through all of the data points. So this system is NOT consistent. This means that **foo** is NOT in the span of the columns of A.

  Make sure you understand this - what comes next will be mighty confusing if you doing understand this.

- Hmm, now what is the span of the columns of A? We had a name for this space. And we had a way to describe it by giving a basis. In a paragraph, explain what the space is, write down the basis and give a description using the words "subspace, dimension". For instance you might say "this is a 5 dimensional subspace of $\mathbb{R}^7$ with basis." Hint: Your basis should consist of vectors whose entries are all whole numbers and you shouldn't have to do any calculations to find it. If you find yourself doing calculations to find the basis, STOP and check with me.

- Draw a picture of this situation. You are working in a high dimensional space, but that's ok - your subspace is only 2 dimensional, so that means it is a [line/plane/hallowe'en dog]. You can draw a picture of this, and you can also draw the vector **foo** that is NOT in that space. As you draw your picture, incorporate a point in your subspace that is CLOSEST to the point **foo**. Draw a little line connecting **foo** to your plane. I'm looking for something like this, but with labels for your vector and your subspace.



- Recall that there are lots of different bases that we can find and use for a subspace. The one you wrote down above is a perfectly fine basis but the vectors in it are not **perpendicular**. At the moment let's fix some notation. Your basis should have two vectors in it. One of those vectors should have different entries. Let's call that one $\mathbf{v}_1$ and the other one $\mathbf{v}_2$. Our goal in the next step is to find a **basis for your subspace with perpendicular vectors.** A synonym for perpendicular is **orthogonal**. The formula for doing that is as follows:

To turn a basis $\{v_1, v_2\}$ into an orthogonal basis, you will

– Keep $\mathbf{v_1}$ the same. Replace $\mathbf{v_2}$ with $\mathbf{v_2} - \dfrac{\mathbf{v_1} \cdot \mathbf{v_2}}{\mathbf{v_1} \cdot \mathbf{v_1}} \mathbf{v_1}$ where $\cdot$ denotes the dot product of two vectors. Note that that fraction is a scalar, and you are multiplying it by $\mathbf{v}_1$.

By using this process, please find an orthogonal basis for your subspace. You answer should have two vectors, one whose entries are all integers and one whose entries involve fractions with denominator 11. Double check that your vectors are perpendicular by verifying that their dot product is zero. Let's call your new vectors $\mathbf{w}_1, \mathbf{w}_2$.

- If only there were a formula for how to find the closest point to a subspace. Oh wait, there is, provided you have an orthogonal basis.. Phew, good thing we did that!

  If you want to project a vector $\mathbf{v}$ onto a subspace that has **orthogonal basis** $\{\mathbf{w}_1, \mathbf{w}_2\}$ then you will get:

  $$\text{Proj}\,\mathbf{v} = \frac{\mathbf{v} \cdot \mathbf{w}_1}{\mathbf{w}_1 \cdot \mathbf{w}_1} \mathbf{w}_1 + \frac{\mathbf{v} \cdot \mathbf{w}_2}{\mathbf{w}_2 \cdot \mathbf{w}_2} \mathbf{w}_2.$$

  Use this formula to figure out what vector is $\star$ in the picture above. Your vector should have numbers that have no more than 1 place after the decimal. This formula only works if you have an orthogonal basis, that's why we needed the earlier steps.

- Ok now, you are almost done. Your vector, call it $\mathbf{p}$ is IN the subspace. Go back and label that closest point in your picture $\mathbf{p}$. Now, although your original equation

  $$A\mathbf{blah} = \mathbf{foo}$$

  had NO solution since **foo** was not in the span of the columns of A. You have now found the "closest point to foo that is actually in the span of columns of A." That's what $\mathbf{p}$ is. This means that the equation

  $$A\mathbf{blah} = \mathbf{p}.$$

  Will have a solution, (and that the solution will give a the line $y = mx + b$ that is the "closest" line, in terms of least squares.) Solve this system, and if everything went well, you should see the numbers 1.3 and 1.9 popping out.

