Problem 1:

Given the linear model $f(u) = b + ku$, and a series of data points (u_i, f_i), the error for any i^{th} data point may be quantified in the following way:

$$e_i = f_i - f(u_i) = f_i - (ku_i + b).$$

In this equation e_i represents the difference between a data value and our linear model evaluated at the same input value u_i.

If the aim is to construct a line of best fit for all (u_i, f_i), in the case that there are multiple data points, we must develop a multivariate function, $E(b, k)$, such that different choices of b and k will yield different linear models when applied to the data set, allowing us to optimize this function so that our model fits our data "well" based on some metric (to follow).

Before stating this equation it must be specified more what is meant by "best-fit". In an intuitive sense, a model will more accurately represent experimental data, when e_i small. Since our data set is not a single point in $\mathbb{R}^2$, however, in order for our linear model to fit the data, "well" we must seek for the lowest possible error "on-average", or with respect to all the data. We will apply a method known as *Least-Squares Regression* to account for all data points along with some other benefits.

$$E(b, k) = \sum_{i=1}^{5} (e_i)^2 = \sum_{i=1}^{5} [f_i - (ku_i + b)]^2.$$

Which for the given data set consisting of five e_i, becomes:

$$\begin{aligned}
&E(b, k) = [0.100 - (k(.041) + b)]^2 + \\
&[0.197 - (k(.086) + b)]^2 + \\
&[0.298 - (k(.128) + b)]^2 + \\
&[0.395 - (k(.173) + b)]^2 + \\
&[0.492 - (k(.218) + b)]^2
\end{aligned}$$

The above equation, while somewhat simple, is useful and appropriate for the following reasons.

Fantastic combination of verbal, visual, and symbolic representations: I see you making connections between various parts of your work! This is a great start to the type of evidence for learning that I am looking for! I appreciate how you are really exploring the full set-up for this problem and demonstrating your knowledge about how this problem relates to one of the major themes in this class: that of creating multivariable functions!

Your descriptions of the optimization problem below is great. This gets into a much larger topic of optimization and least-squares… The absolute value minimization problem is a famous one that fits into a topic known as convex optimization (https://web.stanford.edu/~boyd/cvxbook/bv_cvxbook.pdf). When we return to the polynomial multivariable minimization problem, there are some fun connections between this idea and Taylor series expansions of multivariable functions…

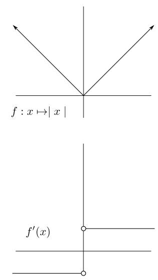


Figure 1: Ah, There's the Rub! $f'(x) = \frac{x}{|x|}$

1.We consider the sum of the errors since, in general, a smaller sum represents a line that is "on-average" as close as possible to the provided data. That is, it allows us to consider the model's closeness to all the data rather than a single point.

2.Each term e_i of the sum should be recorded as a "distance" or magnitude, that is they should be non-zero, since data that might lie below our linear model geometrically is not any less "close" to our model than points lying above our model. If we allow for a sum with $e_i < 0$ we run the risk of having an inaccurate function to optimize(we would in effect be calculating the difference between two $e_i, e_j, i \neq j$ and in general would not be accurately measuring accumulative error sizes). In some way we need to ensure error values are not signed.

We might at first consider the absolute value operator as our means by which to obtain non-negative e_i. Using e_i^2 is better for two reasons(there may be more). Firstly, since this is an optimization problem which implies using differentiation on our function, we want our function to have a derivative that is everywhere defined. Any function of the form $f : x \mapsto | ax + c |$ will not will not have a continuous derivative as shown in figure above.

Secondly, $\sum_{i=1}^{n} e_i^2$ will produce a polynomial-multivariate function. A polynomial function will be continuously differentiable,and furthermore, will be easy to work with, since any polynomial is composed with only multiplication and addition which are easy to differentiate, and computationally efficient to differentiate in the case that n is large.

Problem 2:

$\vec{v}$ is a vector with an initial point $A(1, 0, -1)$ and a terminal point $B(-5, 6, 4)$. Thus,

$$\vec{v} = \begin{bmatrix} -5-1 \\ 6-0 \\ -4+1 \end{bmatrix} = \begin{bmatrix} -6 \\ 6 \\ -3 \end{bmatrix}.$$

To produce a new vector that is in the same direction (colinear) to $\vec{v}$, that has an opposite orientation, and a length of 6, we will.

1. Find $\hat{v}$ (colinear unit vector to $\vec{v}$)
2. Scale this unit vector $\hat{v}$ by 6, and orient it such that it is opposite to $\vec{v}$ by

Missing a negative sign here. That error propagates through the rest of your solution… I bet this is a LaTeX typesetting error

scaling it by a negative factor.

