Matrices, Vectors, and Matrix Multiplication

by Doug Corey, BYU

This paper is to give an intuitive and meaningful introduction to basic ideas in linear algebra. These include a vector, matrix, linear independence, span, matrix multiplication, dot product, subspace, null space, and others. My hope is that if we start off thinking about concrete situations in a natural way, then the mathematics that we build will make sense and be more easily understood. I start off with a different way of thinking about vectors – not as directed arrows, but as packets of numbers. (OK, I do both, but emphasize the packet of numbers idea). Of course the two ideas are related, but thinking about the vectors as packets of numbers in a specific context helps the computations and algebra make more sense. Moreover, in many applications, the vectors are more likely to be recognized as a packet of numbers, not as an arrow or another representation.

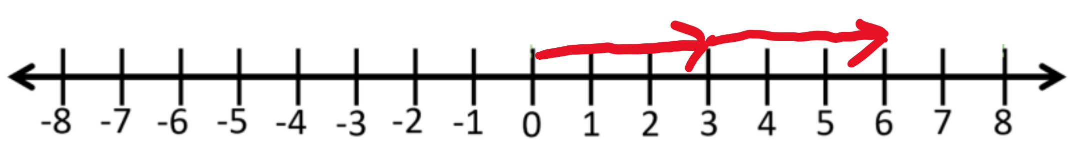
Chapter 1

Vectors as Packets of Quantities . . .Stuck together . . . with Peanut Butter

Let's think about some simple arithmetic, which I will use to make some analogies later on.

Multiplication and Division (with packets): Start with 2*3. What does it mean? 2 copies of groups of three. We don't usually make a distinction between the function of the two numbers – one is the number of groups (2) and the other the size of the group (3) – because the end result is the same. However, it helps to make this distinction because there is a difference with matrices and vectors.

My teacher in elementary school used an arrow notation. You may not have seen this before, but it is common when students start adding integers (positive and negative whole numbers, plus zero). I was taught to think about the three as an arrow going from 0 to 3. Adding 3 is moving to the right three numbers on the number line. 2*3 can be represented as two arrows, of length three, placed end to end, beginning at 0. Since the last arrow ends at 6, we take 6 to be the answer to 2*3. The arrows show the process of going to the right 3 units, two times, for a total of six units.



Note: As you may have picked up from the introduction, arrows are a way to represent a vector so we will be using them a lot.

We can also think about 2*3 where three is a packet of 3 items. A packet is a group that can't be broken up or is easier to work with when together. Lots of things come in natural "packets": a dozen doughnuts, six packs of soda, etc, but also conversions and rates can be thought of as packets: inches in a foot, feet in a mile, lbs in a gallon, etc. A lot of the multiplication and division problems that we have students do are based on different kinds of "packets."

How many sodas do you buy if you need 20 sodas? 20. That is just a counting problem. But how many six-packs of sodas do you buy if you need 20 sodas? Now we are dealing in a "packet" problem. There are many ways to solve this, but we can use the arrow representation earlier. How many arrows of length six do we line up (starting at 0) do we need until we get to 20? That may not be the easiest way of solving this problem, but being able to think about solving this kind of "packet" problem this way will be very helpful later.

Ok, no sweat.

Different kinds of packets: Almost all of the packets that you have run into in your mathematical career only has one quantity in it. They are just packets with one number. However, you don't have to look very far to realize that there are different kinds of quantitative relationships in the world that are worth understanding where the packets contain many quantities. Consider a serving of peanut butter and the nutritional facts associated with each serving.

If we just pick ONE quantity to associate with a serving, Calories for example, then we are still dealing with a packet of one number, a serving has 200 Calories. But if we consider a serving of peanut butter and ALL of the nutritional quantities then we are in a different situation. I have a packet (one serving) associated with MANY quantities (Calories, tablespoons, grams, grams of protein, grams of fat, milligrams of sodium, grams of carbohydrates, etc.) This is a different kind of "packet" and is the kind of packet that fits the idea of a vector. This packet is not just a group size, like a dozen doughnuts, but it is combining multiple quantities and tying them all up together. There are situations where the quantities are, in a sense, stuck to each other. If you want one, then you will get the others. If you want to increase the amount of protein you get by eating peanut butter, then you also increase your amount of all of those other "things" in peanut butter – whether good or bad. You increase your fat, your calories, your sodium, etc. If you want to lower the amount of fat that you eat by decreasing the amount of peanut butter that you eat, then you also lower your protein, your carbohydrates, your fiber, etc.

Packets of quantities are used all the time, especially to model or analyze phenomenon:

- A force has a direction and strength and is often represented as a packet of three numbers, the direction in each of the three coordinate directions x, y, z.
- Each subject in a study has a packet of quantities associated with them. For example, in a study predicting the likelihood of a heart attack, researchers would gather medical information like height, weight, percent body fat, bad cholesterol intake, frequency of exercise, age, sex, etc.
- Populations are often modeled as a packet of quantities. In ecology a population has three natural groups, pre-reproduction stage, reproduction stage, and post reproduction stage. Models use the amount in each group (and the change between groups over time) to understand long term effects of environmental damage or strategies to increase population.
- Color in computers is represented as the amount of Red, Green, and Blue needed to make the color.
- Location on a computer screen or in a computer game is represented by a list of two or three coordinates.
- The status of a character in computer games often requires multiple quantities (life, power, speed, size, money, etc.).
- Dynamical systems (systems that change over time or in response to forces) represent the quantities and how they change in packets (positions, temperature, velocity, acceleration, force, pressure, torque, rotational velocity, etc.)
- The weather at a given time and location can be represented by several quantities (temperature, air pressure, wind speed, wind direction, percent of cloud cover, chance of precipitation, humidity, dew temperature, etc.)
- In a network each person is either associated with others (or not). This is often represented as a list of numbers (or packet) for each person with 0 representing no association with a particular person, and a 1 representing an association.
- In sports, individuals have different rates of production for different statistics. In basketball, for example, we can estimate the rates of points/48 minutes, rebounds/48 minutes, assists/48 minutes, steals/48 minutes, fouls/48 minutes etc. But these are all one packet. You can't mix and match to get Steph Curry's points and Kevin Durrant's rebounds.

Notation: How should we represent a packet of numbers like the nutritional facts in peanut butter? Suppose we are focusing on the three quantities in our diet: the total number of grams, the grams of fat, and the grams of protein. For each serving of PB, we get the values 32, 15, 8. Should we just work with them in a list: 32 15 8. (This doesn't work well in practice since it is hard to keep the spacing consistent.) Or maybe separate them with commas 32, 15, 8 ? Well, there are a few agreed upon ways to write down a packet of numbers. One is using notation that looks like an ordered pair (x,y), but with possibly more numbers: (32, 15, 8). Another is to list the numbers vertically (easier to tell where one number ends and another begins) but to put them in a set of brackets (so we can tell where one list ends and another begins). It looks like this: $\begin{bmatrix} 32 \\ 15 \\ 8 \end{bmatrix}$. Most teachers will use both at times throughout a class, but the latter one is used more often when doing computations. Now I have already done a mathematician trick of taking off the units. We will be using units throughout this adventure, but since the mathematical ideas related to matrices and vectors are generalizations across a lot of contexts, mathematicians don't usually associate their matrices/vectors with units. I wrote the vectors without units because that is how you will see them most of the time in a typical class.

Combining Packets

Scalars: Suppose you take 2 servings of peanut butter well then you get 2 times as much of each of the amounts given per serving. The mathematics related to this point is the proportional relationship between changing one amount and all of the other amounts. To change your diet of one nutritional element by a factor of K they all need to change by a factor of K. We can represent this with the following:

$k\ servings * \begin{bmatrix} 32\ g\ Total/serving \\ 15\ g\ of\ Fat/serving \\ 8\ g\ of\ Protein/serving \end{bmatrix} = \begin{bmatrix} k * 32\ g\ Total \\ k * 15\ g\ of\ Fat \\ k * 8\ g\ of\ Protein \end{bmatrix}$ or without units $k * \begin{bmatrix} 32 \\ 15 \\ 8 \end{bmatrix} = \begin{bmatrix} k * 32 \\ k * 15 \\ k * 8 \end{bmatrix}$

We call this a scalar multiple of a vector. Double the number of servings, you get double the amount of each quantity associated with that serving.

Note: It is called **scalar multiplication** because you can think of it as changing the scale, or resizing, a group. If we think about the 2*3 where the three is represented by an arrow that points from 0 to 3, then 2*3 can be thought of as the 3-arrow stretched to be twice as long, from 0 to six. 1.1*3 is an arrow that is stretched just a little, to point to 3.3. The multiplication -0.2*3 shrinks the 3-arrow to point from 0 to -0.6, so switching the opposite direction. Because we can think of these multiplications as changing the size, or scale, of the arrow, but it doesn't change the line it lives on, we call it a scalar multiplication. This kind of multiplication has a stretching or shrinking effect.

This scalar multiplication matches the only multiplication students are familiar with – the kind we started off talking about: 2*3. In the same way that we used arrows on a number line to figure out what 2*3 should equal (what number are we pointing to if we line up two arrows of length 3 in the positive direction?) we can use the same strategy to make sense of scalar multiplication of a vector, except our vector arrow no longer lives on a number line, but on a 2-dimensional plane, or 3D space, or some higher dimension – depending on the number of entries in our vector. Let's start with an easy case, a 2 dimensional vector of protein per serving and fat per serving from our PB jar: $\begin{bmatrix} 8\ g\ of\ Protein/serving \\ 15\ g\ of\ Fat/serving \end{bmatrix}$ or $\begin{bmatrix} 8 \\ 15 \end{bmatrix}$ without units. The vector points from the origin to the point representing 8 grams of Protein and 15 grams of Fat. See Figure 1(a). If we take two servings of PB how much protein and fat would we get? We can line up two of the same vector arrows (see Figure 1(b)) to see that we should end up with 16 grams of Protein and 30 grams of Fat. The computation should look like this:

$$2\ servings * \begin{bmatrix} 8\ g\ of \dfrac{Protein}{serving} \\ 15\ g\ of \dfrac{Fat}{serving} \end{bmatrix} = \begin{bmatrix} 2\ servings * 8\ g\ of \dfrac{Protein}{serving} \\ 2\ servings * 15\ g\ of \dfrac{Fat}{serving} \end{bmatrix} = \begin{bmatrix} 16\ g\ of\ Protein \\ 30\ g\ of\ Fat \end{bmatrix}$$

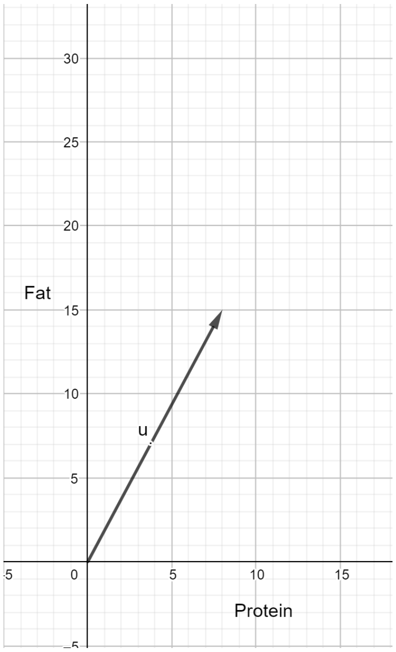


Figure 1(a)

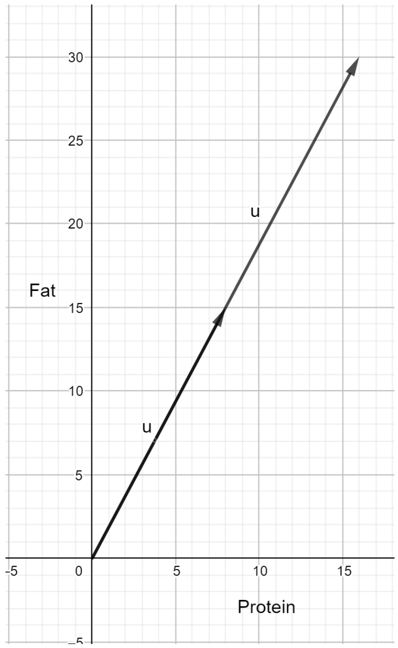
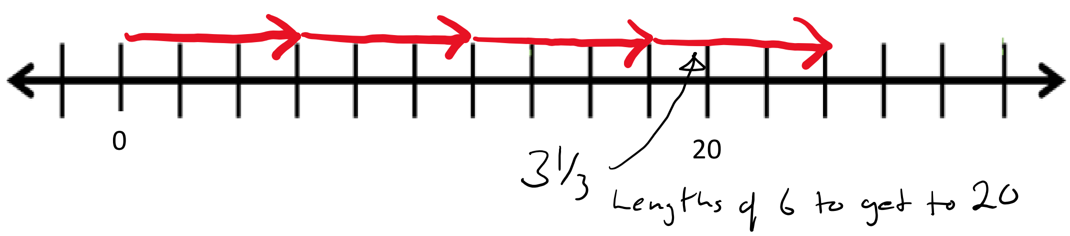


Figure 1 (b)

Division as Lining up vectors end-to-end: Let's go back to the division problem we looked at earlier. How many 6-packs of soda do you need to buy to get 20 sodas? We can think about this as having a point at 20 on the number line and figuring out how many packets of size 6 (or arrows of length six) we need to line up to get to the number 20 by starting at 0.



In this problem we had a goal (20 cans) and we had a packet (six pack) and needed to know how many packets are needed to reach the goal. In our PB context there is a similar problem. If we set a dietary goal, say to increase Protein by 24 grams and Fat by 45 grams, how many packets (servings of PB) that contain 8 grams of Protein and 15 grams of Fat per serving are needed to reach that goal (starting at (0,0), because with no change in packets, the Fat and Protein don't change)? We can still think of this in the same way, but the point that we are trying to reach (24 Protein, 45 Fat) is placed on a coordinate plane, not just a real number line. Also, the vector, or arrow representation of the vector, is drawn from the origin to the point (8, 15). Now we can use the same process as the soda problem, line up vectors end-to-end until we get to the point (24, 45). The vectors that we line up though ALWAYS have to increase the protein and fat in a ratio of 8 to 15! Why? Because that is the ratio of protein to fat in a serving of peanut butter. To get some amount of protein – you need to take in 15/8 that amount in fat. When you take one, you take the other, and if you take k as much of one, you get k as much as the other. In this problem, we need to take 3 servings of PB to reach our goal.

In our mathematical notation we can state this problem as an equation which we need to solve for k:

$k * \begin{bmatrix} 8 \\ 15 \end{bmatrix}$ = $\begin{bmatrix} 24 \\ 45 \end{bmatrix}$, which can by solved like this: $\begin{bmatrix} k * 8 \\ k * 15 \end{bmatrix}$ = $\begin{bmatrix} 24 \\ 45 \end{bmatrix}$ → $\begin{bmatrix} k * 8 = 24 \\ k * 15 = 45 \end{bmatrix}$, then solving either one of the equations gets us k=3.

[You may be wondering why we didn't divide. Afterall, in the soda problem you would have found the answer by dividing 20 by 6 to get 3 and 1/3. But this would require us to divide a vector by a vector. How would we do it? And what should we end up with? Should a vector divided by a vector equal a vector? a scalar? a matrix? It seems like from above it should be a scalar, but as we will learn later, there is a matrix that we could have substituted in for k that would have also worked. Because of these (and other) issues, it is best to avoid division and think about these problems as multiplication problems: What can I multiply this vector by to get that vector?]

But that is too much Fat!: OK, well if you are a typical American (like me), you already have too much fat in your diet (although fat does get a bad rap and some dieticians would say sugar is worse for you than fat). What if someone wants to increase their Protein by 16 grams, but only increase their Fat by 10 grams? (Or 4? Or even decrease it? But for now let's stick with 10 as our desired goal). This is a problem I want you to think about for a minute. How many servings of PB would you need to change your diet by 16 grams of Protein and 10 grams of Fat?

PLEASE THINK ABOUT THE PROBLEM BEFORE MOVING ON

Well I have sad news for you. You can't do that with peanut butter. Why? Well there are different ways to see this. Graphically, what we are dealing with is that we are lining up vectors end to end trying to reach a point that does not lie on the same line as the vector. No matter how many vectors you line up (going either direction, you can take a negative amount and start lining up vectors in the opposite direction) you won't hit that point. See Figure 2.

[Some of you are wondering how you eat a negative serving of PB. You can't. But if you are already eating some PB you can eat a serving less, or two servings less, etc, and the change it will make in your diet is represented by the arrows going in the opposite direction. For this food example, we are assuming you are eating some PB already]

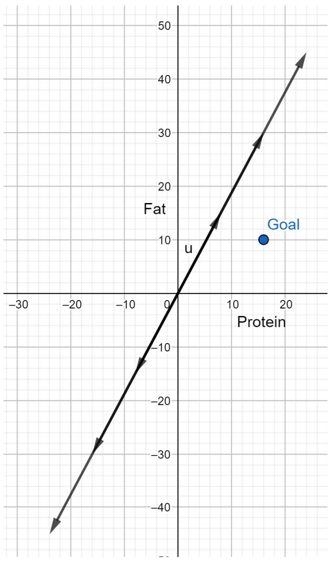


Figure 2. Multiples of the vector always lie on a single line, but the goal is off of the line.

Another way to see this is trying to solve the same kind of equation we solved earlier:

$k * \begin{bmatrix} 8 \\ 15 \end{bmatrix} = \begin{bmatrix} 16 \\ 10 \end{bmatrix} \rightarrow \begin{bmatrix} k * 8 \\ k * 15 \end{bmatrix} = \begin{bmatrix} 16 \\ 10 \end{bmatrix} \rightarrow \begin{bmatrix} k * 8 = 16 \\ k * 15 = 10 \end{bmatrix}$

However, from this point we have a problem, since k needs to equal 2 in order to make the top equation true, but k needs to equal 2/3 for the bottom equation. So there is no solution.

So what can you do to meet your 16 Protein-10 Fat goal? You have three options: 1) find a different food with the ratio of protein to fat that matches the ratio of protein and fat in your goal. We can call this the **Vector Match** strategy. 2) Find another food (or other foods) in addition to peanut butter that you are already eating and are willing to change the amount in your diet, but has a different ratio of protein to fat than peanut butter. The hope is that you can eat some amount of one and some amount of another to reach your goal. We can call this the **Combination** strategy. 3) Change the amount of PB in your diet to get as close as possible to your goal, but give up on meeting the exact goal. We could call this the **Best Approximation** strategy. Option 1, vector matching, might be hard since it might be hard to find a food whose Protein-Fat ratio fits your goal. Option 2 sounds doable, after all, you don't just eat PB (I hope). Option 3 might be the easiest (after all, isn't changing the goal the easiest way to succeed in life?) but trying to decide how much to eat to get you "closest" to your goal actually just raises another math problem. (It turns out to be a very important math problem that we will get to later.) In the next chapter we will explore the combination strategy.

Chapter 2

Combinations of Vectors . . . or Making a PB Sandwhich

Getting the Recipe Right. Let's start by exploring the **Combination** strategy. We need to find another food that has a different ratio of Protein to Fat. (If we don't, then we just get another food that has a vector that lies on the same line as PB, so it doesn't help us to achieve any goals that we couldn't already reach with PB. There is a special name for this, see the note below). Figure 3 shows the nutritional facts for Oroweat Country White Bread. Notice that in one serving of the bread there are three grams of protein and 1.5 grams of fat. Remember that we are shooting for an increase of 10 grams of fat to go along with our 16 grams of protein. It is easy to check that this ratio is different than the peanut butter, since there was more fat than protein in serving of PB, and in the bread there is less fat than protein in a serving, so they can't be the same.

NOTE: In the previous paragraph I mentioned that adding in a certain food might not help you achieve any more dietary goals than the food(s) you already have. This is connected to a fundamental idea related to vectors: ***linear independence*** and ***linear dependence***. (We will come back to this later but I wanted to start you thinking about the idea now.) If you have a set of foods, then those foods would allow you to reach certain goals. If you could remove a food from the set and still reach all of the same dietary goals, then that set of food is considered to be dependent (or linearly dependent). In this case, no single food gets you anything that a combination of the other foods wouldn't get you. [It turns out if you can remove one and it doesn't change the goals you can achieve, then you can remove any one of the foods and it won't change the dietary goals you can reach.] If you remove any food and it changes (reduces) the number of dietary goals you can achieve, then the set of foods is called linearly independent. In this case each food contributes something that the combination of the other foods can't get you. Of course, I have described these ideas in the food context, but the idea is the same for general vectors. What do combinations of vectors look like? Keep reading.