- To summarize what you have just done: You have taken a system that was inconsistent, because the vector **foo** was not in $\text{Col}(A)$. But then you found the "closest point in $\text{Col}(A)$ to **foo**" and then you solved the **consistent** system with this new point. Along the way, you used some formulas to find these closest points.

- I think this is so cool, because to solve a "minimize the error" problem, we didn't need any calculus, just a way to "find the closest point" which we did with these projection formulas. In **Part 2** you will see how to prove these formulas.

- However, the story gets even better. In a surprising plot twist, there is an even quicker way to find the least-squares line.

  To find the least-squares solution to $A\mathbf{x} = \mathbf{b}$, you just solve the system

  $$A^T A\mathbf{x} = A^T\mathbf{foo}.$$

  Go ahead and solve this system (showing all of your work, and row-reducing the matrices by hand) and verify that again you get the same $y = mx + b$ as you got before.

Story Time: This method with $A^T A$ is called the normal equation method and is how computers can quickly find least-squares approximations. This method is great because you don't need to do any projections or deal with orthogonal bases directly. In other words all this is done behind the scenes. However, to really understand what is happening, and why this 'normal equation' works relies on the geometric insight we developed throughout the class. In Parts 2 and 3 you will go deeper into the theory. If you think this stuff is fun, I really encourage you to give it a try and see what happens!

Part 2:

In this part you will make a presentation. You can either record a video of your explanation on a whiteboard, with slides, or with writing on paper, OR you can come by my office sometime and present your explanation. Your presentation should include an explanation of both Tasks 1 and 2.

How do you find an orthogonal basis?

Let's start with two vectors in the plane, say $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$. These vectors are not orthogonal. But what if you wanted to build an orthogonal basis for $\mathbb{R}^2$ using your vectors.

Task 1: Find an orthogonal basis for $\mathbb{R}^2$ of the form

$$\{\mathbf{v}, \mathbf{w} + a\mathbf{v}\}.$$

In other words your job is to find what the scalar a needs to be in order to make these vectors orthogonal. Before consulting any books, I want you to think about how your would do this. Can you draw a picture? Would you solve some equations? In your presentation, I want to know how your first thought about this. Draw a picture of $\mathbf{v}$ and think about what vectors are perpendicular to it. How could you find one of those that is of the form $\mathbf{w} + a\mathbf{v}$?

STOP

Ok, after you've thought about the problem above, here's an idea that can help.

Dot product to the rescue:

You want the dot product of the two vectors in your basis to be equal to zero. Write this down as an equation involving $a, \mathbf{v}, \mathbf{w}$. (Don't use any numbers yet) Then solve your equation for a and see what you get. You should get a formula similar to something in Part 1 of this project. Verify that this is the same as the method you used above.

Pause: before moving on to task 2, make sure you understand what i'm getting at. When you dot perpendicular vectors you always get 0, make sure you see how this gets rid of terms in the example above. This will help a lot with Task 2.

Task 2: Now suppose that we have a basis for a two dimensional plane H in $\mathbb{R}^3$, say the vectors are $\mathbf{v_1}, \mathbf{v_2}$ and we have a random vector $\mathbf{v}$ in $\mathbb{R}^3$ that we want to project as a vector in H plus a perpendicular vector $\mathbf{p}$. This means we want something like:

$$\mathbf{v} = a_1\mathbf{v_1} + a_2\mathbf{v_2} + \mathbf{p}.$$

Using the same idea of "dot product" to the rescue. Can you find a way to solve for a_1, a_2 and $\mathbf{p}$ in terms of $\mathbf{v}, \mathbf{v_1}, \mathbf{v_2}$?

Some hints: What happens if you take the dot product of both sides with $\mathbf{v}_1$? Hmm, what would happen if you dotted $\mathbf{p}$ with $\mathbf{v}_1$? Will a picture help?