Step 1:

$$\hat{v} = \frac{\vec{v}}{\|\vec{v}\|} = \frac{1}{((-6)^2 + 6^2 + (-3)^2)^{\frac{1}{2}}} \cdot \begin{bmatrix} -6 \\ 6 \\ 3 \end{bmatrix} \frac{1}{\sqrt{81}} \cdot \begin{bmatrix} -6 \\ 6 \\ 3 \end{bmatrix} = \frac{1}{9} \cdot \begin{bmatrix} -6 \\ 6 \\ 3 \end{bmatrix} = \begin{bmatrix} -\frac{6}{9} \\ \frac{6}{9} \\ \frac{3}{9} \end{bmatrix} = \begin{bmatrix} -\frac{2}{3} \\ \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}.$$

Step 2:

$$-6 \cdot \hat{v} = -6 \cdot \begin{bmatrix} -\frac{2}{3} \\ \frac{2}{3} \\ \frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{12}{3} \\ -\frac{12}{3} \\ \frac{-6}{3} \end{bmatrix} = \begin{bmatrix} 4 \\ -4 \\ -2 \end{bmatrix}$$

This is the vector we hoped to establish.

Remark 1. *This procedure relies on the true proposition $\|c\vec{x}\| = \mid c \mid \|\vec{x}\|, c \in \mathbb{R}$, which expresses a few very important properties for $\mathbb{R}^n$. For the above problem, it says that any vector can be stretched or shrunk to a desired length $\mid c \mid \|\vec{x}\|$ simply by scaling the vector by that that factor. Conversely, it specifies that all that is guaranteed by this operation is that the vector's scaled norm is the same, and whether or not $\vec{x}$ is given a new orientation depends on whether $c > 0$ or $c < 0$. Finally, this equation reinforces that idea that a vectors components scale with length, which helps to explain why the formula for producing unit vectors works, $\frac{\vec{v}}{\|\vec{v}\|}$. This result follows from the definition of scalar-vector multiplication and the above equation.*

Once again: great work describing in words the ideas behind this math.

I would like to discuss metric and normed spaces with you further, but want some time to explore further before writing more here. I think understand how the above property can be can be proven or at least more deeply understood from the properties that a metric space must have when they are combined with linear spaces to form normed spaces(for a normed space it is desirable and intuitive to reality to combine linear properties with our metric, so we want linear scaling(above) and we also want translational invariance (think two beads connected by a string, translated in space). The two norm definition for $\mathbb{R}^n$ provides for these properties.

Did we address your questions in our last conversation? Check this out: https://people.math.sc.edu/sharpley/math555/Lectures/MetricSpaceIntro.html Notice that the basic inequalities in that book I shared with you show up in this (and so many other) expositions on normed linear spaces.

Problem 3:

[ht]

Let $\vec{p} = proj_{\vec{y}}(\vec{x})$, and assume that $\vec{r}$ denotes a vector orthogonal to $\vec{y}$, and equal to the difference between $\vec{x}$ and $\vec{p}$. Then the following may be stated:

1.Let $\vec{r}$ be both orthogonal to $\vec{y}$ and the vector difference between $\vec{p}$ and $\vec{x}$.

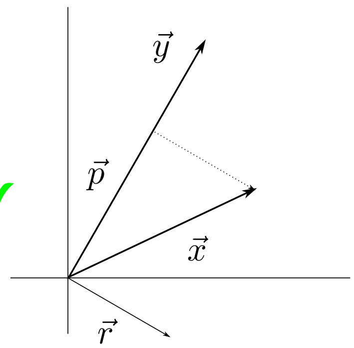
$proj_{\vec{y}}(\vec{x}) = \vec{x} - \vec{r} \leftrightarrow \vec{r} = \vec{x} - \vec{p}$

2.Since $\vec{r} \perp \vec{y}$,

$\vec{y} \cdot \vec{r} = 0$

3.$\vec{p}$ and $\vec{y}$ are colinear meaning that each one representable as scalar multiple of the other.