Nutrition Facts

Serving Size: 1 slice (38g)

Amount Per Serving	
Calories 100	Calories from Fat 14
	% Daily Value*
Total Fat 1.5g	**2%**
Saturated Fat 0g	**0%**
Trans Fat 0g	
Cholesterol 0mg	**0%**
Sodium 140mg	**6%**
Total Carbohydrates 19g	**6%**
Dietary Fiber 0.5g	**2%**
Sugars 2g	
Protein 3g	

Vitamin A	0%
Vitamin C	0%
Calcium	4%
Iron	6%

* Percent Daily Values are based on a 2000 calorie diet.

INGREDIENTS: Unbleached Enriched Flour [Wheat Flour, Malted Barley Flour, Reduced Iron, Niacin, Thiamin Mononitrate (Vitamin B1), Riboflavin (Vitamin B2), Folic Acid], Water, Sugar, Yeast, Soybean Oil, Salt, Butter (Milk), Nonfat Milk, Wheat Gluten, Honey, Monoglycerides, Calcium Propionate (Preservative), Ascorbic Acid (Dough Conditioner), Soy Lecithin.

Figure 3. Nutritional Facts of Oroweat Country White Bread.

We now have a point to reach (16, 10) or (16 grams of protein, 10 grams of fat). But we also have 2 foods, or two different packets. We have a lot more options now! We can vary the amount of peanut butter and bread that we eat and we can do it independently from each other. As mentioned earlier, if we are already eating some servings each of bread and peanut butter, we can even change by a negative value of servings. Figure 4 shows 2 vectors representing a serving of PB and Bread. Each vector in Figure 3 shows the result of just eating one serving of each of the foods separately, not together. What would happen if we ate one serving (a slice) of bread and one serving of PB? How much Protein and Fat would we get?

8g of Pro+3 g of Pro=11 g of Pro

15 g of Fat+1.5 g of Fat= 16.5 g of Fat

We would have a total of 11 g of Protein and 16.5 grams of Fat.

Vector Addition. Another way of thinking about what we just did is that we just added two vectors. If we let the units guide us, like above, then there seems to be only one natural way to add vectors: add the corresponding values.

$\begin{bmatrix} 8\ g\ of\ P \\ 15\ g\ of\ F \end{bmatrix} + \begin{bmatrix} 3\ g\ of\ P \\ 1.5\ g\ of\ F \end{bmatrix} = \begin{bmatrix} 11\ g\ of\ P \\ 16.5\ g\ of\ F \end{bmatrix}$ Or without units: $\begin{bmatrix} 8 \\ 15 \end{bmatrix} + \begin{bmatrix} 3 \\ 1.5 \end{bmatrix} = \begin{bmatrix} 11 \\ 16.5 \end{bmatrix}$

In Figure 4 we see the two vectors on the left, and the addition of the two vectors on the right (green vector). I have lined up the PB and Bread arrows end to end, then drew a new vector to the point that the last vector was pointing to. Of course, that point is the same one that represents the sum from above. Lining up vectors end to end for addition makes sense because that is one way to represent addition of real numbers on a number line.

Your turn to think: What if we lined up the arrows in the figures below in a different order, would they still point to the same point, (11P,16.5F)?

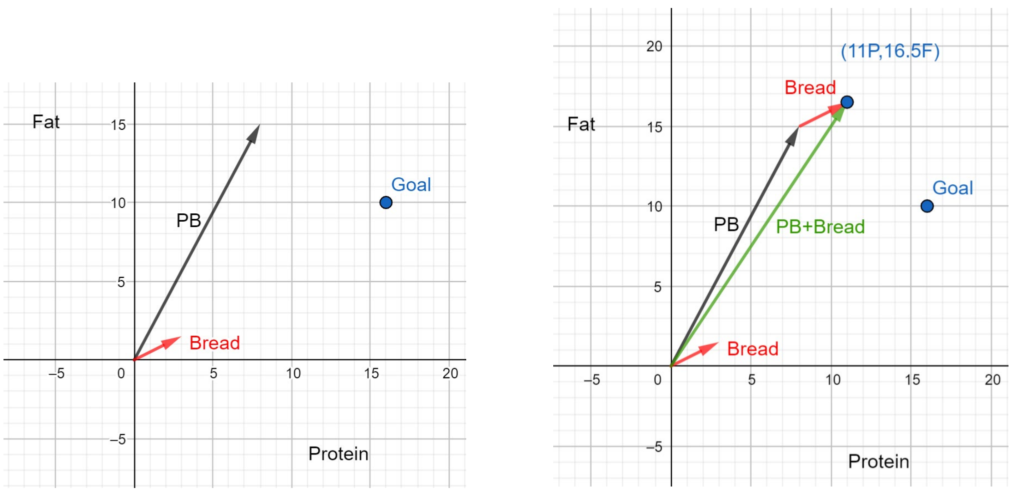


Figure 4. (Left)Vectors representing a serving of PB and a serving of Bread. (Right) The green vector shows the result of PB+Bread

Your turn to think (again): What would PB-Bread look like? (Where would your Protein-Fat intake end up if you increased your diet by one serving of PB and decreased it by one serving of Bread?) Can you generalize the process of how to subtract vectors and how to represent them on the coordinate plane?

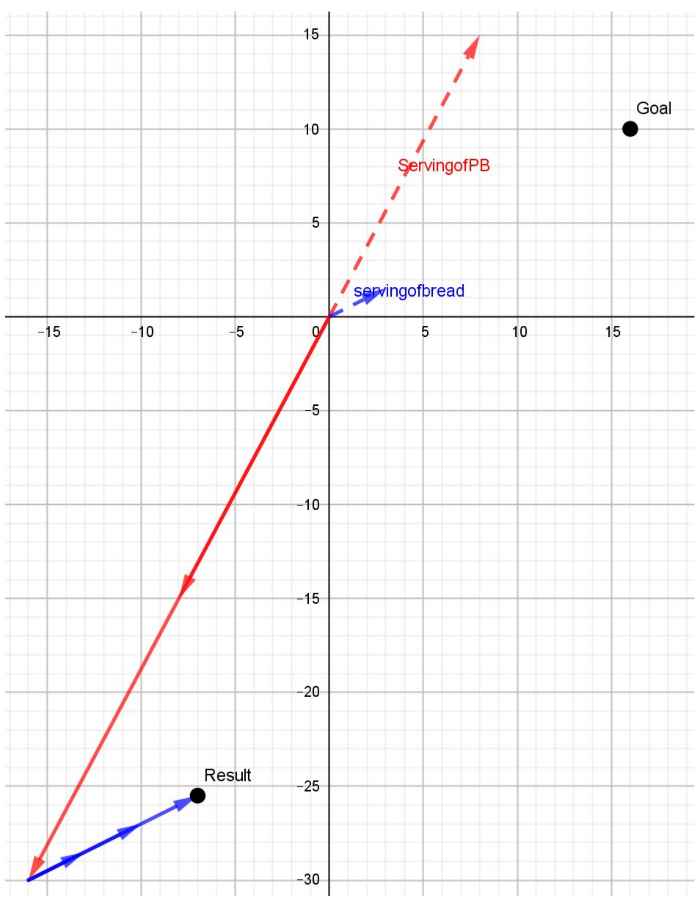
Different Combinations: We are trying to figure out how many servings of PB and Bread are needed to get to our new nutritional goal of Protein up 16 and Fat only up 10. But before we do that, let's get a feel for what happens as we take different amounts of servings for PB and Bread. For example, let's think about what happens if we eat two fewer servings of peanut butter in our daily diet and add three more servings of bread to our diet. We have already dealt with representations on how to handle each one of these individually (if we are only changing the amount of one food, we are in a scalar multiple situation). We also know how to add one serving of each. What does adding multiple servings of each look like?

Computationally, we can calculate how our nutritional intake will have changed. The computation will tell us what point we end up on the Protein-Fat coordinate plane. The calculation is below. I included the units at the beginning and at the end but dropped them in the middle to save space. It would be a good activity to rewrite all of this with the exact units for each number (not just something like grams of fat per serving, but grams of fat per serving of peanut butter).

$$-2\ serv * \begin{bmatrix} 15\ g\ fat/serv \\ 8\ g\ prot/serv \end{bmatrix} + 3 * \begin{bmatrix} 1.5 \\ 3 \end{bmatrix} = \begin{bmatrix} -2 * 15 \\ -2 * 8 \end{bmatrix} + \begin{bmatrix} 3 * 1.5 \\ 3 * 3 \end{bmatrix} = \begin{bmatrix} -2 * 15 + 3 * 1.5 \\ -2 * 8 + 3 * 3 \end{bmatrix} = \begin{bmatrix} -25.5\ g\ of\ fat \\ -7\ g\ of\ prot \end{bmatrix}$$

Your turn to think: Where are the peanut butter quantities? Where are the bread quantities? Where are the fat quantities? Where are the protein quantities?

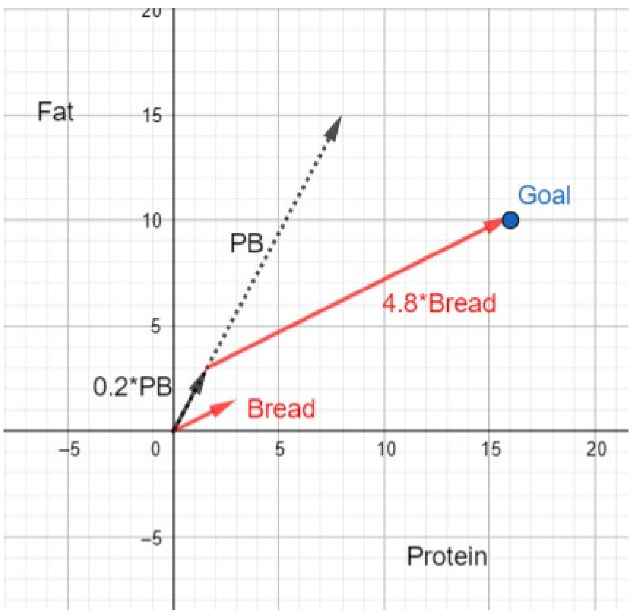
Below is an image that shows the result and another way to find it: lining up arrows end to end. If you didn't already do this earlier, you should convince yourself that the order in which we lined up the arrows doesn't matter. This is really nice, otherwise the order you eat the servings of food would determine your Protein-Fat intake. (OK, on second thought, maybe that isn't so nice. It would be really cool if you could lower your Fat intake just by eating the same food, just in a different order!) This is nice mathematically because it means that order doesn't matter in vector addition, that is, vector addition is commutative; -2PB+3B=3B-2PB=B-2PB+2B= B-PB+B-PB+B= etc.



Well, it looks like we didn't do a good job of choosing how to change the number of servings of each food to get close to our 10 grams of fat and 16 grams of protein. We actually dropped in both! How can we find the amount of PB and the amount of Bread that we need to reach our new goal? OR maybe there are multiple combinations that could get us to the same spot, in that case, we need to find at least one. We could get an estimate from the graph by lining up the arrows of PB and B to try to hit (or get close to) our goal.

[Link to geogebra App to find the solution]

In the figure below it looks like we get close with 4.8 slices of Bread and .2 servings of PB. (Note: In the figure below I didn't represent 4.8 servings of Bread with 4.8 vectors lined up end to end. I just drew a vector that goes in the same direction as Bread but made it 4.8 times as long. Similarly with the PB, I just scaled the PB vector down to be .2 of the length of the vector that represents a single serving of PB. Shrinking and stretching a vector is an equivalent way to think about scalar multiplication of a vector as lining up vectors end to end.)



Well how close did we get with the estimates from the graph? We can actually do the calculation to figure out exactly where we ended up. Actually I will let you do that. Before continuing on try to find the exact coordinates of the point that we ended up at with .2 servings of PB and 4.8 servings of Bread.

One way to figure this out is to do similar calculations to what we did above:

$$0.2\ serv * \begin{bmatrix} 15\ g\dfrac{fat}{serv} \\ 8\ g\dfrac{prot}{serv} \end{bmatrix} + 4.8 * \begin{bmatrix} 1.5 \\ 3 \end{bmatrix} = \begin{bmatrix} 0.2 * 15 \\ 0.2 * 8 \end{bmatrix} + \begin{bmatrix} 4.8 * 1.5 \\ 4.8 * 3 \end{bmatrix} = \begin{bmatrix} 0.2 * 15 + 4.8 * 1.5 \\ 0.2 * 8 + 4.8 * 3 \end{bmatrix} = \begin{bmatrix} 10.2\ g\ of\ fat \\ 16\ g\ of\ prot \end{bmatrix}$$

There are two things that we should address. First, This gets us pretty close. We got the 16 g of Protein exactly, but just a touch high on the fat (we were shooting for 10 g of Fat). This is probably close enough for this context, but not all contexts are as forgiving (try telling the accountant he/she is off by 2% and see if they say, "Well, that is close enough!"). Two percent error is enough to ruin many chemical experiments or engineering devices. So we are still left with an important question: How do we find out exactly (or much more accurately if not exactly) how many of each packet we need to reach a specified goal? Or stated another way, how can we take a combination of vectors to get a new vector that points to a desired point?

Second, I just want to highlight how we are combining the vectors. We have already done this a few times so far, but it is a key idea moving forward in linear algebra. Notice that we are adding up scaled packets. We could also subtract scaled packets but that is really no different than adding packets and scaling the packet in the opposite direction (changing the sign). What I mean by "adding up scaled packets" is the first expression in the calculation above (shown below).

$$0.2\,serv * \begin{bmatrix} 15\,g\dfrac{fat}{serv} \\ 8\,g\dfrac{prot}{serv} \end{bmatrix} + 4.8 * \begin{bmatrix} 1.5 \\ 3 \end{bmatrix}$$

To keep from getting lost in the numbers and just to see the structure of the calculation I am going to represent the scalars (.2 and 4.8) as the letters a_1 and a_2. I will represent our packets/vectors as just p_1 and p_2. Now we can write down the calculation in a way that is perhaps a little easier to see the structure: $a_1* p_1+ a_2* p_2$. This represesents a_1 of packet p_1 plus a_2 of a packet p_2. In our dietary context we could start adding food (like a serving of Jam) and then eating different amounts of PB, Bread, and Jam could be represented as:

$a_1* p_1+ a_2* p_2+ a_1* p_1+ a_3* p_3$

Throw in a glass of milk and we would get:

$a_1* p_1+ a_2* p_2+ a_1* p_1+ a_3* p_3 + a_4* p_4$

Mathematically we could keep on going. This structure is called a **linear combination** of packets. A linear combination has scalars (a_1,a_2,etc.) and a set of packets (p_1, p_2, etc) that are then combined by multiplying each of the packets by one of the scalars then adding them up. The scalars are called the **coefficients** of the linear combination. Now I am sounding a lot like a mathematics textbook, but let's get back to the ***idea*** of a linear combination – that we take some of one thing, some amount of a second thing, some amount of a third thing etc. and combine them. That is what we have been doing with our food. We have taken some amount of PB and some amount of Bread (and some amount of other foods if needed) and add up the quantities associated with each (in this case Fat/serving and Prot/Serving) to find the total amount of Fat and Protein that those foods contribute to our diet.

Why is it called a ***LINEAR*** combination? Think of the equation of a line in standard form Ax+By=C. Can you see the linear combination of x and y on the left hand side? Graphs of lines are generated by plotting all possible linear combinations of A and B that equal a given constant, C. (A and B are the coefficients of x and y, but it in the way we have been thinking about it, it might be easier to consider the packets as the values A and B and the x and y's [which are the amount of A and B we take, respectively] to be the scalars.)

We will get back to our diet problem in a moment, but I want to share something that just might blow your mind. (OK, you should never start off like that unless you really have something amazing to share. This isn't earth shattering, but it may make you see things you thought you knew in a different way.) Here it is: You have been dealing with linear combinations for most of your mathematical life. Wuh? Yes, Look at the list below:

- Let's start with writing down a 3 digit number, say 345. What does that represent? 3 100's, 4 10's, and 5 1's; or 3*100+4*10+5*1. You may have called it "expanded form". Here, the packets are powers of ten.

- Telling time also has packets. Look at the duration of a movie you are streaming and it will say something like 1:47:34, or 1 hour, 47 minutes, and 34 seconds.
- To calculate how much money you have with 3 five dollar bills, 2 ones, 5 quarters, 2 dimes and nickel, what would you do? Something equivalent to 3*5.00+2*1.00+5*.25+2*.1+1*.05; a linear combination of the denominations where the packets are the values of each bill or coin.
- The perimeter of a rectangle can be represented as a linear combination of the lengths of the sides: P=2*L+2*W.
- A polynomial is a linear combination of basic functions, all whole number powers of x. All quadratics, for example are of the form of ax^2+bx+c, or rewritten as $a*x^2+b*x+c*1$ it might be easier to see the linear combination of x^2, x, and 1.
- The total cost of the items you buy at the grocery store is a linear combination of the prices of the items.

Ok, I think that is enough. You probably get the point that linear combinations is something you have been doing for a while, but it hasn't been called that. It is a process that happens in a lot of different contexts for a lot of different reasons, but mathematicians abstract the core idea and give it a name and definition.

Special linear combinations: Dot Products. Although a lot of linear algebra is built around understanding linear combinations of packets, there are two different linear combinations worth talking about. First, the linear combinations of vectors (which we are temporarily considering as packets of two or more quantities), and second the linear combination of individual quantities. Although at the core they are still both linear combinations, they have different enough properties that mathematicians give them different names. We will get to the first later on when we discuss a process called matrix-vector multiplication. The second is a kind of vector-vector multiplication that matches some of the examples above like 345 as a linear combination of 100's, 10's, and 1's. The vector-vector multiplication is called a ***dot product***.

If we write 345=3*100+4*10+5*1 then you might be wondering to yourself, "Where are the vectors?" After all, I did say it was a kind of vector-vector multiplication, but it doesn't look like there are any vectors anywhere in the expression. The six numbers used on the right side of the equation can be written as two, three-quantity packets. I have written them below. The first packet are the digits in the numeral 345. The second packet is the value of that corresponds to the place of each numeral in the place-value system.

$$\begin{bmatrix}3\\4\\5\end{bmatrix} \bullet \begin{bmatrix}100\\10\\1\end{bmatrix} = \begin{matrix}3X100\\4X10\\5X1\end{matrix}$$

Ok, you may be thinking that I am making the number 345 much more difficult than it needs to be. Yes, I am, but the use of dot products comes in very handy as we will see further on and has a property, that to me, is near magical and will save us both a lot of time doing more complex calculations.

The dot product of these two vectors is the sum of the products of the corresponding quantities. In more basic English, to take the dot product multiply the two top numbers of each vector, the two numbers in the second positions, etc, then add up all of those products.

$$\begin{bmatrix}3\\4\\5\end{bmatrix} \bullet \begin{bmatrix}100\\10\\1\end{bmatrix} = \left.\begin{matrix}3X100\\4X10\\5X1\end{matrix}\right\}\text{SUM}$$

The dot product is called the dot product because it is signified by a dot, ●, and is one of three different kinds of products that we use in linear algebra. The others are scalar multiplication and matrix multiplication.

Chapter 3

Using a Place Mat(rix)

Back at the Kitchen: Linear combinations are at the center of linear algebra and we will use this idea throughout the rest of our exploration. They are actually at the center of our current problem about PB and Bread (which you have probably forgotten about.) We are trying to find a combination (see how I threw that word in here) of PB and Bread that would help us reach our goal of 16 g of Protein and 10 g of Fat. We can represent this problem as an equation:

$$a\ serv\ of\ PB * \begin{bmatrix} 15\ g\dfrac{fat}{serv\ of\ PB} \\ 8\ g\dfrac{prot}{serv\ of\ PB} \end{bmatrix} + b\ serv\ of\ Bread * \begin{bmatrix} 1.5\ g\dfrac{fat}{serv\ of\ Bread} \\ 3g\dfrac{prot}{serv\ of\ Bread} \end{bmatrix} = \begin{bmatrix} 10\ g\ fat \\ 16\ g\ prot \end{bmatrix}$$

Where we are trying to find the values of a and b that make the equation true (or make the left side equal the right side). Without the units the equation looks like:

$$a * \begin{bmatrix}15\\8\end{bmatrix} + b * \begin{bmatrix}1.5\\3\end{bmatrix} = \begin{bmatrix}10\\16\end{bmatrix}$$

So how can we solve for a and b?