Part 3:

There are actually ways to define dot products in spaces other $\mathbb{R}^n$, even in infinite dimensional spaces! For instance, if our space is V the space of all functions $f : \mathbb{R} \to \mathbb{R}$, then one way to define the "dot product" of two functions is as follows (I'm going to use $\star$ to make it clear this is a different operation)

Definition: If f, g are two vectors in V then

$$f \star g = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)g(x)dx$$

As a warmup, use Desmos to calculate: $x^2 \star \cos(x)$, $\sin(x) \star \cos(2x)$, and $\sin(3x) \star \sin(3x)$.

You should have found that two of the vectors in the previous problem were **orthogonal**. In fact, ALL of the following vectors are ALL orthogonal to each other AND they all dot with themselves to give 1.

$$\mathcal{B} = \{\frac{1}{\sqrt{2}},\ \cos x,\ \sin x,\ \cos 2x,\ \sin 2x,\ \cos 3x,\ \sin 3x,\ \ldots,\}$$

Whoa: That's some really nice property - no matter which two vectors you take, either $v \star w = 0$ if we take two different vectors and if we dot a vector with itself we get $v \star v = 1$.

Task 1: Verify (in a video or in my office) some of the claims in the above Whoa statement. You don't have to prove everything in general, but I want to get a sense that you understand the profundity of this claim, that there are different cases, that sometimes the integral is zero, that sometimes it is 1. That there's some sine and cosine stuff going on and what's going on with the $\frac{1}{\sqrt{2}}$?

So the set above is actually extremely useful. We can use this space to find really good approximations for functions. For example, imagine you have the function:

$$f(x) = 2\cos(x) + 3\sin(x)$$

this can be written a combination of elements of our set $\mathcal{B}$. You just take 2 of the $\cos x$ and 3 of the $\sin x$. But what if you had a mystery function $M(x)$ that you knew was built out of these functions. (Say you have a wave pattern in a recording) below:

and you want to know how this is built out of the elements of B. Well guess what...

.......The Dot Product Comes to the Rescue

Being the curious scientist that you are, you use a compute (computers can numerically calculate integrals given a table of values, after all) a BUNCH of integrals with your Mystery function M and the elements of $\mathcal{B}$

$M \star \frac{1}{\sqrt{2}} = 0$

$M \star \sin x = 1$

$M \star \cos x = 1$

$M \star \sin(2x) = 0$

$M \star \cos(2x) = 3$

$M \star \sin(3x) = 0$

$M \star \cos(3x) = 0$

$M \star \sin(4x) = 0$

$M \star \cos(4x) = 0$

$M \star \sin(5x) = -4$

$M \star \cos(5x) = 0$

And this information leads you to see: "Hmm, my function seems like it's orthogonal to a lot of these functions, but it's actually got some nontrivial involvement with some of these things... I wonder if... "

And sure enough

$$M(x) = \sin(x) + 3\cos(2x) - 4\sin(5x) \ + \cos(x).$$

Do you see where those coefficients come from?

Task 2: Take the function $M(x)$ above, and knowing what its formula is, go ahead and take the $\star$ product

$$M(x) \star \sin(5x)$$

and do it WITHOUT using any integrals. Use the fact that $\star$ is distributive, so you can just calculate $\star$ with each term individually. AND use the fact that you know that all the properties of $\mathcal{B}$ from the Whoa box above.

After showing how you calculated that product, consider the following. If you had a function that you knew was of the form

$$P(x) = a\frac{1}{\sqrt{2}} + b\sin(x) + c\cos(x) + d\sin(2x) + e\cos(2x)$$

what would you do to try and figure out what the coefficients of a, b, c, d, e are ? Can you relate this to what you did in Part 2?

That's it - you're done.... BUT if you want to do something really fun...