$\alpha\vec{y} = \vec{p}, \alpha \in \mathbb{R}$



Taking these equations above in conjunction:

$$
\begin{aligned}
0 &= \vec{y}\cdot\vec{r} = \vec{y}\cdot(\vec{x}-\vec{p}) && \text{By substitution}\\
&= \vec{y}\cdot\vec{x} - \vec{y}\cdot\vec{p} && \text{Right-Distributivity(linearity) of dot product}\\
&= \vec{y}\cdot\vec{x} - \vec{y}\cdot\alpha\vec{y} && \text{Substitution}\\
&= \vec{y}\cdot\vec{x} - \alpha(\vec{y}\cdot\vec{y}) && \text{Homogeneity of dot product}\\
&= \vec{y}\cdot\vec{x} - \alpha\|\vec{y}\|^2 && \text{Definition of two-norm}\\
&\leftrightarrow \alpha\|\vec{y}\|^2 = \vec{y}\cdot\vec{x}\\
\alpha &= \frac{\vec{y}\cdot\vec{x}}{\|\vec{y}\|^2}. && \text{Rearrange and solve for scalar } \alpha
\end{aligned}
$$

Then,

$$proj_{\vec{y}}(\vec{x}) = \vec{p} = \alpha\vec{y} = \frac{\vec{y}\cdot\vec{x}}{\|\vec{y}\|^2}\cdot\vec{y} = \frac{\vec{y}\cdot\vec{x}}{\|\vec{y}\|}\cdot\frac{\vec{y}}{\|\vec{y}\|}.$$

In this final equation, the farthest right-hand size may be interpreted and visualized as a unit vector co linear to $\vec{y}$, scaled by a real number equal to the size of the projection of $\vec{x}$ onto $\vec{y}$. This can be seen geometrically with the knowledge that $\frac{\vec{y}\cdot\vec{x}}{\|\vec{y}\|} = \|\vec{x}\|\cos\theta$ (how to scale the unit vector).

Problem 4:

$$a =< 1, 0, 1 >$$
$$b =< 0, 1, -1 >$$

Find $\vec{b} = \vec{p} + \vec{r} : p \| a \wedge r \perp a$.

$$p \| a \implies (c\vec{p} = \vec{a}, c \in \mathbb{R} \wedge \vec{a} \times \vec{p} = 0)$$
$$r \perp a \implies \vec{r} \cdot \vec{a} = 0$$

Using this information and a series of substitutions:

$$\vec{r} \cdot \vec{a} = 0 \leftrightarrow r_1 + r_3 = 0 \leftrightarrow r_1 = -r_3$$
$$b = ca + r =< c, 0, c > +r =< 0, 1, -1 >$$
$$\implies$$
$$0 = c + r_1 \leftrightarrow c = -r_1 = r_3$$
$$1 = r_2$$
$$-1 = c + r_3 \leftrightarrow -1 = r_3 + r_3 \leftrightarrow r_3 = -\frac{1}{2} \implies r_1 = \frac{1}{2}$$
$$\therefore r =< \frac{1}{2}, 1, -\frac{1}{2} >, p = -\frac{1}{2}a =< -\frac{1}{2}, 0, -\frac{1}{2} >$$

then,

$$\vec{p} + \vec{r} = \begin{bmatrix} -\frac{1}{2} \\ 0 \\ -\frac{1}{2} \end{bmatrix} + \begin{bmatrix} \frac{1}{2} \\ 1 \\ -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} = \vec{b}.$$

$$\vec{r} \cdot \vec{a} = \|\vec{r}\| \|\vec{a}\| \cos\theta$$
$$\vec{r} \cdot \vec{a} = (\frac{1}{4} + 1 + \frac{1}{4})^{\frac{1}{2}} \cdot (1 + 1)^{\frac{1}{2}} \cos\theta$$
$$= \sqrt{\frac{3}{2}}\sqrt{2}\cos\theta$$
$$\theta = \cos^{-1}(\frac{\frac{1}{2} + 0 - \frac{1}{2}}{\sqrt{\frac{3}{2}}\sqrt{2}}) = \cos^{-1}(0) = 90°$$

$\therefore \vec{r}$ is indeed orthogonal to $\vec{a}$.

There is a simpler way to do this that fundamentally is saying the same thing. I didn't think of it first for whatever reason so that is why it is second. If $\vec{p}$ is colinear to $\vec{a}$, and $\vec{r}$ is perpendicular to $\vec{a}$, then b can be given as the sum of $\vec{p}$ and $\vec{r}$ with p as the projection of b onto a and r as the residual vector orthogonal to the projection. Solving the problem this way looks as follows:

$$proj_{\vec{a}}(\vec{b}) + r = b$$

$$\frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\|^2} \cdot \vec{a} + r = b$$

$$\frac{-1}{2} \cdot < 1, 0, 1 > +r =< 0, 1, -1 >$$

$$r =< 0, 1, -1 > - < -\frac{1}{2}, 0. -\frac{1}{2} >=< \frac{1}{2}, 1, -\frac{1}{2} > .$$

✓ This is the same result as above.