System of equations: Some of you astute readers may have recognized earlier that we could solve this problem if we thought about it in a different way, a way that many of you may be comfortable with. We could think of this problem as a system of equations with two unknowns and two variables. Let me write down the problem in a typical boring system-of-equations word problem.

> Sara wants to change the amount of protein and fat she has in her diet. She eats a lot of peanut butter and bread and wonders how she should change the amount of each to reach her goal of increasing her protein by 16 grams of fat a day and increasing her grams of fat by 10 grams of fat per day. A serving of peanut butter has 15 grams of fat and a slice of bread has 1.5 grams of fat. A serving of peanut butter has 8 grams of protein and a slice of bread has 3 grams of protein. How many servings of peanut butter and how many slices of bread should she eat to reach her goal?

To solve this problem we can write down two equations (one for Fat and one for Protein) where a and b represent the servings of PB and Bread respectively.

$$15a + 1.5b = 10$$
$$8a + 3b = 16$$

You can use your favorite method to solve this (elimination or substitution) to solve this equation. If we both did it right, we should have found a = 2/9≈0.181818 and b=160/33≈4.8484. Pretty close to our a=.2 and b=4.8 solutions we estimated from the graph.

Now if all we cared about was that burning question about PB and Bread and Fat and Protein then we could stop here. In fact, if all we cared about was this particular problem then you are wondering why we have spent so much time talking about packets, and linear combinations and linear independence when all we needed to do was something that most of you learned years ago. Boy, did I just waste a bunch of your time. Pretty much all I have done so far is (maybe) convince you that a linear combination of vectors is the same as a system of equations. I haven't even taught you a better way to solve a system of equations! How sad!

Most of you have had a unit on matrices in college algebra or pre-calculus. After the unit, you may have felt like matrices are just a harder way to solve a system of equations. It is reasonable to think this because that is basically all matrices are used for in those lower level classes. There is a strong connection between systems of equations and matrices and we will use systems of equations to help us understand what is going on with matrices, and when we do, we will actually gain a lot of insight about systems of equations. Moreover, we will end up learning a lot about the mathematics of packets that allow us to model situations that go beyond systems of equations.

A Matrix Emerges. Well here I am talking about matrices and we haven't even seen one yet in our exploration, but we are just on the verge. Let's go back and look at the equation we set up with vectors and our unknowns, a and b.

$$a * \begin{bmatrix} 15 \\ 8 \end{bmatrix} + b * \begin{bmatrix} 1.5 \\ 3 \end{bmatrix} = \begin{bmatrix} 10 \\ 16 \end{bmatrix}$$

This entire equation is almost all represented as packets. We have a PB packet that represents the grams of Fat and Protein per serving. We have a similar packet for Bread. We have our goal packet that shows how many grams of Fat and Protein we are shooting for. The only other things in the equation are the unknown serving sizes, a and b. Should they be a packet too? We could call it the serving size packet, or since they are unknown and represented by variables we could call it the *variable* packet. But there is something weird here. The serving sizes aren't together in the equation, so if we write the serving size packet like the other packets, where would we put the packet in the equation? And how would we write the equation?

Maybe something like these?

$\begin{bmatrix} a \\ b \end{bmatrix} * \left(\begin{bmatrix} 15 \\ 8 \end{bmatrix} + \begin{bmatrix} 1.5 \\ 3 \end{bmatrix}\right) = \begin{bmatrix} 10 \\ 16 \end{bmatrix}$ or $\left(\begin{bmatrix} 15 \\ 8 \end{bmatrix} + \begin{bmatrix} 1.5 \\ 3 \end{bmatrix}\right) * \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 10 \\ 16 \end{bmatrix}$

But this doesn't work because then we would just add the two PB and B vectors together to get

$\begin{bmatrix} a \\ b \end{bmatrix} * \begin{bmatrix} 16.5 \\ 11 \end{bmatrix} = \begin{bmatrix} 10 \\ 16 \end{bmatrix}$ or $\left(\begin{bmatrix} 16.5 \\ 11 \end{bmatrix}\right) * \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 10 \\ 16 \end{bmatrix}$

And now there is no way to get a different amount of servings for the PB and B. Also, what does that multiplication sign mean between the two vectors (we don't yet know how to multiply vectors). If we think of just multiplying the corresponding values (top * top and bottom* bottom) we don't get the same answer as we did before:

$$\left(\begin{bmatrix} 16.5 \\ 11 \end{bmatrix}\right) * \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 10 \\ 16 \end{bmatrix}$$

$$\begin{bmatrix} 16.5 * a \\ 11 * b \end{bmatrix} = \begin{bmatrix} 10 \\ 16 \end{bmatrix}$$

Means a=20/33 and b=16/11. So this does not seem right, not just because we don't get the same answer, but because in our context "a" and "b" represent the number of servings of PB and Bread respectively, but in the computation above a is just multiplied by Fat value and b is multiplied by a Protein value.

It seems like it would be better to write the serving packet (with a and b) horizontally and put it over the PB and Bread packets.

So $a * \begin{bmatrix} 15 \\ 8 \end{bmatrix} + b * \begin{bmatrix} 1.5 \\ 3 \end{bmatrix} = \begin{bmatrix} 10 \\ 16 \end{bmatrix}$ becomes $\begin{matrix} [a \quad b] \\ \left(\begin{bmatrix} 15 \\ 8 \end{bmatrix} + \begin{bmatrix} 1.5 \\ 3 \end{bmatrix}\right) \end{matrix} = \begin{bmatrix} 10 \\ 16 \end{bmatrix}$. At least this way it makes more sense that "a" is associated with the left packet, but "b" is associated with the right packet. However, we still have the problem with the addition of the PB and B packets. It seems like now we could add them before we do any multiplying by a or b, but we just tried that and it didn't make sense. Furthermore, if we add them together, then we lose the advantage of putting the a and b packet over the two other vectors so we could see that "a" goes to the left one and "b" goes to the right one.

Let's go back to writing the a and b packet either before or after the PB and Bread packets. The addition seems to cause us a lot of problems when we try to "factor out" the a and b packet. If we drop the addition then we get:

$\begin{bmatrix} a \\ b \end{bmatrix} * \begin{bmatrix} 15 \\ 8 \end{bmatrix}\begin{bmatrix} 1.5 \\ 3 \end{bmatrix} = \begin{bmatrix} 10 \\ 16 \end{bmatrix}$ or $\begin{bmatrix} 15 \\ 8 \end{bmatrix}\begin{bmatrix} 1.5 \\ 3 \end{bmatrix} * \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 10 \\ 16 \end{bmatrix}$

We now have the two packets (PB and B) next to each other. They aren't being added or multiplied together. They are just . . . together. Well here is a bit of a crazy idea, so far when we had two quantities that belonged together (fat and protein in a serving of peanut butter, goal of fat and protein intake, etc) we put them together in a packet. What if we made a (drum roll here da da da da da da . . .) a packet of packets! (Wuh?) Yes, mind blowing I know, but we could put the PB and B packets into one packet, and write it like: $\begin{bmatrix} 15 & 1.5 \\ 8 & 3 \end{bmatrix}$. Then our equations would look like: $\begin{bmatrix} a \\ b \end{bmatrix} * \begin{bmatrix} 15 & 1.5 \\ 8 & 3 \end{bmatrix} = \begin{bmatrix} 10 \\ 16 \end{bmatrix}$ or $\begin{bmatrix} 15 & 1.5 \\ 8 & 3 \end{bmatrix} * \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 10 \\ 16 \end{bmatrix}$

Does it matter which way we write it? Well, we are not sure yet, but the convention is to write it like the one on the right. We need to take a step back now because we just did two big things. First, we wrote down our first matrix (OK, technically a vector is a matrix, but we wrote down a matrix that looks different than any vectors that we have used so far). Second, we wrote down a multiplication between a matrix and a vector so we better make sure we know what that means.

Matrices: One way to think about a matrix is that it is a packet of packets, or maybe an easier way to think about it is a packet of vectors. We represent matrices in a similar way to the way we represent vectors but allow them to have multiple columns, not just multiple rows. Technically a vector is a matrix with just one column. All of these below are examples of matrices**.**

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 5 \\ 7 \\ -3 \end{bmatrix}$$

$$\begin{bmatrix} 6 & -2 & -5 \end{bmatrix} \qquad \begin{bmatrix} -1 & 0 & 2 \\ 0 & 4 & 7 \\ -3 & -4 & -2 \end{bmatrix} \qquad \begin{bmatrix} i & 0 \\ 0 & 1-i \end{bmatrix}$$

The entries could be integers, fractions, decimals, even imaginary numbers. There is no limit to the size a matrix can be. Some applications deal with matrices that are more than a thousand rows long and columns wide. (Thank goodness for computers). We name the size of matrices by listing the number of rows then the number of columns. For example the upper-left matrix in the previous figure is called a "2 by 3" matrix and it is written as 2X3.

Matrices are nice ways to hold a lot of quantitative information. The 2X2 matrix that we wrote down in our equation, $\begin{bmatrix} 15 & 1.5 \\ 8 & 3 \end{bmatrix}$, is quite concise, but the structure and numbers represent a lot of information in our context. I have tried to illustrate what this matrix is trying to represent in our particular problem in the table below:

$$\begin{bmatrix} 15 & 1.5 \\ 8 & 3 \end{bmatrix} \; means:$$

	Food	
Nutrients	Peanut Butter	Bread
Fat	*15 g/serving*	*1.5 g/serving*
Protein	*8 g/serving*	*3 g/serving*

Well, if you were doing calculations which one would you like to be working from? The bare matrix captures the needed information for calculations and can then be interpreted in the meaning of the context after the calculations are done. If the table above reminds you of a statistics class, it should. Statisticians use matrices to organize and do calculations with their data. We will see some of the basic ways in which matrices, vectors, and their operations are used in data analysis a little later on.

Matrix Multiplication Unmasked: If there is one point that is more frustrating to students than any other in a typical school unit on matrices, it is the idea of matrix multiplication. It is hard to find a student that could explain why matrices are multiplied the way they are. Some mathematics teachers can't explain the "why" behind it either. This leads to a frustrating situation for students that are forced to learn something that seems to have no application, something they do but don't understand why, and something that, even if they cared to know the "why," their teacher can't explain. (I know a lot of teachers and curriculum have made great headway here, but still too many students suffer this frustration). Most get something like the following which seems to have no rhyme or reason.

Matrix	Multiplication
$\begin{bmatrix} 92 & 100 & 89 & 80 \\ 72 & 85 & 80 & 75 \\ 88 & 78 & 85 & 92 \end{bmatrix} \times \begin{bmatrix} .4 \\ .15 \\ .25 \\ .2 \end{bmatrix} = \begin{bmatrix} 90 \\ 77 \\ 87 \end{bmatrix}$	Think of turning the first matrix to the right and sideways, multiplying each number by the numbers in the second matrix, and then adding them together. For example, $(92\times.4)+(100\times.15)+(89\times.25)+(80\times.2)=90.05$ $(72\times.4)+(85\times.15)+(80\times.25)+(75\times.2)=76.55$ $(88\times.4)+(78\times.15)+(85\times.25)+(92\times.2)=86.55$

(https://www.shelovesmath.com/algebra/advanced-algebra/matrices-and-solving-systems-with-matrices/)

With what we have done so far we have a couple different ways to make sense of matrix multiplication. First, if we connect a few of our equations, then the matrix multiplication should be definable without the smoke-and-mirrors-I-don't-know-why-just-do-it frustration some of us experienced.

If $\begin{bmatrix} 15 & 1.5 \\ 8 & 3 \end{bmatrix} * \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 10 \\ 16 \end{bmatrix}$ and $a * \begin{bmatrix} 15 \\ 8 \end{bmatrix} + b * \begin{bmatrix} 1.5 \\ 3 \end{bmatrix} = \begin{bmatrix} 10 \\ 16 \end{bmatrix}$ represent the same equation, then the left hand sides should be equal, so $\begin{bmatrix} 15 & 1.5 \\ 8 & 3 \end{bmatrix} * \begin{bmatrix} a \\ b \end{bmatrix}$ is the same as $a * \begin{bmatrix} 15 \\ 8 \end{bmatrix} + b * \begin{bmatrix} 1.5 \\ 3 \end{bmatrix}$. But using what we know about scalar multiplication of vectors we could rewrite the latter as: $\begin{bmatrix} a * 15 \\ a * 8 \end{bmatrix} + \begin{bmatrix} b * 1.5 \\ b * 3 \end{bmatrix}$. Adding the two vectors we get: $\begin{bmatrix} a * 15 + b * 1.5 \\ a * 8 + b * 3 \end{bmatrix}$. This is a similar representation to the process explained in the graphic above (but we have a couple of variables instead of all numbers). The completely general form would be: $\begin{bmatrix} w & x \\ y & z \end{bmatrix} * \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} aw + bx \\ ay + bz \end{bmatrix}$. If this is all students are taught, and none of the in-between. When I was first taught it, this made no sense at all.

Let's use what we have done to make sense of this. For me, the key is in the first two expressions, but let's take care of the easy one's first. From (2) to (3) we are multiplying vectors by a scalar. As we saw earlier, this takes each quantity in the vector packet and multiplies it by the scalar. This corresponds to saying "If we take 'a' servings of PB then we get 'a' times the nutritionals amounts in PB." From (3) to (4) we are adding two vectors. As we saw earlier, adding two vectors corresponds to adding the corresponding values of each vector to get a new vector the same size as the originals. In our food context, if you have a serving of PB and a serving of Bread, then the total number of Fat and Protein will be the Fat from both and the Protein from both.

(1) (2) (3) (4)

$$\begin{bmatrix} 15 & 1.5 \\ 8 & 3 \end{bmatrix} * \begin{bmatrix} a \\ b \end{bmatrix} = a * \begin{bmatrix} 15 \\ 8 \end{bmatrix} + b * \begin{bmatrix} 1.5 \\ 3 \end{bmatrix} = \begin{bmatrix} a * 15 \\ a * 8 \end{bmatrix} + \begin{bmatrix} b * 1.5 \\ b * 3 \end{bmatrix} = \begin{bmatrix} a * 15 + b * 1.5 \\ a * 8 + b * 3 \end{bmatrix}$$

I have pasted (1) and (2) below. The right side, (2), of the equation represents 'a' copies of one vector plus 'b' copies of another. We have called this a linear combination of the two vectors. For example, 'a' servings of PB and 'b' servings of Bread. The left-hand side, (1), is simply a shorthand notation for the right (why we need a shorthand is addressed below), (2). It represents a linear combination of the columns of the 2X2 matrix with 'a' copies of the first column and 'b' copies of the second. This is not the way you may have been taught to think about matrix-vector multiplication. If you have been taught before, it was the process that takes you right from (1) to (4) like we saw in the figure from the website earlier. With the understanding that the entries of the vectors represent the amount (multiple) of each column to add up to each other, the idea of matrix multiplication can be demystified.

$$\begin{bmatrix} 15 & 1.5 \\ 8 & 3 \end{bmatrix} * \begin{bmatrix} a \\ b \end{bmatrix} = a * \begin{bmatrix} 15 \\ 8 \end{bmatrix} + b * \begin{bmatrix} 1.5 \\ 3 \end{bmatrix}$$

This process works for different size vectors and matrices. There are some examples below.

$$\begin{bmatrix} 1 & .3 & .2 \\ .1 & .23 & 1.2 \\ .5 & 2 & 3.1 \end{bmatrix} * \begin{bmatrix} 3 \\ 4.4 \\ 5 \end{bmatrix} = 3 * \begin{bmatrix} 1 \\ .1 \\ .5 \end{bmatrix} + 4.4 * \begin{bmatrix} .3 \\ .23 \\ 2 \end{bmatrix} + 5 * \begin{bmatrix} .2 \\ 1.2 \\ 3.1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 1 & -1 \\ 1 & 1 & 0 & 1 \end{bmatrix} * \begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \end{bmatrix} = 1 * \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 0 * \begin{bmatrix} 0 \\ 1 \end{bmatrix} + (-1) * \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 1 * \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

There is no limit to the size that a matrix or vector can be. Some applications, like genetic research with DNA data, deal with matrices that are thousands of rows and thousands of columns.

Is there a limit to the sizes of matrices and vectors that can be multiplied together? To explore this question, we might try swapping the vectors and matrices in the examples above. What should these products equal?

$$\begin{bmatrix} 1 & .3 & .2 \\ .1 & .23 & 1.2 \\ .5 & 2 & 3.1 \end{bmatrix} * \begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \end{bmatrix} =? \qquad \text{Or} \qquad \begin{bmatrix} 1 & 0 & 1 & -1 \\ 1 & 1 & 0 & 1 \end{bmatrix} * \begin{bmatrix} 3 \\ 4.4 \\ 5 \end{bmatrix} =?$$

Well in either case there seems to be a problem. Let's try rewriting the first one as a linear combination.

$$\begin{bmatrix} 1 & .3 & .2 \\ .1 & .23 & 1.2 \\ .5 & 2 & 3.1 \end{bmatrix} * \begin{bmatrix} 1 \\ 0 \\ -1 \\ 1 \end{bmatrix} = 1 * \begin{bmatrix} 1 \\ .1 \\ .5 \end{bmatrix} + 0 * \begin{bmatrix} .3 \\ .23 \\ 2 \end{bmatrix} + (-1) * \begin{bmatrix} .2 \\ 1.2 \\ 3.1 \end{bmatrix} + 1 * \begin{bmatrix} ? \\ ? \\ ? \end{bmatrix}$$

The left hand side says "Take 1 of the first column, 0 of the second column, -1 of the third column, and 1 of the fourth column, then add them up." BUT THERE IS NO FOURTH COLUMN! This product is not defined because there is no way to make sense of how the three columns should match up with the four rows of the vector.

Similarly, this product: $\begin{bmatrix} 1 & 0 & 1 & -1 \\ 1 & 1 & 0 & 1 \end{bmatrix} * \begin{bmatrix} 3 \\ 4.4 \\ 5 \end{bmatrix}$ is also undefined because we can't match up the columns of the matrix to the rows of the vector.

Why Do We Need a Shorthand? If you are wondering why we would want a shorthand notation (after all, we can just right down the right-hand side every time). Well here are two quick reasons why we might want one. First, it comes in really handy to rename matrices and vectors as variables because it saves a lot of writing and can make it easier to see what operations are being performed. For example let $A = \begin{bmatrix} 15 & 1.5 \\ 8 & 3 \end{bmatrix}$ and $x = \begin{bmatrix} a \\ b \end{bmatrix}$, then the left hand side of the equation above, $\begin{bmatrix} 15 & 1.5 \\ 8 & 3 \end{bmatrix} * \begin{bmatrix} a \\ b \end{bmatrix}$, can simply be written as $A * x$ (or just Ax). This makes it much nicer to do algebra with matrices and vectors than writing out all of the entries all of the time (especially when you don't know what the entries actually are!)

The second reason a shorthand notation comes in handy is that in a lot of contexts the quantities of 'a' and 'b' that we put into a vector is already easily represented as a packet. It might be a set of coordinates, a packet of values in a data set, or any of the other things that were suggested earlier that could be packets. There is no real reason why 'a' and 'b' are any different than the packet of '15' and '8' associated with Fat and Protein of PB, and so we ought to be able to represent them the same way.

Why Multiplication? Some of you may be wondering why we should even call this process of linearly combining vectors (or packets) "multiplication." Looking at the right side of the equation above you can see there is an addition happening, so why call it "multiplication"? Well, isn't multiplication repeated addition? 5*4 = 4+4+4+4+4 and $3 * \begin{bmatrix} 15 \\ 8 \end{bmatrix} = \begin{bmatrix} 15 \\ 8 \end{bmatrix} + \begin{bmatrix} 15 \\ 8 \end{bmatrix} + \begin{bmatrix} 15 \\ 8 \end{bmatrix}$. The kind of multiplication that happens when we have packets is a repeated addition with PACKETS OF DIFFERENT SIZES! Apple +Apple + Apple+Orange +Orange=3*Apple+2*Orange. Can you see the linear combination? Some Apples plus some Oranges? 4+4+4+4+4-2-2-2=5*4+3*(-2) a linear combination of 4's and -2's. $\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 4 \\ 5 \end{bmatrix} + \begin{bmatrix} 4 \\ 5 \end{bmatrix} + \begin{bmatrix} 4 \\ 5 \end{bmatrix} = 4 * \begin{bmatrix} 1 \\ 2 \end{bmatrix} + 3 * \begin{bmatrix} 4 \\ 5 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 2 & 5 \end{bmatrix} * \begin{bmatrix} 4 \\ 3 \end{bmatrix}$, a linear combination of the two vectors $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 4 \\ 5 \end{bmatrix}$. Viewed in this way, linear combinations of different packets is one way of generalizing multiplication.