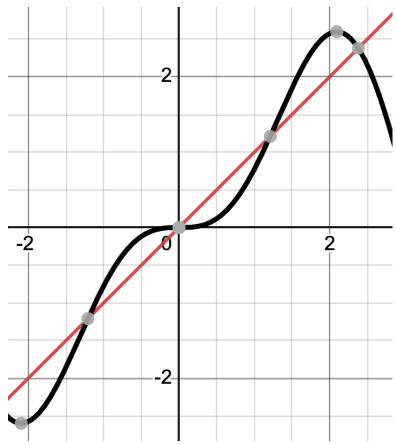
Not every function can be built from the functions in $\mathcal{B}$. For instance, all of those functions are periodic with period 2π. However, if we allow **infinite sums** and **only look in the range from $-\pi$ to π** we can actually get pretty far. For instance, if we HOPED that the function x were a combination of these things we might try:

$$x = a\frac{1}{\sqrt{2}} + b\sin(x) + c\cos(x) + d\sin(2x) + e\cos(2x)$$

and using $\star$ product we could estimate what happens with the coefficients and get

$$x = 0\frac{1}{\sqrt{2}} + 2\sin(x) + 0\cos(x) - 1\sin(2x) + 0\cos(2x)$$

And this doesn't look too bad!



If we wanted to try more terms, say going all the way to sin/cos of $(5x)$ we could get

$$x = 2\sin x - \sin 2x + \frac{2}{3}\sin 3x - \frac{2}{4}\sin 4x + \frac{2}{5}\sin 5x$$

You may have noticed a pattern - that the coefficient of $\sin nx$ is the same as $\pm\frac{2}{n}$. (and that all the coefficients of $\cos(nx)$ are zero. You could verify this by calculating e.g. $\frac{1}{\pi}\int_{-\pi}^{\pi} x\sin(nx)dx$.

This is true! And so if the pattern continues, we get that

$$x = \frac{2}{1}\sin x - \frac{2}{2}\sin 2x + \frac{2}{3}\sin 3x - \frac{2}{4}\sin 4x + \frac{2}{5}\sin 5x - \frac{2}{6}\sin 6x + \frac{2}{7}\sin 7x + \dots$$

Now if you $\star$ both sides with themselves, you'll get:

$$x \star x = (\frac{2}{1}\sin x - \frac{2}{2}\sin 2x + \frac{2}{3}\sin 3x - \frac{2}{4}\sin 4x + \frac{2}{5}\sin 5x \cdots) \star (\frac{2}{1}\sin x - \frac{2}{2}\sin 2x + \frac{2}{3}\sin 3x - \frac{2}{4}\sin 4x + \frac{2}{5}\sin 5x \cdots)$$

But oh my gosh, what happens on the right hand side - all the $\star$ products where we have different terms will be zero (by the Whoa property). So that the simplified right hand side is:

$$x \star x = (\frac{2}{1})^2\sin(x)\star\sin(x) + (\frac{2}{2})^2\sin(2x)\star\sin(2x) + (\frac{2}{3})^2\sin(3x)\star\sin(3x) + (\frac{2}{4})^2\sin(4x)\star\sin(4x) + \cdots$$

But the Whoa property also says that $\sin(nx) \star \sin(nx) = 1$ so we have that

$$x \star x = 4 + \frac{4}{2^2} + \frac{4}{3^2} + \frac{4}{4^2} + \cdots$$

$$x \star x = 4(1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots)$$

But wait a minute, let's calculate

$$x \star x = \frac{1}{\pi} \int_{-\pi}^{\pi} x \cdot x dx = \frac{1}{\pi}\frac{1}{3}(\pi^3 - (-\pi^3)) = \frac{2\pi^2}{3}$$

So we get that

$$\frac{2\pi^2}{3} = 4(1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots)$$

which means:

$$1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \cdots = \frac{\pi^2}{6}.$$

(This is one of the most beautiful facts I've ever seen in all of math - and you can understand it just using the powers of linear algebra! Dot products, perpendicularity, and a little optimism. It's been a great semester working with everyone. I hope you enjoyed this project.)