Problem 5:

An equation representing the area of the parallelogram spanned by $\vec{x}$ and $\vec{y}$ in $\mathbb{R}^2$ may be produced by a geometric difference which is equal to $\mid det[\vec{x}, \vec{y}] \mid$ (Proof that is is magnitude here requires more work, provided later).

Four geometric shapes with simple formulas for their areas, may be specified by the components of $\vec{x}$ and $\vec{y}$.(Figure 2) They are:

- 2 pairs of triangles given by $\frac{x_1 y_1}{2}, \frac{x_2 y_2}{2}$
- A large rectangle given by $x_1 y_2$
- A smaller rectangle given by $x_2 y_1$

The area of the parallelogram may then be given by the difference:

$$[x_1 y_2 + \frac{x_1 y_1}{2} + \frac{x_2 y_2}{2}] - [x_2 y_1 + \frac{x_1 y_1}{2} + \frac{x_2 y_2}{2}] = x_1 y_2 - x_2 y_1 = det[\vec{x}, \vec{y}].$$ ✓

✓ Notice that the last two terms of the binomial on the left are identical to the last two terms of the right hand binomial, meaning that their difference is zero (they cancel), leaving only the terms representing a larger rectangle and a smaller one, as shown in the figure, and giving the area of the parallelogram in terms of vector components.

Problem 6:

Before considering how to represent the area of a parallelogram spanned by two vectors, we will state the formula for area for parallelograms (Figure 3).

As is show in the figure, the base of a parallelogram spanned by $\vec{x}, \vec{y}$ may be given by, $\|\vec{x}\|$, and the height by $\|y\| \sin \theta$ where θ is the angle between the two vectors. As figure 4 shows, it is arbitrary whether a vector is chosen to represent the base or height of the parallelogram.

I was hoping you'd show this in the easier case of $\mathbb{R}^2$… In this case, your argument holds in 3D. Not bad, but more general than the question that I was asking.

How did you make these diagrams in Figures 1, 2, and 3? Is this TikZ? I'd love to get access to your source code, if you're willing to share…

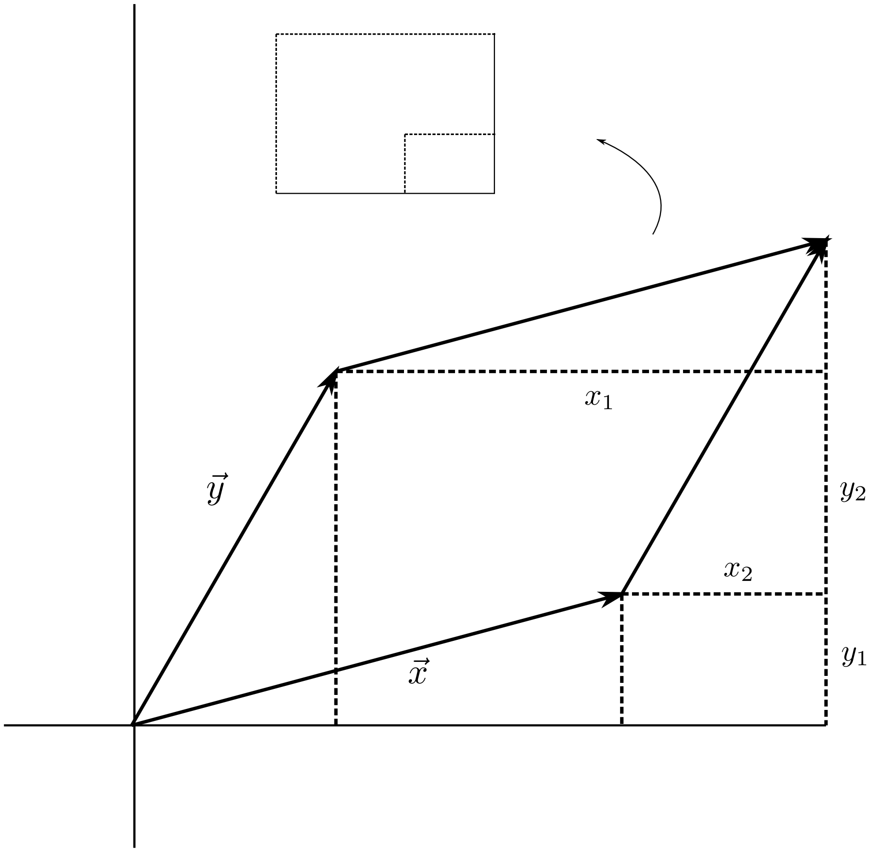


Figure 2: Parallelogram composed of simpler geometry

✓ However, this result made be extended to show that this formula is equivalent to the length of a vector given by the cross product of $\vec{x}$, $\vec{y}$. This result is somewhat abbreviated to avoid expanding two trinomial squares and 18 subsequent terms. The procedure follows analogous lines for showing that determinants in $\mathbb{R}^2$ correspond to areas(see end of quiz).