Another reason why it makes sense to call this process a multiplication comes from the units of the quantities. I have labeled the units in the matrix-vector product below based on our PB and Bread problem. Each of the entries in the matrix are grams per serving (for example, the 15 represents the 15 grams of fat per serving of PB). The vector multiplied by the matrix shows the number of servings (2/9 represents 2/9ths servings of PB). Finally, the resulting product, $\begin{bmatrix} 10 \\ 16 \end{bmatrix}$, shows the resulting grams of each nutrient (10 represents 10 grams of Fat). For the units to work out, grams/serving needs to be multiplied by # of servings so the units will result in grams. This argument gives us further credence to think of this process as a multiplication.

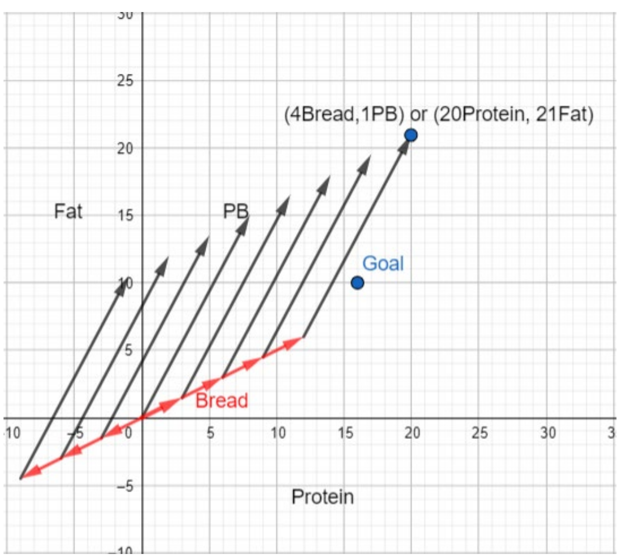
grams/serving * Servings = grams

$$\begin{bmatrix} 15 & 1.5 \\ 8 & 3 \end{bmatrix} * \begin{bmatrix} 2/9 \\ 160/33 \end{bmatrix} = \begin{bmatrix} 10 \\ 16 \end{bmatrix}$$

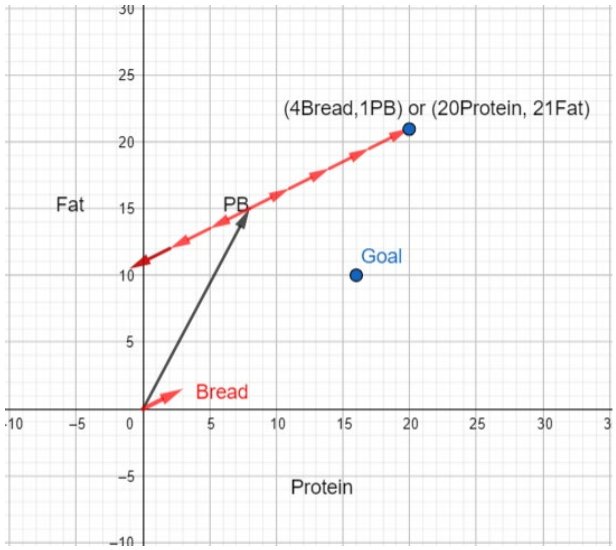
Was it just luck? Let's get back to our problem at hand, PB and Bread and Fat and Protein. We were able to find a linear combination of PB and Bread that allowed us to reach our goal, but was that just luck? Did we just happen to pick a goal that was a possibility with the PB and Bread packets? Or can we reach any Fat and Protein goal with just these two foods? What if we didn't want to increase our fat at all and just increase our protein by 16 grams. Is that possible with PB and Bread? Or would we need to add in another food (or multiple other foods)?

What I am wondering about can be said another way. Where could all of the linear combinations of servings of PB and Bread get us on the Fat-Protein plane?

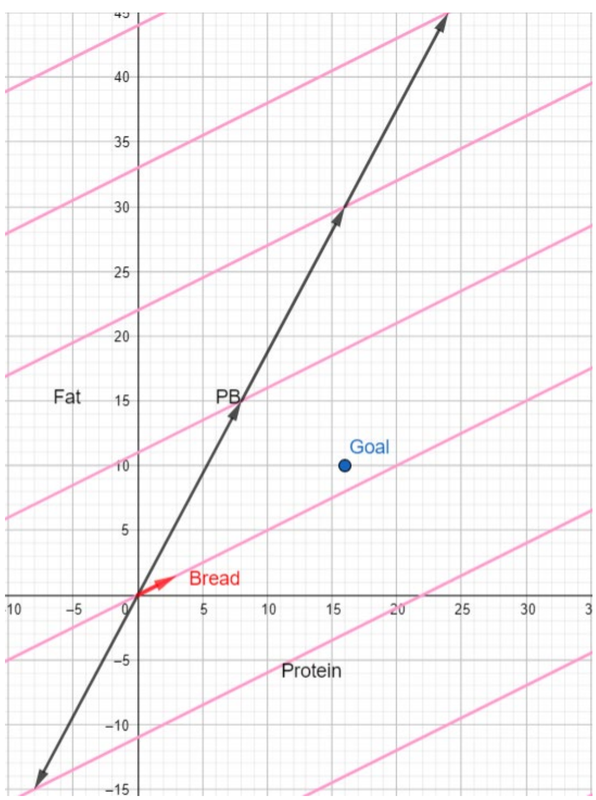
Let's explore a little bit by looking at a few different combinations. We already know that as we change the number of servings of PB, we end up on a line, the line that the PB vector falls on. Similarly with Bread, if we only adjust the amount of bread in our diet we can change the Fat-Protein content along the line that the Bread vector falls on. As we take combinations of these we start to see a pattern. Let's look at (0 Bread, 1 PB), (1 Bread, 1 PB), (2 Bread, 1 PB), etc. Let's also include combinations where we subtract bread servings from our diet as well, for example: (-1 Bread, 1 PB), (-2 Bread, 1 PB), etc. These combinations are similar to the ones where we just changed the number of Bread servings, but now we are always including 1 serving of PB. The graph below shows the Points on the Fat-Protein plane where we end up for a few of these combinations. (I have represented them by going along the Bread vectors first and then going up one PB). I have labeled one point that represents the combination of 4 Breads and 1 PB, or from the cartesian coordinates (20,21), which represents 20 grams of Protein and 21 grams of Fat. As you look at the graph below, what do you notice?



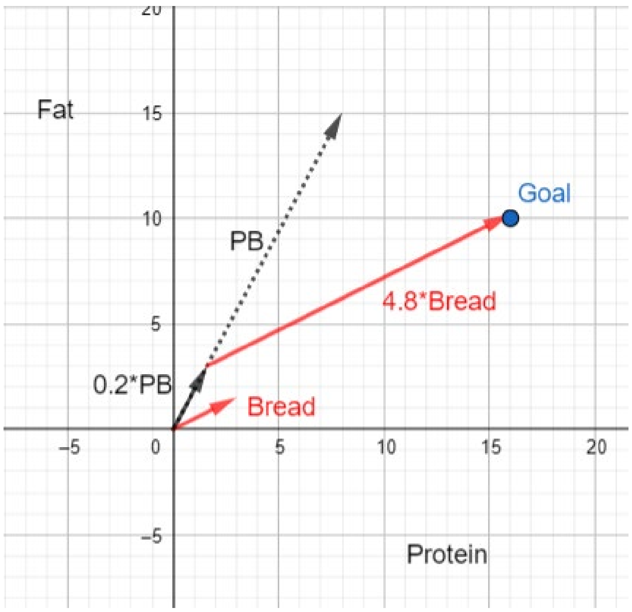
One thing that pops out at me is that the resulting combinations (represented by the ends of the black arrows) all lie on the same line. Additionally, that line is parallel to the line that the Bread vector falls on. This might be easier to see if I represent it on the graph another way, going up one PB, then along different multiples of the Bread vector. That graph lies below.



So adding one serving of PB to all possible Bread servings gets us all values on another line that is parallel to the original Bread-only line. What would we get if we added 2 PB's to all possible servings of Bread? Another line parallel to the Bread-only line but shifted in the direction of 2 PB arrows end to end. What if we decreased our diet by one PB and considered all possible Bread servings? We would get another line parallel to the Bread-only line but shifted down one PB vector. These lines (and a few others) are illustrated in the figure below.

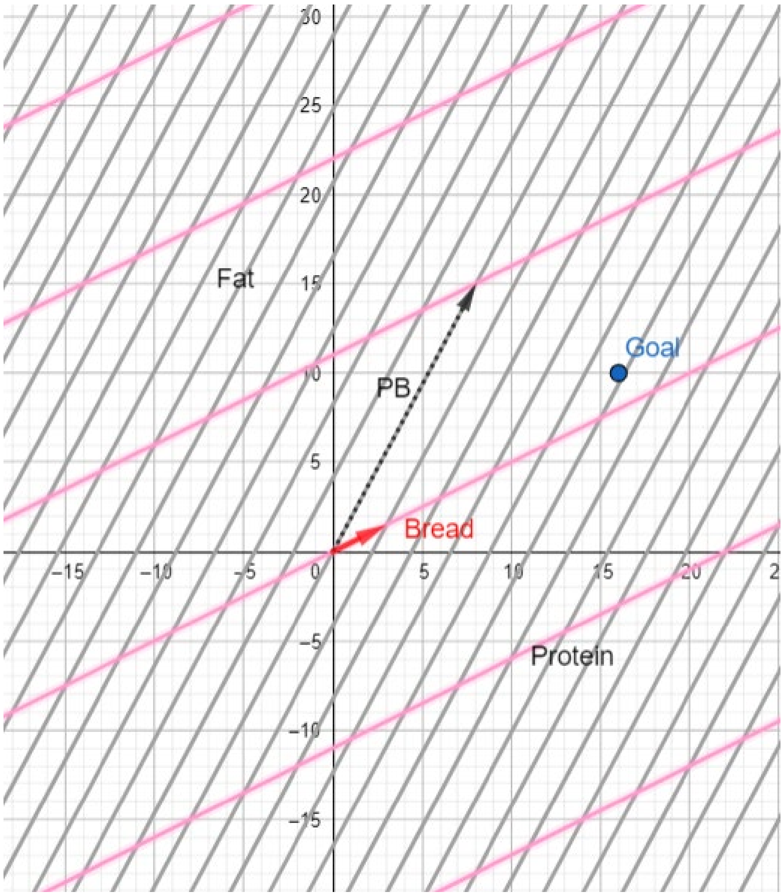


We can get any combination of Fat-Protein that lie on the PB line (line of black arrows) or on the pink lines, but what about in between the pink lines? Well since we can take partial servings of PB it seems like we should be able to shift the lines to any level along the PB-only line, then we can stretch as far as we want along the lines parallel to the Bread-only lines. This argument makes it seem like we can reach any point on the Fat-Protein plane. One process to find the amount of PB and Bread needed to find any combination of Fat-Protein, would be to shift the Bread-only line until it goes through the desired point, then measure how many PB servings you shifted the Bread-only line and count how many Bread servings you need to go from the PB-only line to the point. This is close to what we did to find an approximate solution to our problem graphically. I have posted that figure again below.



If we think of all possible PB servings with just one Bread, or two Breads, etc. then we generate another set of lines (gray lines in the figure below). These gray and pink lines illustrate a kind of grid that works similarly to our x-y cartesian grid. Although we only draw lines periodically, we can still find coordinates of points that don't fall on our drawn lines. Additionally we can find "coordinates" of points in terms of serving of PB and servings of Bread as well as grams of Fat and grams of Protein.

Give it a try! Using the gray and pink lines can you find the combination of PB and Bread needed to lower Protein by 14 grams and increase fat by 15 grams? How about just raising the Protein by 22 grams but keeping Fat the same?



Spanning (not Spamming): We found that we can get to any point on the plane with the linear combination of the vectors $\begin{bmatrix}15\\8\end{bmatrix}$ and $\begin{bmatrix}1.5\\3\end{bmatrix}$. In food talk this translates to, "We found that we can get any combination of Fat and Protein by independently adjusting the number of servings of PB and Bread in our diet." The answer to the question of "Where do linear combinations of vectors allow me to reach?" has a name. It is called the *span* of the vectors. We could say that "the PB and Bread vectors span the Fat-Protein plane." Since mathematicians usually aren't studying a particular context, they would just keep the units out of it and would say "The vectors $\begin{bmatrix}15\\8\end{bmatrix}$ and $\begin{bmatrix}1.5\\3\end{bmatrix}$ span the plane."

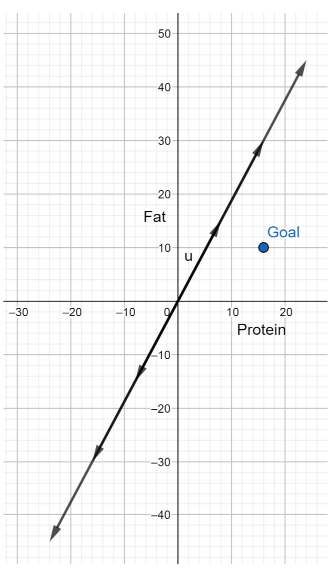
Unlike many names in mathematics, this name is actually appropriate. (Logarithm is the classic case of unhelpful names in mathematics, although the Latin meaning can help some. It would have been better to call it something like the "lost exponent" instead of "logarithms.") Sorry for the side note. Back to *span*. We talk about the "span of a bridge" or a "wing span" which are both connected to reaching farther and going new places. The span of vectors is all the places you can reach to with combinations of those vectors. Sometimes a set of vectors can get you anywhere you may want to go (the space of Fat and Protein for instance) but sometimes the set of vectors can't get to all possible places in the space you care about (we will see this soon). Sometimes two vectors are as good as three (or four or five or . . .) because their span is the same. If we added Smuckers Raspberry Jam to the mix (0 Fat and .1 gram of Protein per serving) with the PB and Bread, it doesn't help us reach any more places than just the PB and Bread. Why? Well first, we can already get everywhere in the Fat-Protein plane with PB and Bread so there is nowhere else to go (sure, we can make a sweeter sandwich but mathematically it isn't needed to reach any Fat-Protein goal). Second, PB, Bread, and Jam vectors are not linearly independent, recall that this means that at least one of the vectors is a linear combination of the others. We can take any one of them out and still be just fine (mathematically that is, not yummily) at achieving our Fat-Protein goals.

Basis: Mathematicians have a name for a set of vectors that span the important space for a particular context (whether that is Fat-Protein plane, the general 2-dimensional plane, 3-dimensional space, or more complex entities like the space of all sound waves). It is called a ***basis*** set of vectors. A basis set for the plane means you have enough vectors that will span the plane. We can have more than enough, so a basis refers to the smallest set of vectors that span a given space. The smallest set is a set that is linearly independent. A minimal set of spanning vectors (a basis) is enough to model the space but also avoids unnecessary complexity by including any unnecessary vectors. Our PB and Bread vectors span the Fat-Protein plane, but if we remove either one of these vectors, the set (which will just be one vector) won't span the entire Fat-Protein plane. Therefore the PB and Bread vectors are a basis for the Fat-Protein plane. We need them both, but we don't need any more. Certainly there are 2 other foods that could also serve as a basis for the Fat-Protein plane (minimal basis are not necessarily unique), but PB and Bread are one of them.

Chapter 4

Just Do the Best You Can

What happens when our goal is not in the span of our vectors? We actually have already bumped into this problem. When we first set our goal of 16 g of Protein and 10 grams of Fat, we only had PB to work with. The span of the PB vector didn't include the goal, as you can see from the image below. The span of the PB vector only includes multiples of the PB vector so it can only reach places on the line on which the PB vector falls. The span of the PB vector is a one-dimensional line (all lines are one dimensional).



Recall that there were three strategies we could try to reach our goal when we only had PB:

Vector Match Strategy: Find a different food with the ratio of protein to fat that matches the ratio of protein and fat in your goal.

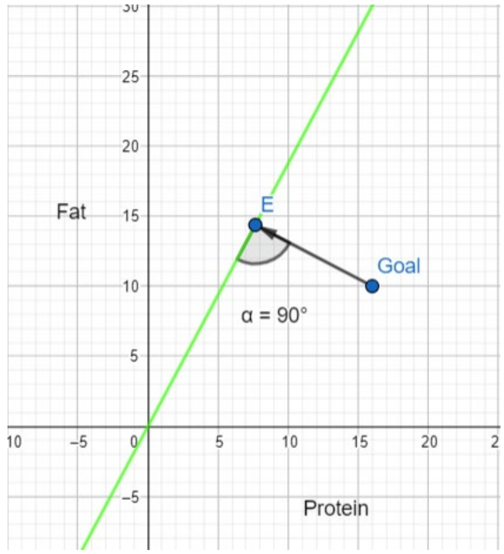
Combination Strategy: Find another food (or other foods) in addition to peanut butter that you are already eating and are willing to change the amount in your diet, but has a different ratio of protein to fat than peanut butter. The hope is that you can eat some amount of one and some amount of another to reach your goal.

Best Approximation Strategy: Change the amount of PB in your diet to get as close as possible to your goal, but give up on meeting the exact goal.

We already explored the **combination strategy** by including bread as a food we could add to, or remove from, our diet. We won't delve into the **vector match strategy**. I mentioned earlier that the **best approximation strategy** might be the easiest but trying to decide how much to eat to get you "closest" to your goal actually just raises another math problem. We now get to explore this math problem. The best approximation strategy is a very viable strategy for many linear algebra applications and is at the basis of such important technology as being able to quickly pass voice data across cellular networks or predicting outcomes from data.

Let's begin by looking at our image of the graph again that shows the span of the PB vector and our goal (shown below on the left). We can only get values of Fat and Protein that lie on the PB-only line. Which point on that line is the closest to the goal? There may be a few defensible answers. We could find the point on the line that matches our Protein goal but is off on the Fat goal (EqualProtein). That would get us exact on one of our nutrients. Of course there is no reason we have to match Protein exactly. We could match Fat exactly and be off on Protein (EqualFat). We could also go out on the line until we eat the same number of combined grams (16 grams of Protein +10 grams of Fat=26 total grams) as the goal (SameTotalAmount26gms), which looks like a better option than the first two. However, looking at the image to the left, it looks like there might be even a better choice.

If we take "closest" literally we can try to use what we know about distances between a point and a line to find another option. This allows us to use geometry to find the closest point. From geometry you may recall that the closest point on line to a point that is not on the line is found by going toward the line on a perpendicular path to the line. The image on the right shows a vector pointing from the Goal point to the PB line (or the span of the PB vector). How can we find the coordinates of point E? [Actually, we don't need the coordinates, we need the number of PB vectors (servings of PB) that would fall on that Protein-Fat point, but once we get the coordinates, it is pretty easy to find the number of servings.]

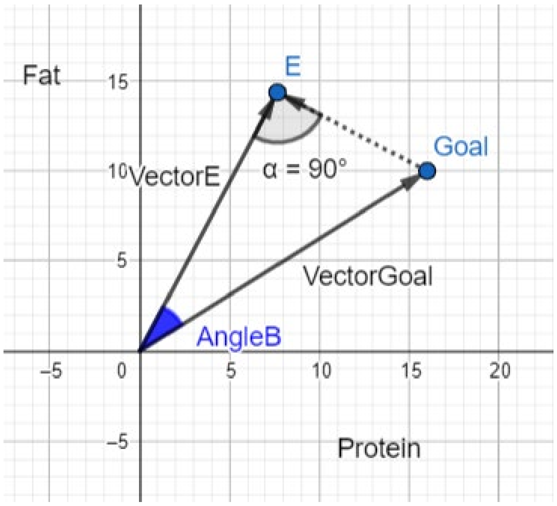


Finding the coordinates of point E or the number of servings of PB can be done in a couple of ways that use tools you already know. We can write down the equation of two lines (the PB line and the line that goes through the Goal point that is perpendicular to the PB line) and then solve a system of equations. Yes, the system shows up again! It will continue to come up, and we will use it to continue to make sense of working with packets.