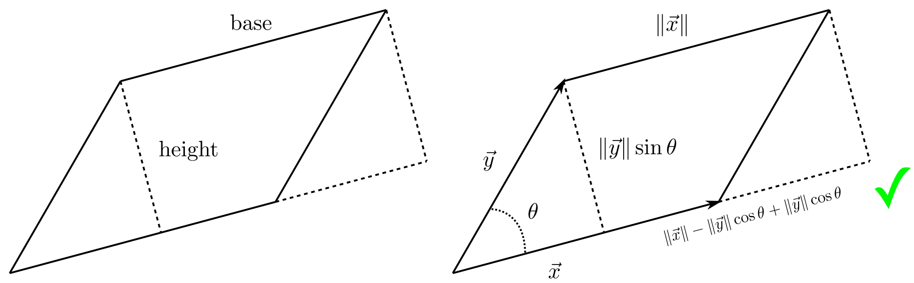


Figure 3: Analog of Parallelogram area for Vectors

$A_p = \|\vec{a}\|\|\vec{b}\|\sin\theta$

$Consider : \sin\theta = \sqrt{1-\cos^2\theta}$

$$\sqrt{1-\frac{(a_1b_1+a_2b_2+a_3b_3)^2}{(a_1^2+a_2^2+a_3^2)(b_1^2+b_2^2+b_3^2)}}$$

$$\sqrt{\frac{(a_1^2+a_2^2+a_3^2)(b_1^2+b_2^2+b_3^2)-(a_1b_1+a_2b_2+a_3b_3)^2}{(a_1^2+a_2^2+a_3^2)(b_1^2+b_2^2+b_3^2)}}$$

Considering the whole equation and moving $\|\vec{a}\|\|\vec{b}\|$ under the square root,denominators cancel yielding

$$\sqrt{(a_1^2+a_2^2+a_3^2)(b_1^2+b_2^2+b_3^2)-(a_1b_1+a_2b_2+a_3b_3)^2}$$

which yields a lovely 18 terms.But, 3 pairs will cancel $a_1^2b_1^2, a_2^2b_2^2, a_3^2b_3^2$
the remaining terms can be rearranged into 3 binomial squares

$$\sqrt{(a_1b_2-a_2b_1)^2+(a_1b_3-a_3b_1)^2+(a_2b_3-a_3b_2)^2} = \|\vec{a}\times\vec{b}\|$$

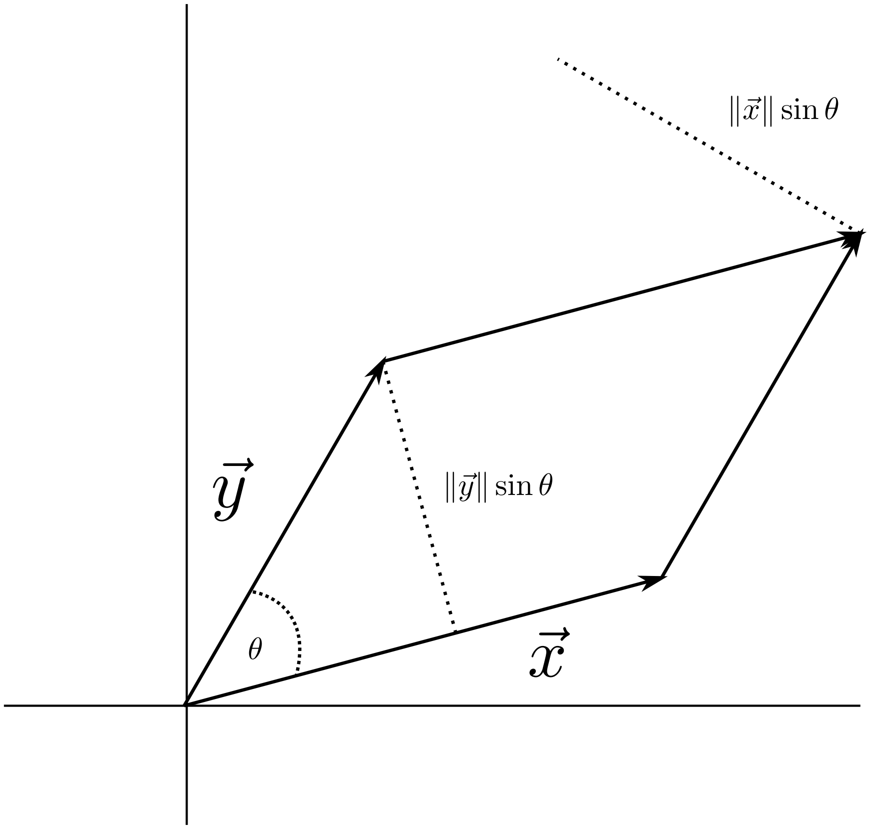


Figure 4: Geometric Importance of cross product

Problem 7:

I like to imagine the component form of the cross product as an example case of using relationships learned in simpler cases($\mathbb{R}^2$) and transferring them to a more complex space ($\mathbb{R}^3$). The cross product is a response to the following questions.

1. If the dot product provides a standard for orthogonality in $\mathbb{R}^2$ how might I consider vector orthogonality in $\mathbb{R}^3$, or ask a similar question about how "aligned" two vectors are?
2. How might I produce an orthogonal vector to a plane in $\mathbb{R}^3$ specified by two other 3-D vectors?
3. Algebraic equations for geometry in $\mathbb{R}^3$

Let's start with the first question. If we were starting in $\mathbb{R}^2$ with knowledge of linear combinations, dot products, and vector projects, then we understand that two vectors of various lengths when scaled, and added, span particular parallelograms. We also know, from both $\vec{x}\cdot\vec{y} = \|\vec{x}\|\|y\|\cos\theta$ and $\mid det(\vec{x},\vec{y}) \mid$

that parallelogram area is smaller when the vectors are closer (with respect to θ), that is when they are less orthogonal/more parallel.
Vectors in $\mathbb{R}^3$ still define a plane and θ still exists, although the plane may not be parallel to a coordinate plane. Therefore, we might think to talk about 3-D vector orthogonality by talking about the plane spanned by two vectors in $\mathbb{R}^3$. However, the component way to do this in $\mathbb{R}^2$ was for vectors in $\mathbb{R}^2$, so to use components in $\mathbb{R}^3$ we need something more.

But!, we can use our 2-D component form if we combine it with our knowledge of projections, to encode 2-D information in a 3-D structure, which will utilize all components in $\mathbb{R}^3$ using information from $\mathbb{R}^2$.This idea will be summarized for brevity.

- Consider two vectors $\vec{x}, \vec{y} \in \mathbb{R}^3$, and then consider their plane projections in xy-plane, yz-plane, xz-plane. These are mathematically given by taking the coordinates of the desired plane from $\vec{x}, \vec{y}$, with third coordinate set to 0.(The same result follows using projection formula from problem 3).

- Consider these projections as vectors in $\mathbb{R}^2$, taking those coordinates corresponding to the plane projected in, and excluding the unused dimensions.Furthermore, consider the parallelogram spanned by these vectors. Consider what happens to this 2-D parallelogram when our starting $\vec{x}, \vec{y} \in \mathbb{R}^3$ approach co linearity. As $\vec{x} \to \vec{y}$, each 2-D vector will become closer and the area of the parallelogram will shrink. Thus we have found a property in 2-D that is related to the orthogonality(or co linearity) of our 3-D vectors.

- . Do this procedure for each pair of coordinates for $\vec{x}, \vec{y} \in \mathbb{R}^3$. Since they are in 3-D space we have to consider the multiple ways in which they may not be co linear(For instance if $\vec{x}, \vec{y}$ only have different z component). We can only say definitively that they are co linear, when all computed 2-D determinants(areas) are zero. (This is a way to see why the length of the cross product measures area). Notice too that we have utilized all components of $\vec{x}, \vec{y}$.

When thinking about how to usefully encode these values in a 3-D structure, we consider the second question. Can we use our 2-D determinants to create a new orthogonal vector in $\mathbb{R}^3$ and why does this work? Also is there a way to be "efficient" with our information storage. Yes to both!

- There are three determinants to encode, so let's store them in a vector, which is a structure in $\mathbb{R}^3$

- Let's store each determinant in our new vector in the component not involved in our determinant computation. This is both a "smart" encoding in that information is not repeated/duplicated, and doing so means that the vector we establish will be orthogonal to our original vector. When we store values in the "missing" component(for example we store the determinant computed in the xy-plane as the third component of our vector), we encode the information in a way as to encode a structure that is independent in a technical sense from our existing vectors, in a similar manner as to how each basis vector cannot map to another without additional information that expands the possible linear combinations. In a rough sense, I like to think that we are incrementally constructing an orthogonal vector using the size of our determinants to "tweak" each component of our built vector so that if orients itself orthogonal to the plane in $\mathbb{R}^3$ described by $\vec{x}, \vec{y}$. ✓

✓

- I had really hoped to formally elaborate on this point more, but I think that, one, I have looked at this problem for to long and need space for my thoughts on this interpretation to become less muddled. Furthermore, I think that I lack some of the linear-algebra language to put more elegantly on paper the visuals that I have in my head that I suspect relate to span and independence, so I would like to come back to this discussion in maybe a week or so.