The equation of the PB line is $y = \frac{15}{8}x$, and the equation of the line perpendicular to the PB line that goes through Goal is $y - 10 = \frac{-8}{15}(x - 16)$. For the second equation I use the point-slope formula and the fact that the slope of a perpendicular line is the multiplicative inverse of the slope of the original with the sign changed. You are free to solve this system of equations to find the coordinates of E, but I am going to move on to another method that is . . . harder.

Well, maybe I should call it 'more advanced.' I know what you are thinking, "But why do we need another way!?" We don't need another method to solve our PB problem. But doing it another way is going to get us to some magic. The beginning of a linear algebra class does seem to be a list that could be grouped in a chapter called "how to do things you already know how to do, but in a harder way." After this problem, we will be on the verge of flipping it around and starting the list we could call "How to do things (and understand things) that you already knew, but in an easier way."

Let's use some trigonometry. If we think about the graph differently it becomes a trigonometry problem. Below is an image with two vectors drawn: The vector that points to E and the vector that points to Goal. I kept a third vector in the image just to help remind you that we are working with a right triangle situation. We want to know the length of the vector on the PB line that points to E (or the coordinates of E, either one is sufficient for us to solve the PB-Fat-Protein problem). If we know the angle between the two vectors (VectorE and VectorGoal) then the length of the vector that points to E can be found using basic trigonometry.



The Cos(AngleB) is the ratio of the length of Vector E and the length of VectorGoal. In equation form:

$$Cos(AngleB) = \frac{length\ of\ VectorE}{length\ of\ VectorGoal}$$

If we want the length of VectorE, then we can solve the equation to get:

$$length\ of\ VectorE = Cos(AngleB) * length\ of\ VectorGoal$$

Well, this is now pretty straight forward. . . . . . If we knew what the value of AngleB was. But we do not. It isn't the end of the world, because we can find the angel by doing more trig. We can use the coordinates of E and Goal, find the length between them, then use inverse Sin to find the AngleB. But there is a different way. A magic way.

The Amazing Presto pulls a dot product out of his hat: Remember that dot product? The linear combination of a set of quantities, or it can be thought of as a special way to multiply two vectors (a vector-vector product). As a quick review, the dot product works like this:

$$\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \bullet \begin{bmatrix} 3 \\ 0 \\ 4 \end{bmatrix} = sum \begin{pmatrix} 1*3 \\ 2*0 \\ (-1)*4 \end{pmatrix} = 1*3 + 2*0 + (-1)*4 = -1$$

Here is the magic thing about the dot product. If the vectors in the product have a length of 1, then the dot product equals . . . . I know you are on the edge of your seat . . . You are thirsty but you won't go get a drink because you are so anxious to know . . . Your phone is ringing but you don't care . . . Ok, here it is . . . the dot product equals the cosine of the angle between the two vectors! (So long as the vectors have a length of one).

Yes, I know. I couldn't believe it either. How can this process that uses such elementary operations, multiplication and addition, a process that can be taught to third graders, calculate such a technical property of two vectors in a multi-dimensional space? It sure seems like magic to me.

This means that we can find the number of servings of PB that gets closest to our goal without using trig and without solving a system of equations. But before we get to that, let's test out the dot product and convince ourselves that it has this property.

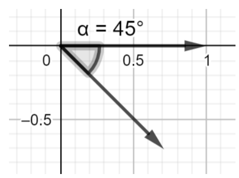
Note: Just a warning, we will spend a bit of time exploring the dot product. It might seem longer than is needed, but the work we do here has an alternative purpose as well, it will allow us to build up some intuition about working *with vectors in 2 and 3 and even higher dimensions as well as better understanding projections. Being able to visualize the spatial relationships of vectors, lines, planes, spanning sets, etc is a powerful tool working with vectors when represented geometrically. It will also give us a good chance to introduce some notation to save some writing.*

Why don't we start with pretty simple vectors: $\begin{bmatrix}1\\0\end{bmatrix}$, and $\begin{bmatrix}0\\1\end{bmatrix}$. These are the vectors that point to the points (1,0) and (0,1) on the coordinate plane. They each lie on the x- or y-axis so we already know that the angle between them is 90 degrees (or $\frac{\pi}{2}$ radians). The cosine of 90 degrees is 0. What do we get with the dot product?

$\begin{bmatrix}1\\0\end{bmatrix}\bullet\begin{bmatrix}0\\1\end{bmatrix} = 1*0+0*1 = 0+0 = 0$

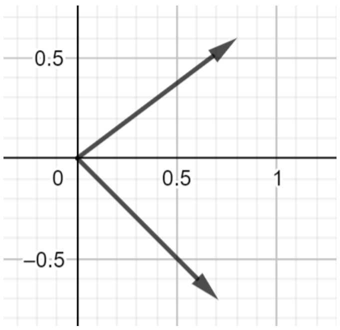
Wow. Pretty slick.

How about something not so straight forward? Maybe $\begin{bmatrix}\frac{\sqrt{2}}{2}\\ \frac{-\sqrt{2}}{2}\end{bmatrix}$ and $\begin{bmatrix}1\\0\end{bmatrix}$? These have an angle of 45 degrees (as illustrated below) so the cosine of the angle between these two vectors should be $\frac{\sqrt{2}}{2}$.



$\begin{bmatrix}\frac{\sqrt{2}}{2}\\ \frac{-\sqrt{2}}{2}\end{bmatrix}\bullet\begin{bmatrix}1\\0\end{bmatrix}=\frac{\sqrt{2}}{2}*1+\frac{-\sqrt{2}}{2}*0=\frac{\sqrt{2}}{2}$. Wow, the cosine of 45 degrees.

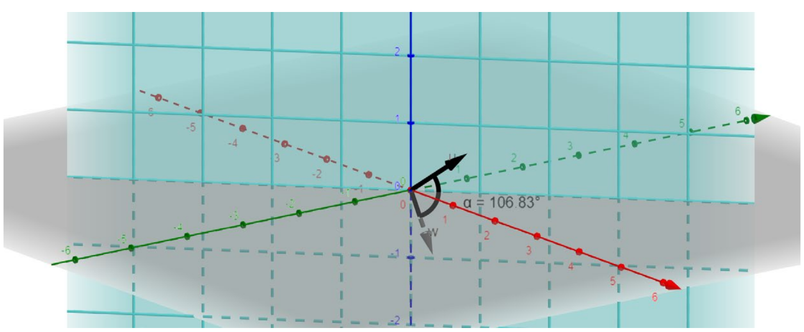
How about something even more extreme? $\begin{bmatrix}\frac{\sqrt{2}}{2}\\ \frac{-\sqrt{2}}{2}\end{bmatrix}$, and $\begin{bmatrix}.8\\.6\end{bmatrix}$. Neither of these are super crazy, and the second one has a unit length of 1 because it is a scaled 3-4-5 right triangle (divide each value by 5 to get .6-.8-1). I have graphed these two vectors in the image below. They look almost perpendicular, but we can tell that the angle is a little less than 90 degrees because the vector in the first quadrant isn't rotated enough to hit the point (.5, .5). We don't know the angle, but we could find it using one of the calculator's inverse trig functions, then we could also find the cosine of the angle using the calculator. But the dot product gives us an easier way. We may still use a calculator to find an approximate decimal form of the cosine of the angle, but it is much easier than the multi-step trigonometry approach. We can actually get quite close without the calculator since using .7 as an approximate value for $\frac{\sqrt{2}}{2}$.

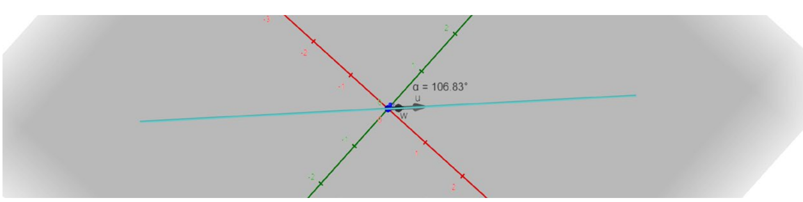


$\begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{-\sqrt{2}}{2} \end{bmatrix} \bullet \begin{bmatrix} .8 \\ .6 \end{bmatrix}$=.8*$\frac{\sqrt{2}}{2} + .6 * \frac{-\sqrt{2}}{2} \approx .8 * .7 - .6 * .7 = .2 * .7 = .14$. Not far from the calculator estimate of .1414214.

Now we can find the angle using cos^{-1}(.1414214), or some prefer arccos(.1414214) which the calculator gives as 81.87 degrees. That seems quite reasonable from the picture above.

Although our examples have only been in the plane, the same process works in higher dimensions. Angles are measured in the plane in which both vectors lie. (Three non-collinear points determine a plane). Consider these two vectors, $\begin{bmatrix} .22 \\ .22 \\ -.95 \end{bmatrix}$, and $\begin{bmatrix} \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \end{bmatrix}$. The angle between these two vectors is about 106 degrees. I have represented the vectors in the image blow. The x-axis is red, y is green, and z is blue. The x-y plane is shaded gray. The light blue plane is the plane the cuts through both vectors. Dashed lines are used to represent position behind, or under, planes. A subsequent image below the first shows a top down view (perpendicular to the x-y plane) that shows the two vectors lining up with the light blue plane.





What does the dot product give us in this case? It should be a small (closer to 0 than -1) negative number since we are just a little (about 15 degrees) more than 90 degrees. Recall that cosine begins entering negative values at 90 degrees.

$$\begin{bmatrix} .22 \\ .22 \\ -.95 \end{bmatrix} \bullet \begin{bmatrix} \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \end{bmatrix} = .22 * \frac{\sqrt{3}}{3} + .22 * \frac{\sqrt{3}}{3} + (-.95) * \frac{\sqrt{3}}{3} = -.41 * \frac{\sqrt{3}}{3} \approx -.2895$$

We won't do an example but the dot product still works in more than three dimensions (when the vectors have four or more values). It might be crazy hard to visualize a plane that slices through 2-20 dimensional vectors, but there is still one and we can still us the dot product to find the cosine of the angle between them.

Why does the dot product calculate the cosine of the angle between the vectors?

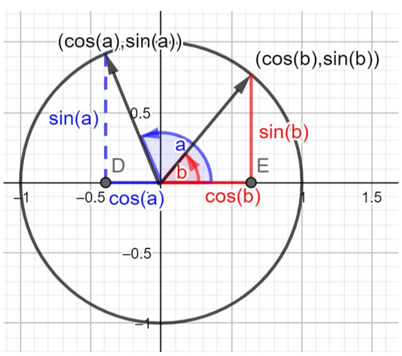
I will give a quick explanation that is valid for the plane. It is based on the unit circle and a trigonometric formula. The first connection to make is that any unit vector that starts at the origin points to a position on the unit circle. Second, the coordinates of a point on the unit circle are $\cos(\theta)$, for the x-coordinate, and $\sin(\theta)$, for the y-coordinate. If we take two unit vectors then we can write them as $\begin{bmatrix} Cos(a) \\ Sin(a) \end{bmatrix}$ and $\begin{bmatrix} Cos(b) \\ Sin(b) \end{bmatrix}$. The dot product of these vectors is:

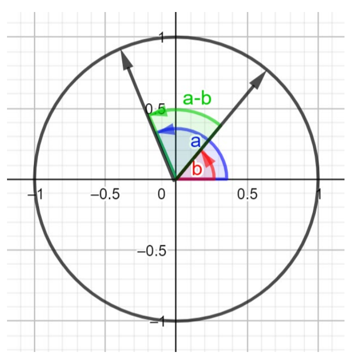
$$\begin{bmatrix} cos(a) \\ sin(a) \end{bmatrix} \bullet \begin{bmatrix} cos(b) \\ sin(b) \end{bmatrix} = cos(a) * cos(b) + sin(a) * sin(b)$$

But the right side is also the formula for the cosine of the difference of two angles (one of those formulas that students, and even professors, don't always have in their memory so don't feel bad if you don't remember it):

$$\cos{(a-b)} = cos(a) * cos(b) + sin(a) * sin(b)$$

If we represent our two vectors on the coordinate plane, then we can see that the difference between angle a and angle b is exactly the angle between the two vectors.



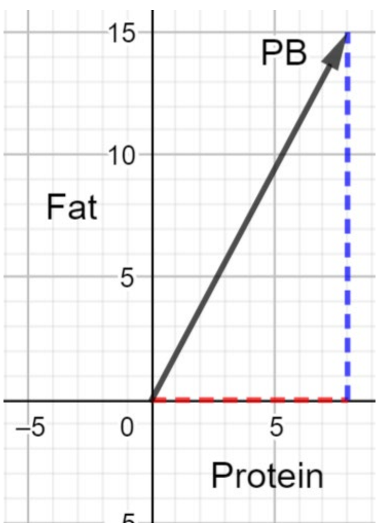


Properties of the Dot Product:

[Maybe paste in here and quickly discuss the algebraic properties that the dot product has that we can use to simplify expressions and solve equations].

But our PB and Goal vectors don't have a length of 1! So we know how to use a dot product to find the cosine of the angle between two vectors AS LONG AS they both have a length of one. But our PB and Goal vectors don't have lengths of 1! Are we still stuck? No, we can just scale them to a length of 1, and use the dot product on the scaled versions of the vectors.

Finding the length of a vector: The PB vector, with 8 grams of Protein per serving and 15 grams of Fat per serving, which we represent as the vector $\begin{bmatrix} 8 \\ 15 \end{bmatrix}$needs to be scaled down to have a length of 1. How long is it right now? Well if we draw the segments represented as the blue and red dashed lines in the figure below, we can see that the PB vector is the hypotenuse of a right triangle. Moreover, the length of the two dashed lines are just the values of the PB vector. The red dashed line shows the vector goes over 8 in the positive Protein direction and the blue dashed line shows the PB vector goes up 15 in the positive Fat direction. We can use the Pythagorean Theorem to find the length of the PB vector: $Length\ of\ PB = \sqrt{8^2 + 15^2} = \sqrt{289} = 17$. (Wow! I am so lucky. I don't know how I happened to pick a food [peanut butter] and two nutrients in that food [fat and protein] that happened to make a vector with a length that is a whole number.) So the PB vector has a length of 17. How do we use this to make a vector that points in the same direction as PB but only has a length of one? Simple. We use scalar multiplication. Scalar multiplication will stretch or shrink a vector, like we saw earlier. The reason that it is called "scalar" is because it can be used to change the scale of the object, larger or smaller. We need to scale PB so it is 1/17 the length, so we can multiply $\begin{bmatrix} 8 \\ 15 \end{bmatrix}$by 1/17: $\frac{1}{17} * \begin{bmatrix} 8 \\ 15 \end{bmatrix} = \begin{bmatrix} \frac{1}{17} * 8 \\ \frac{1}{17} * 15 \end{bmatrix}$ = $\begin{bmatrix} \frac{8}{17} \\ \frac{15}{17} \end{bmatrix}$.



If we take a step back we can take what we have done and write down a to turn any vector into a unit vector (a vector of length one). For any vector $\begin{bmatrix} a \\ b \end{bmatrix}$, the unit vector pointing in the same direction is $\frac{1}{length\ of\ \begin{bmatrix} a \\ b \end{bmatrix}} * \begin{bmatrix} a \\ b \end{bmatrix}$. We also learned that the $length\ of\ \begin{bmatrix} a \\ b \end{bmatrix}$can be found using the Pythagorean Theorem, and is equal to $\sqrt{a^2 + b^2}$. The Pythagorean Theorem generalizes to higher dimensions:

The length of $\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_{n-1} \\ a_n \end{bmatrix} = \sqrt{{a_1}^2 + {a_2}^2 + {a_3}^2 + \cdots + {a_{n-1}}^2 + {a_n}^2}$

So the process of finding a unit vector can be written as:

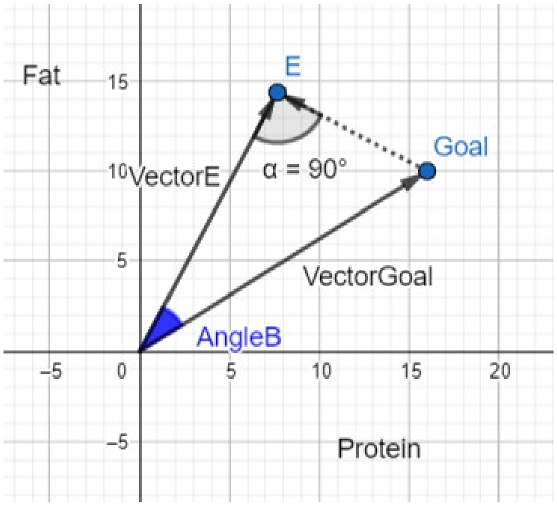
$$\frac{1}{\sqrt{{a_1}^2+{a_2}^2+{a_3}^2+\cdots+{a_{n-1}}^2+{a_n}^2}} * \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_{n-1} \\ a_n \end{bmatrix} \quad \text{or} \quad \frac{1}{Length\ of \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_{n-1} \\ a_n \end{bmatrix}} * \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_{n-1} \\ a_n \end{bmatrix}$$

Let's Get Lazy: You may have noticed that writing down big vectors, or having to write out in words "length of" is long and cumbersome. There are ways to get around this. When I was taking a lot of mathematics classes I hated writing down a lot for my homework problems. Sometimes the problems were computational, sometimes they were proofs, but either way it was a pain to write down words like "diagonalizable" or "homeomorphic" 15 times in the same proof or write down the same expression 15 times during a long computational process. I started to create notation to make things easier and keep my hands from cramping up so much. Mathematicians do this all the time. It saves space and time.

The first thing mathematicians do to save space and time is to just name objects and then use the name. For example, we could shorten up: $\frac{1}{Length\ of \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_{n-1} \\ a_n \end{bmatrix}} * \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_{n-1} \\ a_n \end{bmatrix}$ to $\frac{1}{Length\ of\ u} * u$. If we first said "Let u be any vector." We could have used this trick earlier when talking about the unit vector in the direction of PB, and simply written $\frac{1}{Length\ of\ PB} * PB$.

But what about always writing "length of"? Is there a way to shorten that? It isn't an object so naming it doesn't help. A computer programmer would just define a function called 'length()' that takes in a vector and spits out the length. Mathematicians, however, like to be even more efficient. They use a notation related to absolute value, $|x|$, to signify the length of a vector. After all, absolute value does give the length of a vector on the number line. For general vectors, not just one-dimensional ones, the length of a vector is shown by using double bars instead of single, $\|x\|$. Now with that notation we can shorten up our expressions from above even more: $\frac{1}{\|u\|} * u$, or even just, $\frac{u}{\|u\|}$. To show the process of creating a unit vector in the direction PB, or just as set of symbols that can be used as a name for that unit vector, we can write $\frac{PB}{\|PB\|}$.

Are we there yet? Now we finally have the tools to solve the best approximation problem without having to do a ton of trig and computer calculations by using the magic property of the dot product. Recall that we were using a right triangle to solve for the length of a vector that represents the closest we can get to our goal with PB. I have put the image we were working from below. We had figured out that the length of the vector pointing to point E could be found by:



$$length\ of\ VectorE = Cos(AngleB) * length\ of\ VectorGoal$$

We now know that we can rewrite this equation to:

$$\|VectorE\| = Cos(AngleB) * \|VectorGoal\|$$

Now would be a good time to stop reading and try to apply what we have done so far and solve for $\|VectorE\|$ yourself. Sure, I am going to walk through it, but your learning will be much deeper and last longer if you try to do it yourself. Even if you don't do it correctly or get stuck, spending a few minutes or even a lot more than a few minutes will pay dividends.

Let's start with that Cos(AngleB) term. We know that the cosine of the angle between two unit vectors can be found with the dot product. AngleB is the angle between VectorE and VectorGoal, or the vector PB and VectorGoal (VectorE and Vector PB are on the same line). Since we don't know the coordinates of E, we can't use it in the dot product, so we will use PB. To use the dot product and get the cosine of the angel out, we need the two vectors to be unit vectors. We can find these unit vectors by dividing by their length, as we discussed above. So we can find Cos(AngleB):

$$Cos(AngleB) = \frac{PB}{\|PB\|} \bullet \frac{VectorGoal}{\|VectorGoal\|}$$

Substituting in to our previous equation gets us:

$$\|VectorE\| = \frac{PB}{\|PB\|} \bullet \frac{VectorGoal}{\|VectorGoal\|} * \|VectorGoal\|$$

We can simplify this by using algebraic properties:

$$\|VectorE\| = \frac{PB \bullet VectorGoal}{\|PB\|}$$

Now finding $\|VectorE\|$is a practice in computation (or typing it into a computer/calculator) where PB is $\begin{bmatrix} 8 \\ 15 \end{bmatrix}$, VectorGoal is $\begin{bmatrix} 16 \\ 10 \end{bmatrix}$, and we can find the length of PB through the Pythagorean theorem (length of $\begin{bmatrix} a \\ b \end{bmatrix}$is $\sqrt{a^2 + b^2}$).