✓

The results of the previous discussion provide for a better understanding of the component form of the cross product,shown below. For completion we can show that we now have a mathematical validation for question 2 asked above. Since the result of the cross product is a vector, we can show that it is indeed an orthogonal vector relative to either one of its arguments. Below is dthe definition of the cross product(component form), and proof of orthogonality.

For $\vec{x}, \vec{y} \in \mathbb{R}^3$

$$\vec{x} \times \vec{y} := \begin{bmatrix} det \begin{bmatrix} a_2 & b_2 \\ a_3 & b_3 \end{bmatrix} \\ -det \begin{bmatrix} a_1 & b_1 \\ a_3 & b_3 \end{bmatrix} \\ det \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} \end{bmatrix}.$$

=

$$(a_2b_3 - b_3a_2)i - (a_1b_3 - b_1a_3)j + (a_1b_2 - b_1a_2)k.$$

Note that based on this definition, by "recording" determinants, we can see that the cross product equals zero only if $\vec{x}$, $\vec{y}$ are colinear in $\mathbb{R}^3$ since this is the only case where each projection of $\vec{x}$ equals each projection of $\vec{y}$, meaning that each determinant computes to zero.This gives us a way of talking about angles/information between vectors in $\mathbb{R}^3$ by specifying whether they are parallel or not.In problem 6 it was shown that the cross product also encodes the area of the parallelogram spanned by $\vec{x}$, $\vec{y}$, which extends our connection between geometry and algebra for $\mathbb{R}^3$.

The result of the cross product is an orthogonal vector to the $\vec{x}, \vec{y}$.

$\vec{x} \cdot (\vec{x} \times \vec{y}) = 0$ — same idea follows for other argument

$x_1(x_2y_3 - y_2x_3) + x_2(x_3y_1 - y_3x_1) + x_3(x_1y_2 - y_2x_1)$ — dot product yields a real number ✓

$x_1x_2y_3 - x_1x_3y_2 + x_2x_3y_1 - x_2y_3x_1 + x_3x_1y_2 - x_3x_1y_2 = 0.$ — terms sum to zero, so arguments are orthogonal

I see you having some fun with determinants. I have a fun problem involving the volume of a parallelpiped related to the determinant of a 3 x 3 matrix. If you're interested, please ask me about this during an office hour… I have worked on that problem for over 6 hours and still don't have a great solution. Maybe you can make some progress :)

Remark 2 (Thinking in Terms of Projections)**.** *The previous result, provides another way for thinking about the cross product definition in terms of projections. If the vector representing the cross product is to be orthogonal to $\vec{x}$, $\vec{y}$, than representing it as the sum of a projection vector, and a residual vector should be the same as simply expressing the residual vector since the projection of an vector onto an orthogonal plane is 0. Let c denote $\vec{x} \times \vec{y}$ (using implementation above) and $\alpha, \beta \in \mathbb{R}$*

$$\begin{bmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{bmatrix} - proj_{\vec{x}\vec{y}}(c) = r$$

$$proj_{\vec{x}\vec{y}}(c) = \frac{c \cdot (\alpha\vec{x} + \beta\vec{y})}{\|\alpha\vec{x} + \beta\vec{y})\|^2} \cdot (\alpha\vec{x} + \beta\vec{y})$$

Applying dot product linearity and homogeneity to the numerator

$$\to c \cdot (\alpha\vec{x} + \beta\vec{y}) = \alpha \cdot c \cdot \vec{x} + \beta c \cdot \vec{y}$$

$$= \alpha \cdot 0 + \beta \cdot 0.$$

so the whole equation becomes

$$c - \vec{0} = r$$

✓

Which means that our specification of the cross product has no projection in the plane spanned by $\vec{x}, \vec{y}$, and is orthogonal to this plane.

Remark 3. *A commonly cited reason for why the second component of our cross product is negative is so that is corresponds with our right-hand rule for establishing which of two possible orthogonal vectors is specified by the cross product. I am not a huge fan of this, because it feels a bit vague or like the math should establish the precedence for the right-hand tool rather than the converse. I have an intuition that a more pleasing result could be developed by exploring the nature of determinants father, but can at least understand why we append a negative to the second component of the cross product vector. The determinant between the x and z basis vectors, and other vectors lying in this plane, is positive, since the z basis vector is clockwise from the x basis vector, and we take the determinant with x as the left hand argument. Thus if, we want the right-hand rule to be consistent for all planes, we have to make this determinant negative, otherwise it would point in the direction of the positive y basis vector which is diametrically opposed to the direction specified using the right hand rule, in this case, so we flip it.*

Awesome. I am right with you in developing a more convincing approach to this derivation!

1 Some Extended Results

✓ 1.1 The law of cosines

$$(a - b\cos\theta)^2 + (b\sin\theta)^2 = c^2$$

$$c^2 = a^2 - 2ab\cos\theta + b^2\cos^2\theta + b^2\sin^2\theta$$

$$c^2 = a^2 - 2ab\cos\theta + b^2(\cos^2\theta + \sin^2\theta)$$

$$= a^2 - 2ab\cos\theta = b^2.$$

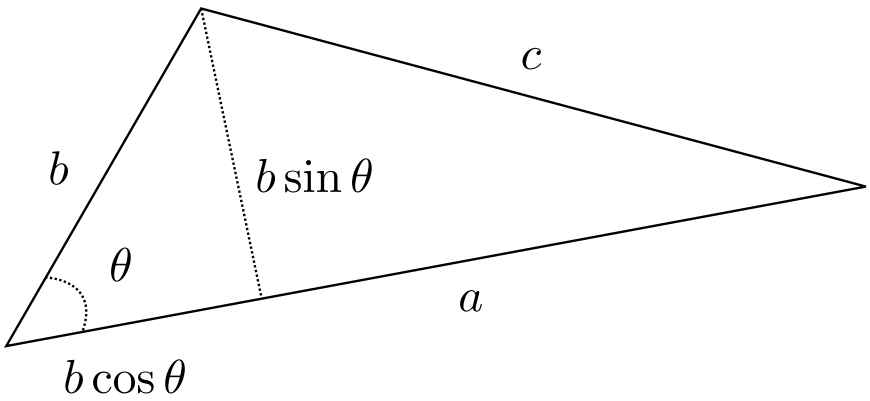


Figure 5: Problem Context

1.2 Geometric Interpretation of the Dot Product

Use law of cosines, with a,b,c specified in relation to vectors.

$$| a - b |^2 = | a |^2 + | b |^2 - 2 | a || b | \cos^2 \theta$$

Expand left side first

$$| a - b |^2 = (a - b) \cdot (a - b) = (a - b) \cdot a - (a - b) \cdot b$$

$$a \cdot a - a \cdot b - a \cdot b - b \cdot b$$

$$= | a |^2 - 2an + | b |^2$$

$$\therefore | a |^2 - 2ab - | b |^2 = | a |^2 = | b |^2 - 2 | a || b | \cos^2 \theta$$

$$a \cdot b = | a || b | \cos\theta.$$

1.3 $\mathbb{R}^2$ Determinants Size is Equal to Area

In problem 5, a component form for parallelogram area was given, and was shown to be equal to the determinant. Area will be non-negative,but the determinant may be negative depending the rotation angle between the vectors involved. In problem 5, we showed that when the area is positive, then the component formula works out to be equal to the determinant, but it would be helpful to show that at any point the determinant will provide useful(accurate) information about a related parallelogram. The following derivation will provide for this since talking about the parallelogram with norms rather than components will allow us to consider generally how the determinant is related. (A expansion step below has been omitted for the sake of showing the result more concisely, but I have it on paper if needed.)

$A_p = \|a\|\|b\| \sin\theta$

$Consider : \sin\theta = \sqrt{1 - \cos^2\theta}$

$$\sqrt{1 - \frac{(a_1b_1 + a_2b_2)^2}{(a_1^2 + a_2^2)(b_1^2 + b_2^2}}$$

$$\sqrt{\frac{(a_1^2 + a_2^2)(b_1^2 + b_2^2)(a_1b_1 + a_2b_2)^2}{(a_1^2 + a_2^2)(b_1^2 + b_2^2)}}$$

Intermediate step of expanding product of numerator, and canceling like terms

$$\sqrt{\frac{(a_1^2b_2^2 + a_2^2b_1^2 - 2a_1a_2b_1b_2)}{(a_1^2 + a_2^2)(b_1^2 + b_2^2)}}$$

Adding in $\|a\|\|b\|$, denominators cancel $\implies \sqrt{a_1^2b_2^2 - 2a_1a_2b_1b_2 + a_2^2b_1^2}$

$\sqrt{(a_1b_2 - a_2b_1)^2} =\mid a_1b_2 - b_2a_1 \mid=\mid det[a, b] \mid .$