According to my calculations I found that $\|VectorE\|$is about 16.353.

Ok, now what? We found the length of VectorE but we really want to know the number of peanut butter servings to get us closest to our goal. The length of the PB vector, or $\|PB\|$, was 17. We just need to shrink it a little bit, or add just a little less than one serving of PB to our daily intake. (Of course for practical purposes in this particular problem about PB and Fat and Protein, rounding to one serving would be just fine). Let's make sure we know how to find the exact value though for applications that require more precision.

We need to shrink a vector of length 17 to have a length of 16.353. How do we shrink (or stretch) a vector? Or how do we 'scale' a vector? Scalar multiplication. What scalar can we multiply PB by to change the length from 17 to 16.353? Some of you may have realized that the scalar should be $\frac{16.353}{17}$. One way to find this is first to think of shrinking the PB vector down to a unit vector (multiply by $\frac{1}{17}$) then by stretching it out to a length of 16.353 (multiply by 16.535). This is, of course, the same as just multiplying PB by $\frac{16.353}{17} \approx .96$. There are other ways to find the scalar (or scale factor). Some of you may have thought about using an equation like the one below and solving for k in a single step.

$$k * 17 = 16.353$$

The equation above comes from the following:

$k * \|PB\| = \|VectorE\| \rightarrow k = \frac{\|VectorE\|}{\|PB\|}$

And we can generalize it to find the scalar that will stretch/shrink one vector, u, to match the length of another, v, see below. BE CAREFUL! The scaled vector, $k * u$, won't necessarily be pointing to the same point as the other vector, v, unless they lie on the same line and are pointing in the same direction (they could be on the same line but facing opposite directions).

$k * \|u\| = \|v\| \rightarrow k = \frac{\|v\|}{\|u\|}$

So the number of servings of PB to add to our daily food intake is .96 servings, and the actual Fat and Protein content that would be added to our diet, the entries of VectorE or the coordinates of the point closest to our goal, would be:

$$\frac{\|v\|}{\|u\|} * PB = VectorE$$

Withour specific values and units:

$$.96\ servings * \begin{bmatrix} 15\ g\ of\ Fat/serving \\ 8\ g\ of\ Protein/serving \end{bmatrix} = \begin{bmatrix} 14.4\ g\ of\ Fat \\ 7.68\ g\ of\ Protein \end{bmatrix}$$

Or without units:

$$.96 * \begin{bmatrix} 15 \\ 8 \end{bmatrix} = \begin{bmatrix} 14.4 \\ 7.68 \end{bmatrix}$$

Let's do infinite amount of work . . . in finite steps: Take a step back and think about what we have done. We started with two vectors (PB and VectorGoal), and we found the vector along the PB line that was the best approximation to VectorGoal. Or we could also state it as finding the point along the PB line that is closest to the Goal point. Well every time we have the task of finding the best approximation along one line to another vector or point we don't want to have to go through the same work from scratch. One great power of algebra is that it allows us to do infinite amount of work in finite number of steps. If we represent the work we did with general vectors, we should be able to derive an algebraic formula that would take in two vectors, and spit out the best approximation vector. Then we can just use the formula and that can save us time.

Let's backtrack a little and keep all of our work in terms of the labels (variables) PB and VectorGoal, these are our inputs, and VectorE, our output. We found VectorE in two big steps. We first found the length of VectorE:

$$\|VectorE\| = \frac{PB}{\|PB\|} \bullet \frac{VectorGoal}{\|VectorGoal\|} * \|VectorGoal\| \tag{1}$$

Or in the simplified version:

$$\|VectorE\| = \frac{PB \bullet VectorGoal}{\|PB\|} \tag{2}$$

Then we used the length of VectorE and the length of PB to scale PB to the desired length (the length of VectorE).

$$\frac{\|VectorE\|}{\|PB\|} * PB = VectorE \tag{3}$$

We can combine these equations to get a formula for VectorE. We replace $\|VectorE\|$ from equation (3) with the right side of equation (2).

$$VectorE = \frac{\frac{PB \bullet VectorGoal}{\|PB\|}}{\|PB\|} * PB = \frac{PB \bullet VectorGoal}{\|PB\| * \|PB\|} * PB$$

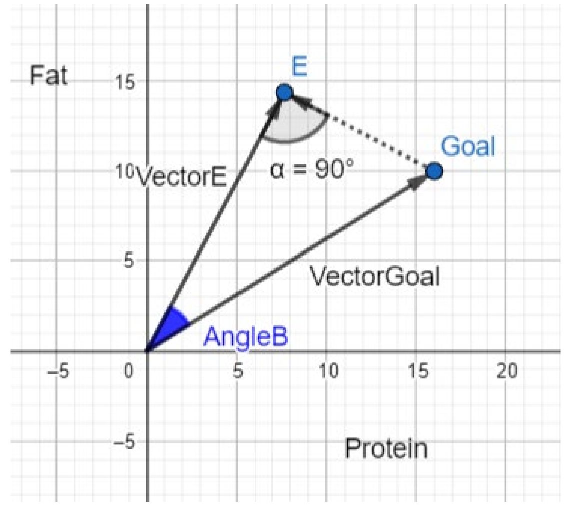
This matches the formula they give in the textbook, but they label u and v for the vectors. The formula is not symmetric, meaning we need to specify which vector is the one we are scaling and which one we are trying to get closest to. In our example, PB is the vector we are scaling and VectorGoal is the one we are trying to approximate. If you switch the PB and VectorGoal vectors in the formula above, then you won't get the same VectorE.

One of the problems with this formula is that in this simplified form it is hard to make sense of. By this I mean that it might be hard to look at it and see what every piece is doing to get you the best approximation vector. We can see that the best approximation vector is simply a scaled version of the PB vector because the expression $\frac{PB \bullet VectorGoal}{\|PB\| * \|PB\|}$ is simply a scalar (the dot product in the numerator produces a scalar, and the two lengths in the denominator or scalars too). However, it isn't clear (at least to me) when I look at it why the right side should produce the best approximation vector. If we unsimplify it, it is easier to see (at least for me) and connect each piece to the process we went through. Instead of substituting in the right part of equation (2) into equation (3), we use the right part of equation (1), which is not yet simplified, we get a different, but equivalent, equation. I have rearranged the terms so they allow us to chunk them into meaningful pieces.

$$VectorE = \frac{PB}{\|PB\|} \bullet \frac{VectorGoal}{\|VectorGoal\|} * \|VectorGoal\| * \frac{PB}{\|PB\|}$$

Here is one way to make sense of this equation (it may help to refer to the image we started with when we were applying trigonometry to the problem. I have pasted it below.)

$$VectorE = \underbrace{\overbrace{\frac{PB}{\|PB\|} \bullet \frac{VectorGoal}{\|VectorGoal\|}}^{\text{Cos(AngleB)}} * \overbrace{\|VectorGoal\|}^{\text{Length of hypotenuse}}}_{\text{Length of Vector E}} * \overbrace{\frac{PB}{\|PB\|}}^{\text{Unit vector in the Direction of PB}}$$



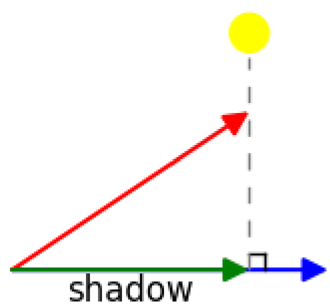
Now for me this makes a lot more sense. (I know that if it makes sense to the professor that it doesn't necessarily make sense to the students, but I hope this will help some of you). As a reminder, we are trying to find vectorE because it is the vector that points to the point on the line that PB falls on that is closest to Goal. VectorE and VectorGoal make two sides of a right triangle, with VectorGoal as the hypotenuse. Now looking at the equation we can see that the dot product of the two unit vectors is the cosine of the angle between them. From trigonometry, the length of the adjacent side (VectorE) is the length of the hypotenuse times the cosine of the angle between them (AngleB). The length of VectorE times the unit vector in the direction of PB will give VectorE, which is what we are trying to find.

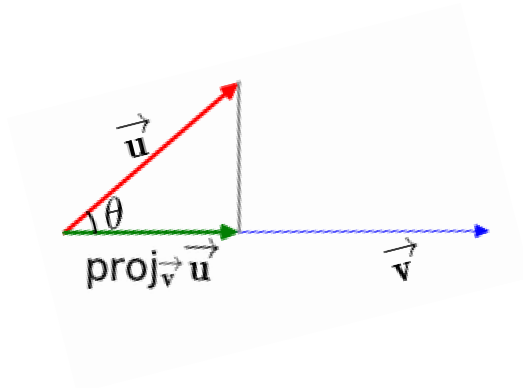
The textbooks don't give this as the equation because it isn't efficient (why divide by $\|VectorGoal\|$ if that means you will just have to multiply by $\|VectorGoal\|$ later? Just save time and don't do either). However, this longer equation is easier to interpret, and perhaps seeing it may make it easier for you to remember the other equation and where it comes from.

Projections=Best Approximation. Our problem was to figure out how to get closest to our goal of 16 grams of Protein and 10 Grams of Fat by adjusting the amount of PB in our diet. We have approached this problem by finding the closest point to our goal that we could access with just different amounts of PB. But the textbook doesn't use the language of best approximation or finding the closest point. It calls this process a **projection**. The idea of a projection is like the idea of a shadow. The images show the relationship of the "shadow" of a vector onto another vector. When we talk about projections it implies a perpendicular projection (the rays of the sun are perpendicular to the surface on which the shadow lies). The perpendicular projection matches the best approximation vector that we have been working with. Now instead of saying "find the vector on the span of PB that points to the closest point to the end of the vector VectorGoal"we can say "find the projection of VectorGoal onto PB," which is much shorter.

The image on the right introduces some new notation. It summarizes the phrase "find the projection of u onto v" with the notation $proj_{\vec{v}}\,\vec{u}$. The arrows over u and v signify that they are vectors, not scalars. I have just tried to be clear which variables are which so far and haven't introduced a different notation, but we will need some coming up. There are three common ways to show that variable or label is representing a vector and not a scalar or a matrix (which we will bump into soon): use of an arrow $\vec{v}$, use of a bar $\bar{v}$, and use of bold text $\boldsymbol{v}$. There are others but these are the most common you will see in this class. The textbook will use bold text or arrows, but on the board it will be easiest for me to use a bar.

Note: Some students find it hard to remember which vector is the subscript and which is not in the notation $proj_{\vec{v}}\ \vec{u}$*. One way to remember is that the projection of u onto v, and v is nearly below u in the notation. It can also help to have the right image in the illustration below in mind, and the projection goes from the upper vector to the lower one, then the lower one is the subscript. The lower one indicates the ground, or surface on which the projection (or shadow) will appear, so it makes sense to put it lower in the notation.*





Many applications of linear algebra can be characterized as finding a projection of a vector onto another vector or onto the span of a set of vectors (a plane or hyperplane). These problems try to find linear combinations of vectors (or a vector) that best approximate a point or another vector. For example, least-squares regression can be thought of as a projection of the outcome vector onto the span of the input variables. Audio compression, such as that used by your cell phone to pass audio data, can be done by projection of an audio data vector (representing a small time clip of voice data) onto a span of sine-waves of different frequencies. I know this might not make much sense right now, but as we become more familiar with the math of vectors, matrices, spans, and subspaces, we will see that linear algebra turns these hard problems into pretty straight-forward computations. We will do both of these problems this semester.

Chapter 5 Vectors as Space Tools

Considering that we have thought about vectors as coordinates and have used vectors to find angles and distances you may suspect, correctly, that vectors and matrices are useful in geometry. Vectors, vector operations, combinations of vectors, and the geometric objects they can represent, cause vectors and matrices to be extremely powerful tools in practical 2D and 3D geometry. They are powerful and practical because they deal with packets of quantities, things that represent real-world phenomena and match well with computations by computers. Locations are represented well as coordinate vectors. Geometric figures can be approximated or created from by packets of coordinates, which is simply a matrix. Manipulating the coordinates are done with vector and/or matrix operations.

In this chapter we attack a problem that will reveal the powerful and practical nature of vector geometry and will help us uncover some important geometric skills working with vectors in 3D. Vectors can be building blocks of points, lines, planes, and space. In higher dimensions, we deal with hyperplanes. It is hard to visualize the geometry beyond 3 dimensions, but the mathematics easily extends to higher-dimensional spaces, but higher dimensional geometry is usually best understood by analogy to 3-dimensional space.

The Y on the Mountain

On a mountain that overlooks Provo, UT is a mountain known by locals as Y-mountain. It isn't a real clever name. There is a very big man-made Y on the mountain so it is an obvious name. The Y represents BYU. When the Y was made the plan was to write all three letters on the mountain but they made it so big and it took so much work that they decided it wasn't worth it to put the other letters on the mountain. BYU has since become to be known as "The Y". The Y is one of the largest and oldest hillside letters (only two colleges have a larger one).

One of the remarkable things about the Y symbol on the mountain is that it looks very close to an actual block Y that would be on a BYU T-shirt, or on a poster. That might not be very remarkable until you realize that someone had to do some pretty careful engineering/surveying to make it look like that from campus. Below are three different pictures of the Y (and thanks to google earth I don't even need to leave my seat to get them). The first and third are views from BYU campus. The middle is a picture that is looking at the y in a direction approximately perpendicular from the mountain. That middle picture is what the actual Y looks like if the mountain was a piece of paper and you picked up the piece of paper, put it vertical, and looked at it. Well it is obviously a Y, but looks very long and skinny. The left fork is wider than the right and it looks like the angels of the forks, and other angles, are a bit off.

Now you may be wondering, if the first and third pictures are both from campus, why do they look so different. Well the first is from the old Academy Campus (the current building that is now the Provo City Library). The third is from the quad in the current campus. These pictures show that when it was created, the center of campus was not the current center of campus. The engineers had either the Academy building or possible the Karl G. Maeser Building as the view point. (In my view it looks best from the Academy).

Well, I think we should redo the Y on the mountain and make it look “correct” from the center of the current campus. Ok, I don’t really think we should, but that is the problem we are going to tackle to investigate the use of vectors, lines, planes, dot-products, etc. in 3D. This would have been similar to the math the original surveyors used (or could have used) to plan out the current Y. We have the amazing tool of google earth that will give us fairly precise coordinates and elevations, something that took much more work to get 120 or so years ago when mapping out the Y.

Below is a blue Y imposed on the image of the Y-view from campus. That is the new image we are looking for on the mountain, one that faces the correct way for the new center of campus.

Let's lay out a couple of assumptions. First, we are placing the new Y so the lower left corner is the same place as the old Y. We will assume the image of the new Y (the blue image above) is in a plane that faces directly towards the center of campus. That means a line from the center of campus to the blue Y is perpendicular to the plane that the blue Y is in. We are also going to assume the face of the mountain near the Y is flat, so it can be modeled with a plane. Let's also choose to make the origin the bottom left corner of the old Y and the blue Y.

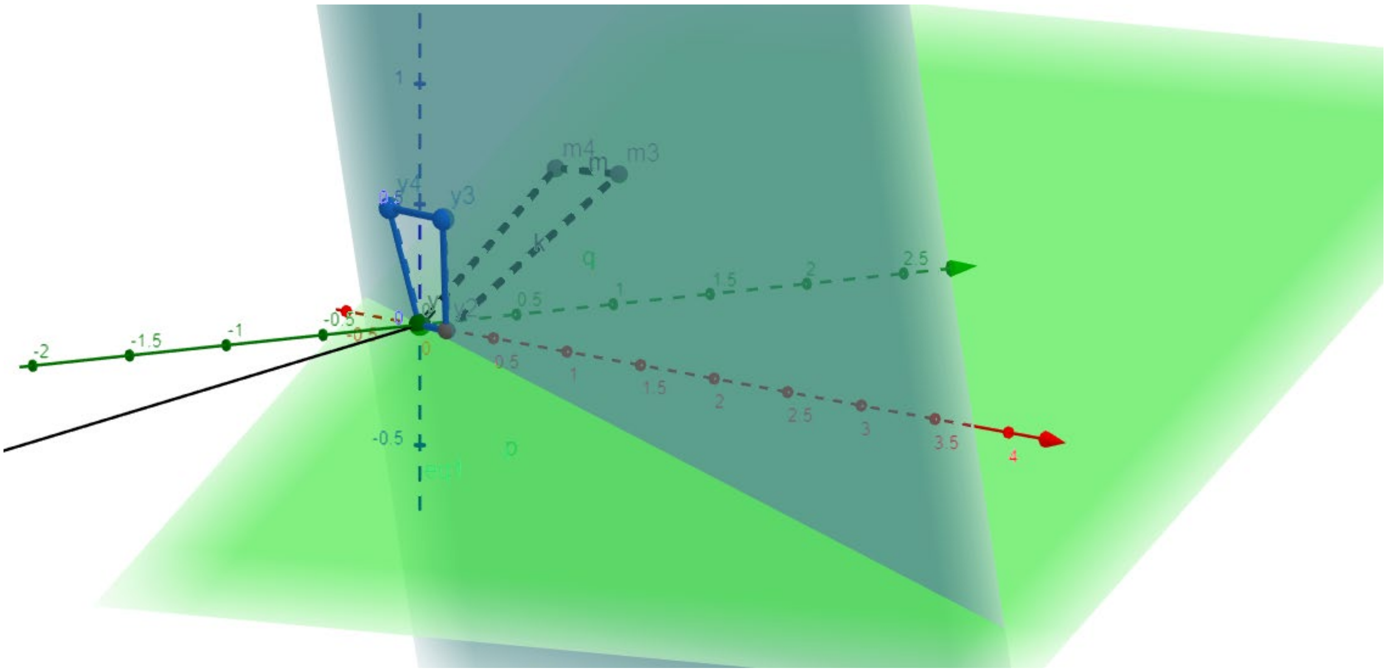
Note: To avoid confusion lets call the old white Y currently on the mountain OldY, the blue Y that we are using to make the new y, BlueY, and the new Y on the mountain NewY.

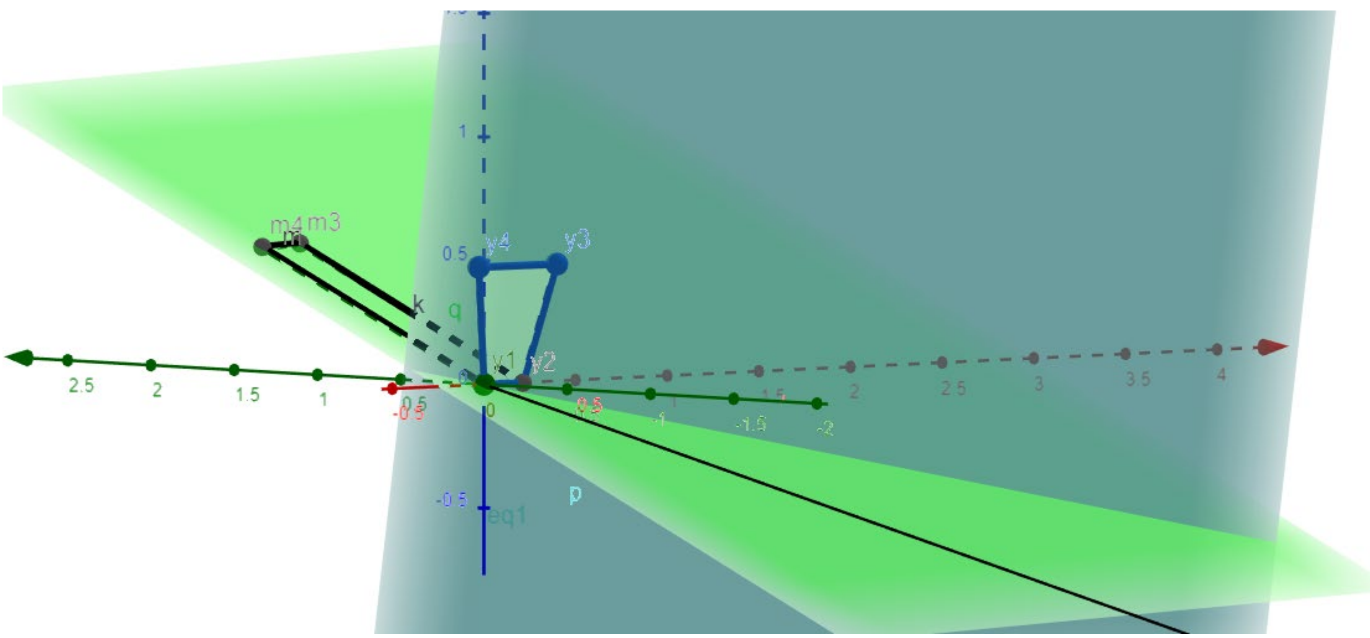
We need a reasonable unit for distance. Meters seems small but kilometers seem large. We will use 100 meter units (officially these are decameters, but for most of us this is about a football field length).

Since we are working in three dimensions we need to decide on a 3-dimensional grid. Which way is x? which way is y? Which way is z? Since I am going to be using geogebra to help us out with the math, it works out for south to be the positive x direction, east to be the positive y direction, and up to be the positive z direction.

For our activity here we are only going to work with four points on the Y, the top left and top right points, and the bottom left and right points. Just working on projecting these four points will give us plenty of opportunities to understand the mathematics. The four points are illustrated in the image below.

Below I have pasted two images that capture some of the features we are dealing with in this problem. The bluish plane represents planeB, the plane with the blue Y that we are using as our model. The quadrilateral with blue edges in the bluish plane represent the four corners of the BlueY. The blueish plane is "facing" campus. From the lower left corner of the BlueY there is a black line going directly away from planeB. That is the vector that goes from the origin to the center of campus. That vector is orthogonal/perpendicular to the planeB. The green plane is planeM, the mountain plane. The black quadrilateral on the green plane, represents the four corners of the OldY. We need to figure out the equation of the two planes, find the points for the blue Y, in the bluish plane, then projecting them onto the green plane. Where to start? Well to start off we need to find the equation of the two planes.





Planes in R^3

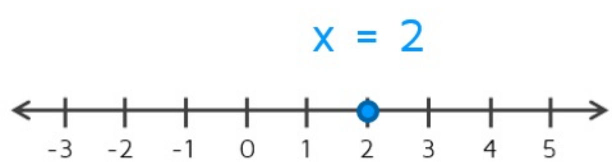
As a quick review a plane is a slice of space that has the same features of a table top, or a non-curved TV screen. It has no curvature or holes, but in mathematics we think of the planes extending infinitely. They are useful in the same way that lines and equations of lines are useful in modeling edges, paths, or relationships between quantities on an x-y plane.

Well, you probably have never even seen the equation of a plane before so you don't even know what we are trying to find. Something you may have heard is that three points determine a plane (as long as they don't all lie on the same line). This is similar to the fact that two distinct points determine a line. The equation of a line in the x-y coordinate plane is often written as $Ax+By=C$ or $y=mx+b$. A standard way to write an equation of a plane in the x-y-z coordinate grid is $Ax+By+Cz=D$ or $z=mx+ny+b$. The first one, with constants A,B,C, and D, is usually preferred. There are other ways to write equations of planes too, but we will start working with $Ax+By+Cz=D$, which is called the *general form* of the plane (or sometimes the *standard equation* of the plane).

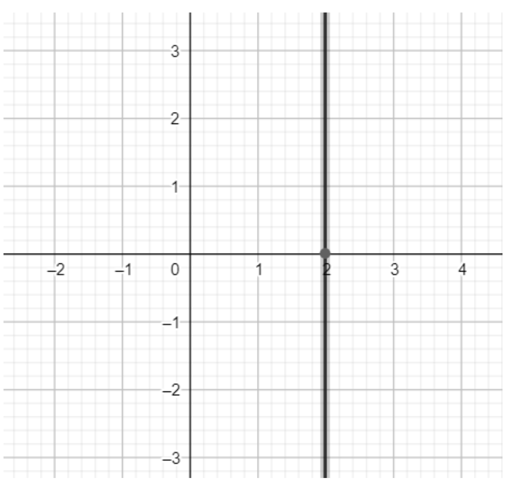
Just so I don't get accused of lying, $y=mx+b$ can also be the equation of a plane. Yes, you are probably only familiar with it as a line in the x-y plane, but if we are in the x-y-z grid then the set of all points (x,y,z) that satisfy the equation $y=mx+b$. "There is no z variable!" I hear you say. That is correct. But $y=3$ or $x=2$ are equations of lines and there is no x or y respectively. We can think about the equation $y=3$ and $y=0x+3$ or the equation $x=1$ and $x+0y+0z$ $=2$. We can even think about them as $y=0x+0z+3$ or $x+0y+0z=2$If we go back for enough in your education $x=1$ was just a point on the number line, not even a line. Now it is also a plane.

Let's take a brief look at how $x=2$ has evolved through your mathematical education:

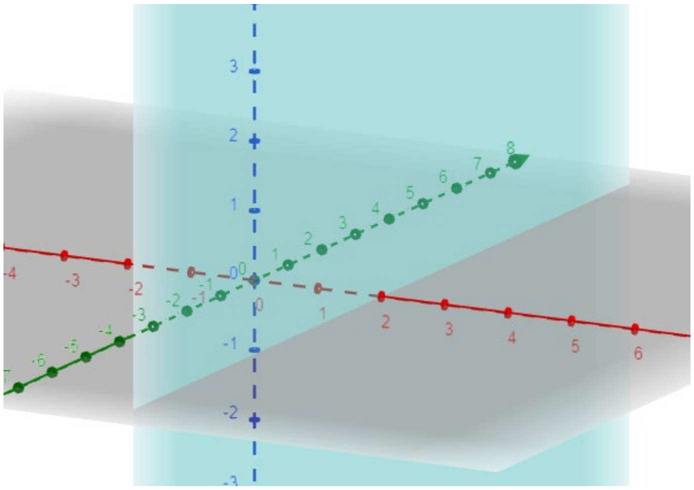
2nd Grade:



Algebra Class:



Linear Algebra:



(x=2 is represented by the light blue plane in the image above. x is the red axis.)

We could actually keep going and consider x=2 in a four dimensional space but I left my four dimensional printer in my other pair of hyperpants.

Enough history let's keep moving on.

Equation of PlaneM, the Mountain Plane

Let's start with the equation of the plane for the mountain. We will call it planeM (M is for mountain).

Well you know that the equation of the plane is Ax+By+Cz=D and it takes three points to determine a plane (and I just mentioned both of those) then you might be able to think of a strategy on how to use the three points to find A,B,C, and D. While you are thinking about that, I am going to go find two more points in planeM. (You already know that (0,0,0) is one point in this plane because OldY is in this plane and the lower left corner of OldY is the origin). In fact, just with that information we can figure out the value of D. See if you can do that while I go try to find some other points.

Ok I am back. It took me a little bit to use google earth to find the coordinates of two more points and to turn them into our XYZ points with our scale. I picked the upper left and upper right corners of OldY. Their coordinates are (-.34, .96, .53) and (-.04, 1.06, .53) respectively. Looking at the z coordinates of both (.53) means that the top of the OldY is about 53 meters higher in elevation than the bottom of the OldY. Also, the very small x-coordinate of the upper right point (-.04) shows that it is almost due East from the lower right point (the point we are taking to be the origin). If you ever get directly west of the Y (which is about the south side of the library), look up to the mountain and you can see that the upper right corner is quite close to being over the lower left.

Ok, while I was gone finding coordinates were you able to find D? To find D, all we have to do is plug in the values of 0 for x, y, and z. Since the point (0,0,0) is a point on the plane we are trying to find, it has to satisfy the equation of the plane (it has to make the equation true). If x,y, and z are all zero, then that means the left hand side of Ax+By+Cz=D is zero, so D must be zero for the equation to be true.

How about being able to devise a way to figure out a way to find A,B, and C? Well since points in the plane satisfy the equation, we can plug in the other points and create a system of equations. (This is different than what most of your teachers taught you to do with finding equations of lines, but it is still a fine strategy. If you have two points plug the values in for x and y in y=mx+b to get two equations in m and b, then solve by elimination or substitution.)

If we plug in our two points (upper right and upper left) into Ax+By+Cz=0 (remember we already used (0,0,0) to get D=0) then we get:

A(-.34)+B(.96)+C(.53)=0

A(-.04)+B(1.06)+C(.53)=0

Now some of you might already recognize a problem. We have two equations and three unknowns so we won't be able to solve for A,B, and C, just two of the three. We have infinite solutions here. But it is not really a problem. The equation Ax+By+Cz=D is quite flexible. It can represent the same plane in multiple ways. In the same way that Ax +By =C can represent the same line in multiple ways. x+y=1 is the same as 2x+2y=2. We have some freedom here and can actually pick the value of A, B, or C to be any non-zero value. We could do that now (say let A=1) or we can solve first and pick a value later based on the free variable. Let's just make A=1 and solve. We get B=-3 and C≈6.08.

Now that is enough for us to write down the equation of the planeM, the plane that models the face of the mountain: x-3y+6.08z=0

KNOW THE PROCEDURE: GIVEN THREE POINTS CAN YOU USE A SYSTEM OF EQUATIONS TO FIND THE EQUATION OF A PLANE IN R^3 (in general form).

Equation of PlaneB

One plane down, one to go. For the plane that the BlueY lies in (we will call it planeB), we don't actually have three points to work with, just one (the origin). But we do have another piece of information, the line that goes from the center of campus to the origin is perpendicular (or orthogonal) to the plane. It might take a little reasoning for us to figure out how we can use this info. What do we know about perpendicular lines? We know something about their slopes if we are just dealing with lines in the x-y plane: If m is the slope of one then -1/m is the slope of the other (m not equal to zero). But we are not in the x-y plane, we are working in the x-y-z grid so that doesn't seem to help us much. We did learn earlier that vectors that are orthogonal have a dot product that is zero. That is something that CAN help us. We can take the vector that points from the origin to the center of campus and find two other vectors that are perpendicular to that vector to find two points in planeB, then we are in the same situation we were in with planeM and we could solve a system of equations.

Mathematically, all vectors from the origin that are perpendicular to the vector from central campus to the origin will lie in planeB and any vector in planeB will be perpendicular to the central-campus-to-the-origin vector. This is actually a very important idea to understand in linear algebra but we will not have time to fully expound on it here.

The coordinates of the center of campus are (-1.68, -23.70, -4.25). So the center of campus is 168 meters north, -2370 meter west, and 425 meters lower in elevation than the lower left corner of the OldY (or origin). How do we use this information to create a vector?

DIFFERNCE BETWEEN A VECTOR AND A POINT

Let's pause just a moment and talk about the difference between a vector and a point. This is something that I have heard my students struggle with even late into a linear algebra class. The difference is not cut and dry, which is why it is a struggle at times to distinguish the difference between them. Sometimes a point and vector are interchangeable, so teachers will use "point" and "vector" interchangeably. Each point can be represented as a vector from the origin, and they will have the same coordinates. The point (1,1,1) can be thought as the end of the vector $\begin{bmatrix}1\\1\\1\end{bmatrix}$ if (AND THIS IS A BIG IF), the vector starts at the origin. When vectors begin at the origin it is called a vector in *standard position*. When vectors are in standard position then the correspondence between points and vectors is as easy as they come, they vectors and points have the same coordinates and are interchangeable for most practical purposes.

A point represents a location, but, a vector represents a direction! (Ok, technically a direction and a distance). The vector $\begin{bmatrix}1\\2\\3\end{bmatrix}$ shows a direction, but not a position. That vector could start at any point, but it represents a direction (the same direction as going from the origin to the point (1,2,3)) and a distance (the distance from (0,0,0) to (1,2,3). A vector captures a relationship between two points. That is why an arrow is a nice image to use to think of a vector. A vector shows the coordinates one point would need to move to get to another point. No matter where the point starts, the move is in the same direction and the same length.

Some things don't depend on location so they match with vectors very nicely. A move or force represented by the vector $\begin{bmatrix}1\\1\\1\end{bmatrix}$ and the move or force represented by the vector $\begin{bmatrix}4\\-2\\-2\end{bmatrix}$ are orthogonal, whether the vectors start at the some point, or end at the same point, or no matter where they are located. The moves/forces are orthogonal.

To draw/graph a vector in a program like geogebra, you can't just specify the vector. That is not enough information for geogebra to know where you, the user, wants the vector drawn/graphed. You also need to specify the starting point of the vector (or equivalent information to the vector and the starting point, like the starting and ending point of the vector).

Back To PlaneB

OK, now that we have a better handle on the difference between points and vectors let's pick up where we left off, How do we get a vector from the point at the center of campus, (-1.68, -23.70, -4.25), and the origin? We subtract one from the other. Which one do we subtract from which? For our problem it doesn't matter, but it will effect the direction the vector is pointing. When subtracting two points p,q to get a vector, then p-q will mean the tail of the vector is at q and the head of the vector is at p. In other words the vector points from q to p. If you want a vector that points from p to q (tail at p and head at q) then calculate the vector q-p. Do we want a vector that points from the center of campus to the origin (lower left point of the Y) or from the origin to the center of campus? It won't matter for us since they will both be perpendicular to all the vectors in planeB. Since most of us prefer to work with positive numbers, let's find the vector that points from the center of campus to the origin:

$$\begin{bmatrix}0\\0\\0\end{bmatrix}-\begin{bmatrix}-1.68\\-23.70\\-4.25\end{bmatrix}=\begin{bmatrix}1.68\\23.70\\4.25\end{bmatrix}$$

So any vector that is perpendicular to $\begin{bmatrix}1.68\\23.70\\4.25\end{bmatrix}$ will be in planeB (if you believe my claim from earlier).

How do find one? Well, that isn't too bad, we can use our knowledge of dot products to find a couple. The easiest way is to pick one of the values of a new vector to be zero, then select the other ones so the sum will be zero. For example consider:

$\begin{bmatrix}1.68\\23.70\\4.25\end{bmatrix}\bullet\begin{bmatrix}0\\a\\b\end{bmatrix}$ and $\begin{bmatrix}1.68\\23.70\\4.25\end{bmatrix}\bullet\begin{bmatrix}c\\d\\0\end{bmatrix}$

If we pick a=-4.25 and b=23.70 then the first dot product will be zero. If we pick c=-23.70 and d=1.68 then the second dot product will be zero. Now we have two vectors orthogonal to the campus-to-origin vector. We can use these to find two points in planeB by adding them to any point in the plane. The only other point we know to be in the plane is the origin, that means our vectors will be in standard positions so we can just use their coordinates to represent their corresponding points: (0,-4.25, 23.70), (-23.70, 1.68, 0). With these two points and the origin we are in the same situation that we were in before. This time I simply typed in "plane through ((0,0,0), (0,-4.25, 23.70), (-23.70, 1.68, 0)) into wolframalpha and it spit out: 0.395294 x + 5.57647 y + z = 0, which I am going to round to .4x+5.58y+z=0.

Now we have an equation for both planes, planeM and planeB.

A Quicker Method for Finding the Equation of Plane when an Orthogonal Vector and Point are Known.

There is actually a much quicker method to find the equation of planeB than the method we took. The work we did was important for you to understand because it captures work and reasoning that is very useful in a linear algebra class that we will continue to draw upon. From now on though if we have a vector perpendicular to a plane and a point in the plane, then this is the technique to use: Step 1 - use the coordinates of your vector and substitute them in for A,B, and C in the equation Ax+By+Cz=D (x-coordinate in for A, y-coordinate in for B, and z coordinate in for C). Step 2 – substitute your point in for x,y,z and calculate D. That is it.

Let's try this out in our example. We know that the vector $\begin{bmatrix}1.68\\23.70\\4.25\end{bmatrix}$ was perpendicular to planeB, so the equation should be 1.68x+23.70y+4.25z=D. Now to solve for D, evaluate the left side at the point (0,0,0), since we know the origin is in the plan, so it needs to satisfy the equation: 1.68*0+23.70*0+4.25*0=D => 0=D. So the equation of planeB is 1.68x+23.70y+4.25z=0. How does this compare to what we found before? .4x+5.58y+z=0? They don't look the same. Do they represent the same plane? Well in the second one the coefficient of z is 1, so one easy way to check is to divide both sides of the first equation by 4.25 (the coefficient of z).

(1.68/4.25)x+(23.70/4.25)y+(4.25/4.25)z=(0/4.25)

The coefficient of z is clearly 1, and the right side is clearly 0. The other two coefficients, upon calculating come out to be .395294 for x and 5.57647 for y, the some ones that were the output of wolframalpha and that we rounded to .4 and 5.58 respectively.

This new strategy works like a charm! In fact, even if we have three points, it is sometimes easier to use the three points to find two vectors (take two pair of points and subtract one from another to create vectors) then find a vector that is perpendicular to both (this might require a system of equations). Then use this method.

Why Does This Method Work?

You may have been wondering why it just so happens that the vector $\begin{bmatrix} A \\ B \\ C \end{bmatrix}$ is perpendicular/orthogonal to one vector in the plane Ax+By+Cz=D. (I know I sure wondered that when I was in linear algebra. By the way, this turns out to be a really nice fact to know, so don't forget it!) Well here is one way to see what is happening. Let's start with the special case when D=0, so the plane goes through the origin. In this case, any point in the plane can be represented as a vector in standard positions with the same coordinates. So if the point (1,2,3) was in the plane, then the vector $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ IN STANDARD POSITION is in the plane. (Remember, a vector itself doesn't give a location, just a direction and distance in that direction so we need to specify where the vector starts). So any point (x,y,z) in the plane means the vector $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ in standard position is also in the plane. (We actually don't need to make a big deal about the vector being in the plane, because even if we moved the vector outside of the plane it wouldn't change the vectors it is orthogonal to. Orthogonality has to do with what direction the vectors are pointing, nothing to do about where they are starting. However, I thought it might be easier for you to think about the vectors being in the plane).

Well here is where we get clever, the left side of the equation (Ax+By+Cz) is an expression that we can rewrite as a dot product: $\begin{bmatrix} A \\ B \\ C \end{bmatrix} \bullet \begin{bmatrix} x \\ y \\ z \end{bmatrix}$. But the right side of the equation is zero, so that means the dot product of the two vectors, $\begin{bmatrix} A \\ B \\ C \end{bmatrix} \bullet \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, is zero, but a dot product is zero ONLY IF THE VECTORS ARE ORTHOGONAL TO EACH OTHER! So when D=0, we can see that the vector made by the coefficients of x,y, and z in the equation of the plane is orthogonal to vectors in the plane.

What if (0,0,0) is not in the plane?

If the origin is not in the plane then we can do something similar, but we can't work with the vectors in standard position. To find a vector in the plane we need two points in the plane then we can subtract one point from another to get a vector. Let (x,y,z) be any point in the plane and let (p_x, p_y, p_z) be some other point in the plane. Then when we subtract them (technically subtracting the corresponding vectors in standard position), we get a vector: $\begin{bmatrix} x \\ y \\ z \end{bmatrix} - \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix} = \begin{bmatrix} x - p_x \\ y - p_y \\ z - p_z \end{bmatrix}$. Any vector that has a dot product with this xyz-minus-pxpypz vector that equals zero will be orthogonal to every vector in the plane. Let's start with the dot product of $\begin{bmatrix} A \\ B \\ C \end{bmatrix} \bullet \begin{bmatrix} x - p_x \\ y - p_y \\ z - p_z \end{bmatrix}$ and see if we can manipulate it into the equation of a plane in standard form:

$$\begin{bmatrix} A \\ B \\ C \end{bmatrix} \bullet \begin{bmatrix} x-p_x \\ y-p_y \\ z-p_z \end{bmatrix} = 0 \to \begin{bmatrix} A \\ B \\ C \end{bmatrix} \bullet \left(\begin{bmatrix} x \\ y \\ z \end{bmatrix} - \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix} \right) = 0 \to \begin{bmatrix} A \\ B \\ C \end{bmatrix} \bullet \begin{bmatrix} x \\ y \\ z \end{bmatrix} - \begin{bmatrix} A \\ B \\ C \end{bmatrix} \bullet \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix} = 0 \to \begin{bmatrix} A \\ B \\ C \end{bmatrix} \bullet \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} A \\ B \\ C \end{bmatrix} \bullet \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix}$$

Evaluating each dot product gives us:

$$Ax + By + Cz = Ap_x + Bp_y + Cp_z$$

Well the left side is exactly what we wanted, but what about the right side? The right side should be a constant, and it is since all six values on the right side are constants, but is it the right constant? What was the second step in the process of finding the equation of a plane when we know an orthogonal vector and a point in the plane? The first step is to take the coordinates of the orthogonal vector and make them the coefficients of x,y, and z. The second step is to use known point in the plane to solve for D. We do that by evaluating the left side at each of the respective coordinates of the known point. So when we have $Ax + By + Cz = D$, we plug in any point in the plane for x,y, and z to find D. But notice, the right side of this equation: $Ax + By + Cz = Ap_x + Bp_y + Cp_z$ is exactly that! It is the left side of the equation evaluated with coordinates of any point in the plane!

Let's quickly walk through an example. Suppose I want the equation of the plane that is has a normal vector (that is another word for a perpendicular vector or an orthogonal vector) of $\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$and goes through the point (1,2,-1).

We know that the equation will be 2x+y+z=D, and we can find D in two ways. We can plug in the point on the left hand side (2*(1)+1*(2)+1*(-1)=D) so D=3. Or we can find the dot product between the normal vector and the point in the plane (with the point represented as a vector in standard position): D= $\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} * \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$ =2*(1)+1*(2)+1*(-1)=3. The calculation is the same either way, but it illustrates why $Ap_x + Bp_y + Cp_z = D$.

A different way to think about a plane. The work we did has actually shown us that a plane is all of the points that have the same dot product with a given vector $\begin{bmatrix} A \\ B \\ C \end{bmatrix}$, where $\begin{bmatrix} A \\ B \\ C \end{bmatrix}$ is the coefficient vector from the general form equation of the plane. (We do have to represent the points as vectors in standard position of course). That might strike you as odd. It did to me. Since the value of a dot product can change just by scaling the vector, it seems like you could find another point, maybe very far away from another, and the two dot products would be different. Nope, you can't, not if they are in the same plane. One reason is that no matter where you move in the plane from point1 to point2, you are moving orthogonal to $\begin{bmatrix} A \\ B \\ C \end{bmatrix}$ (to make things easier I am going to call the vector $\begin{bmatrix} A \\ B \\ C \end{bmatrix}$, vector u. That means any new point you end up at in the plane can be written as a sum of two vectors, one that points from the origin to the starting point (call it v), and the other vector that is orthogonal to $\begin{bmatrix} A \\ B \\ C \end{bmatrix}$, call it w. Since the dot product is distributive over addition (that is $u \bullet (v + w) = u \bullet v + u \bullet w$) then $u \bullet (v + w) = u \bullet v$ because the dot product of u and w will always be zero (because $\begin{bmatrix} A \\ B \\ C \end{bmatrix}$ with a move in the plane represented by w create orthogonal vectors, and the dot product of orthogonal vectors is zero. Since the second term ($u \bullet w$) is always zero, it won't change the dot

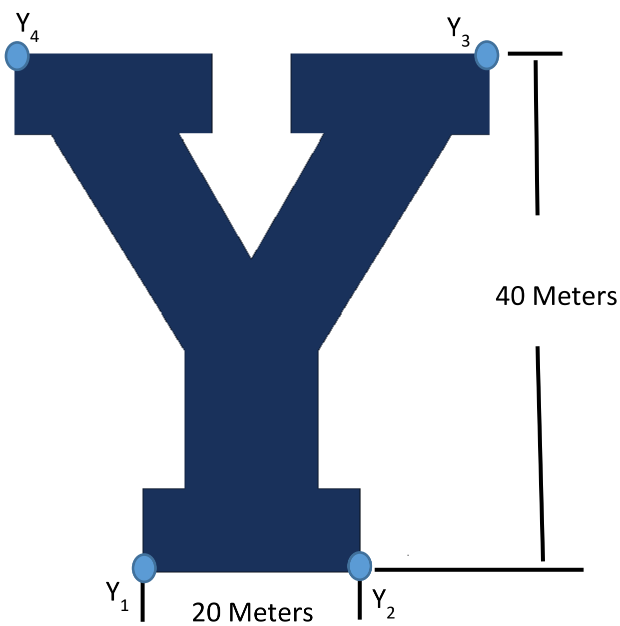
Finding the Coordinates of the Corners of BlueY in PlaneB.

If we want to keep the appearance of NewY from the center of the current campus to be the same as the OldY from the old Brigham Academy building, then we can use the information about OldY to figure out information of BlueY. The base (bottom line segment) of the OldY will be very close to the base of BlueY. From Google earth I measured the length of the bottom line segment of OldY and it was and 20 meters, so in our units, it would be a distance/length of .2 .

The increase in elevation from the bottom to the top of OldY is about 50 meters, so .5 in our units. (These are measurements from google earth). But Blue Y will not be that tall. If the height of BlueY is actually the same as the height of the OldY then NewY will reach higher up the mountain and appear taller than OldY. To be precise We should actually take one of the top coordinates of OldY, connect it with the coordinate of Brigham Academy, and find out where that line intersects the plane through the origin that faces Brigham Academy. We could then use that point of intersection to find the height of BlueY. However, I don't think we need another plane in our analysis, so I found the vector that points from the top-left point of OldY to Brigham Academy: $\begin{bmatrix} 6.18 \\ -41.5 \\ -5.07 \end{bmatrix}$. We can do a quick estimate of the height of BlueY with this vector. The direction to Brigham Academy from OldY is mainly West (-41.5). If we drop 507 meters (-5.07 is the drop in elevation in hundreds of meters) moving 4150 west, then it takes about 8 meters moving west to drop one meter in elevation (4150/507 ≈ 8). From the coordinates of the points at the top of OldY we can see that they are about 100 meters from the bottom of OldY in the eastward direction (the top left corner of OldY has an east coordinate of .96). This means that a 100 meter move westward from the top of OldY would result in a drop of about 12 meters. So the BlueY should be about 12 meters shorter than the raise in elevation of OldY. The raise in elevation was about 53 meters (the coordinates in the up direction were .53). So the height of BlueY should be 53-12=41 meters. We can round down to give us something a little nicer to work with so let's call it 40 meters.

[Image of the geometry of above- I need to put this in sometime]

We now know the dimensions of BlueY, (at least the dimensions that will allow us to get the four corner points).



If the lower-left corner of BlueY, Y_1, is the origin, then the lower-right point, Y_2,will be a point in planeB that is 20 meters to the "right" of the origin. We will assume we want the bottom and the top of BlueY to be horizontal, that means that the move from the lower-left to the lower-right points won't change in elevation. Another way of saying this is that these two points have the same z-coordinates. Since the lower-left is the origin, then both of these points will have a z-coordinate of 0. How do we find the x and y coordinates of the lower-right point?

Well we can use the equation of planeB to help us out here: PlaneB=> $.4x+5.58y+z=0$.

If we substitute zero in for z then we get $.4x+5.58y=0$. We can think of this as just the equation of a line. In fact it is the equation of the line at the intersection of two planes: PlaneB and the x-y plane (the x-y plane is the plane where z=0). We just need to move along this line from the origin to a point that is 20 meters (or .2 of our units) away. We do need to make sure we move in the right direction, after all, there are two points on this line that are 20 meters from the origin.

EQUATION OF A LINE IN R³: The intersection of two non-parallel planes is a line, and mathematicians use this fact as one way to write down the equation of a line in R^3. It may seem weird that the equation of a plane in R^3 is actually less complicated than the equation of a line in R^3 (at least written in general form). It takes two non-parallel planes to define a line so the equation of a line in R^3 is actually a system of equtions. The system is not unique, there could be intersections of a lot of different pairs of planes that intersect at the same line. It is similar to defining a point in the x-y plane as a solution to a system of equations of lines. In fact, that is precisely what is meant by the coordinate (a,b) in the plane. The point that lies on the line x=a and the line y=b.

PLANE: $Ax + By + Cz = D$

LINE: $\begin{Bmatrix} Ax + By + Cz = D \\ Ex + Fy + Gz = H \end{Bmatrix}$

We could use this to write down the equation of the line in R^3 that we found above (the equation of the line at the intersection of planeB and planeM):

$$\begin{Bmatrix} .4x + 5.58y + z = 0 \\ z = 0 \end{Bmatrix}$$

One way to figure out the x and y coordinates is to rearrange the equation of the line .4x+5.58y=0 so that we can get an expression for y in terms of x: $y = \frac{-.4}{5.58}x$. Now any point on the line can be thought of as a point with coordinates $\left(x, \frac{-.4}{5.58}x\right)$, or all vectors of the form $\begin{bmatrix} x \\ \frac{-.4}{5.58}x \end{bmatrix}$, or all multiples (or linear combinations) of the vector $\begin{bmatrix} 1 \\ \frac{-.4}{5.58} \end{bmatrix}$. [As a side note, the line can be represented as **x**=t*$\begin{bmatrix} 1 \\ \frac{-.4}{5.58} \end{bmatrix}$, and in general **x**=t*$\begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} c \\ d \end{bmatrix}$ is called the vector form of the equation of the line. In our case c and d are both zero since the line goes through the origin]. We can now use the distance formula to set up an equation to find the coordinates that we were looking for (the lower right point of BlueY, Y_2). The distance from (0,0) to $\left(x, \frac{-.4}{5.58}x\right)$ is $\sqrt{(x-0)^2 + (\frac{-.4}{5.58}x - 0)^2}$ and we want to find x so that that distance is the same length as the base of BlueY, .2 (or 20 meters). So we set up the following equation and solve:

$$\sqrt{(x-0)^2 + (\frac{-.4}{5.58}x - 0)^2} = .2$$

$$\sqrt{x^2 + (\frac{-.4}{5.58})^2 x^2} = .2$$

$$|x|\sqrt{1 + (\frac{-.4}{5.58})^2} = .2$$

Which finishes up as x=±.199488. First, do we care about the positive, negative or both values of x? Only the positive, since the lower right point is in the south direction and the south direction is moving in the positive x direction. What about the y-coordinate of the lower right point? Substituting .199488 for x in $y = \frac{-.4}{5.58}x$, gives us a y-coordinate of about -.014337.

That is an x-coordinate of nearly .2 and a y-coordinate that is nearly 0. This makes sense because the plane with BlueY, planeB, faces almost directly west, so moving southward in the plane doesn't move much in the east/west direction, almost completely in the north/south direction.

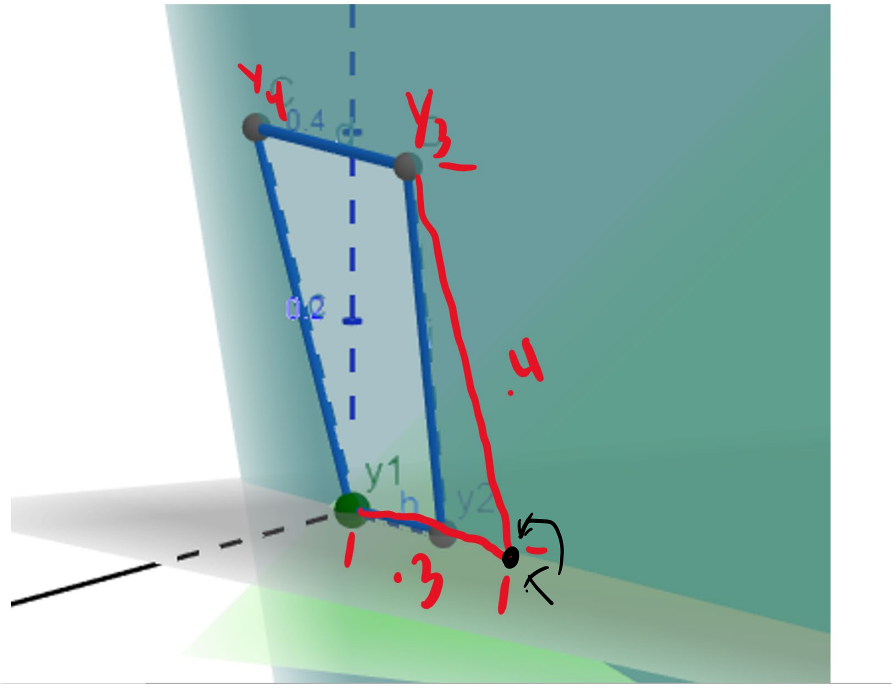
We have now found the coordinate of the lower right point, Y_2= (.2,-.014, 0). Now you may scoff at me rounding the x-coordinate up to .2, because now the distance is clearly more than .2 from Y_1, (0,0,0), to the coordinates Y_2 (since we go .2 units to the south AND .014 units to the west), but if you do the Pythagorean theorem (or calculate the norm of vector Y_2, $\|Y_2\|$, then you will see we are only off by 5 centimeters, about 2 inches, which isn't that big of a deal when you are looking at something from miles away.

The key ideas that we used in this section so far (in trying to find the coordinates so far of the points of BlueY in planeB) are:

- Finding the equation of two planes to find a line.
- Defining the equation of a line with equations of two non-parallel planes.
- Finding a point on a line a certain distance from another point on the line.
- Using the ratio of entries of a vector to think about how much we move in some direction (down for us) as we move in another direction (west).
- Using the context to make sense of the meaning of values of vectors/coordinates.

Finding the Other Points, Y^3 and Y^4, on BlueY

We will just find one of these points, the other one can be found with a very similar technique. Let's tackle Y_4 together. This upper right point is over to the right (southward-ish) 30 meters, so .3 in our units, and it is up-ish a height of 40 meters, or .4 in our units. It is not exactly south and up, because the plane is tilted a little to the north and downward, facing BYU campus. But to travel from origin (Y_1) to Y_4 would be .3 in one direction and .4 in an upward direction if you stayed on the blue plane the entire trip. I have tried to illustrate this in the image below. The moves of .3 to the "right" and .4 "up" are not just in an x,y, or z direction, could be moving in 2 or three of these at the same time.



We will have to find this with some round-about work. We are actually figure out each piece of the path in the figure above (over .3 then up .4) to figure out the final coordinate. Over .3 is pretty easy since we already did the hard work of finding the lower-right corner (Y2), and we move in the same direction. Since we are moving from the origin, the coordinates of Y2 give us the entries of the vector the point from the origin to Y2: $\begin{bmatrix} .2 \\ -.014 \\ 0 \end{bmatrix}$. Multiplying that by 3/2 (to get a length of .3 instead of .2) gives us a vector of $\begin{bmatrix} .299 \\ -.021 \\ 0 \end{bmatrix}$ (with a slight correction for our rounding earlier) that points from the origin to the point I have labeled T in the image above (for short I will just call this vector T).

Now the question is, “How do we find the vector that points from T to Y3?” What do we know about this vector?

- It lies in planeB.
- It is perpendicular to the vector we just found (from the origin to T).
- It has length of .4.
- Since it is in planeB then it is perpendicular to the vector from the origin to the center of campus

We are luck that we know that it is perpendicular to two other vectors. That can help us a lot, since we know it will have dot products with those vector that equal zero. That sounds like a good place to start. Lets pick $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$ to be the vector from T to Y3. Then we know that:

$\begin{bmatrix} a \\ b \\ c \end{bmatrix} \bullet \begin{bmatrix} .299 \\ -.021 \\ 0 \end{bmatrix} = 0$ and that $\begin{bmatrix} a \\ b \\ c \end{bmatrix} \bullet \begin{bmatrix} 1.68 \\ 23.70 \\ 4.25 \end{bmatrix} = 0$.

Expanding them as equations and writing them as equations gives us a system:

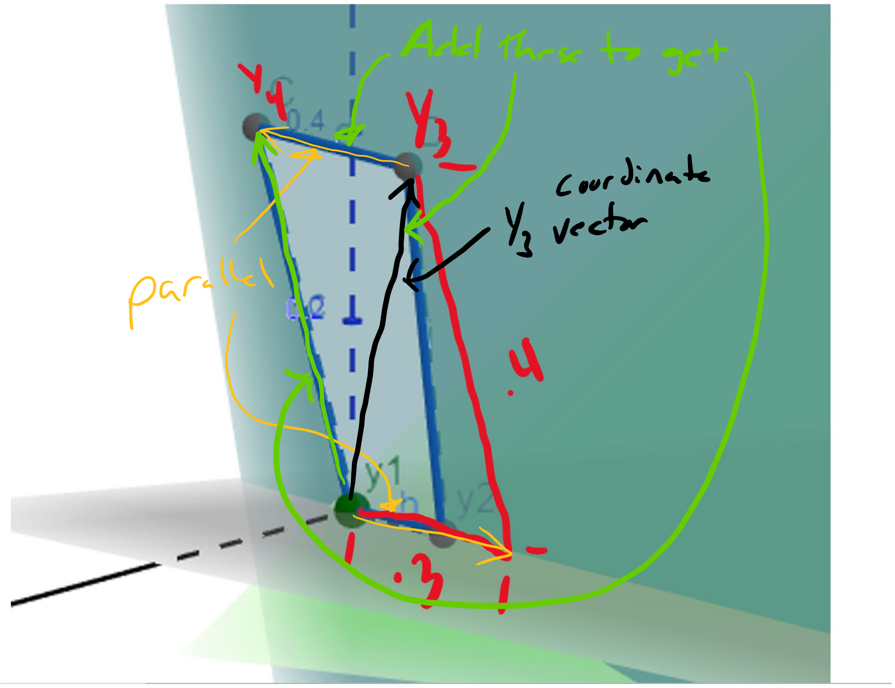
.299a-.021b=0 and 1.68a+23.70b+4.25c=0 which we could solve a, b, and c. We have 2 equations and 3 unknowns so that is not enough information to find a,b, and c exactly. That makes sense because a scalar multiple of vector that is perpendicular to these vectors is also perpendicular to them. We can fix one of the unknowns before we solve, so I will let a=1 and let wolframalpha do the computations for us. I ended up with b≈14.23, c≈-79.80. So the vector $\begin{bmatrix} 1 \\ 14.23 \\ -79.80 \end{bmatrix}$ is one vector that is in planeB (perpendicular to the vector that points from the origin to the center of campus) and is perpendicular to vector T. Our problem is that it might not be pointing to Y3 if we started the vector at point T. In fact it can't, because the z coordinate is negative so this vector is pointing down, not up so the direction is off. The other issue is that the length of the vector needs to be .4 and it is clear that the vector $\begin{bmatrix} 1 \\ 14.23 \\ -79.80 \end{bmatrix}$ is much larger than .4. We can use our technique in finding the coordinates of Y2 to help us here. Which scaled version of $\begin{bmatrix} 1 \\ 14.23 \\ -79.80 \end{bmatrix}$ has a length of .4? Or in other words, for what value of t is $\begin{bmatrix} 1t \\ 14.23t \\ -79.80t \end{bmatrix}$ a vector of length .4? (We used x instead of t when we were searching for the coordinates of Y2 because we had put everything in terms of the x coordinate, but in general we can scale it with any variable).

$$\left\| \begin{matrix} 1t \\ 14.23t \\ -79.80t \end{matrix} \right\| = \sqrt{t^2 + 14.23^2t^2 + 79.8^2t^2} = |t|\sqrt{1 + 14.23^2 + 79.8^2} = .4$$

Using the equation made with the last two expressions and solving for t give t=±.00493 . We know we need the negative value to get the z-coordinate to be the correct sign, so substituting t=-.00493 into $\begin{bmatrix} 1t \\ 14.23t \\ -79.80t \end{bmatrix}$ gives us $\begin{bmatrix} -.00493 \\ -.0702 \\ .3934 \end{bmatrix}$. Now this is NOT the coordinate vector of Y3. This is the vector that goes from T to Y3, not from the origin to Y3. We have to add this vector to vector T to get the coordinates of Y3. The vector from the origin to Y3 is the sum of the vector that goes from the origin to T and the vector from T to Y3.

$\begin{bmatrix} .299 \\ -.021 \\ 0 \end{bmatrix} + \begin{bmatrix} -.00493 \\ -.0702 \\ .3934 \end{bmatrix} = \begin{bmatrix} .295 \\ -.091 \\ .3934 \end{bmatrix}$. Thus the coordinates of Y3 are (.295, -.091, .3934).

Well I said we would just do Y3 and not Y4, but we have actually done all the hard work for Y4 so it is straightforward to compute. It is easy because the vector from y1 to y2 is parallel the vector from y4 to y3. We just need to add the correct vector to the Y3 coordinate vector (the vector that points to Y3 and begins at the origin). I have illustrated these ideas in the image below. The two yellow vectors are parallel, but the top one is longer (.4) then the bottom (.3). Adding the top yellow to the black vector (that points to Y3 from the origin) will result in the green vector that points from the origin to Y4, which is just what we need to find the coordinates of Y4.



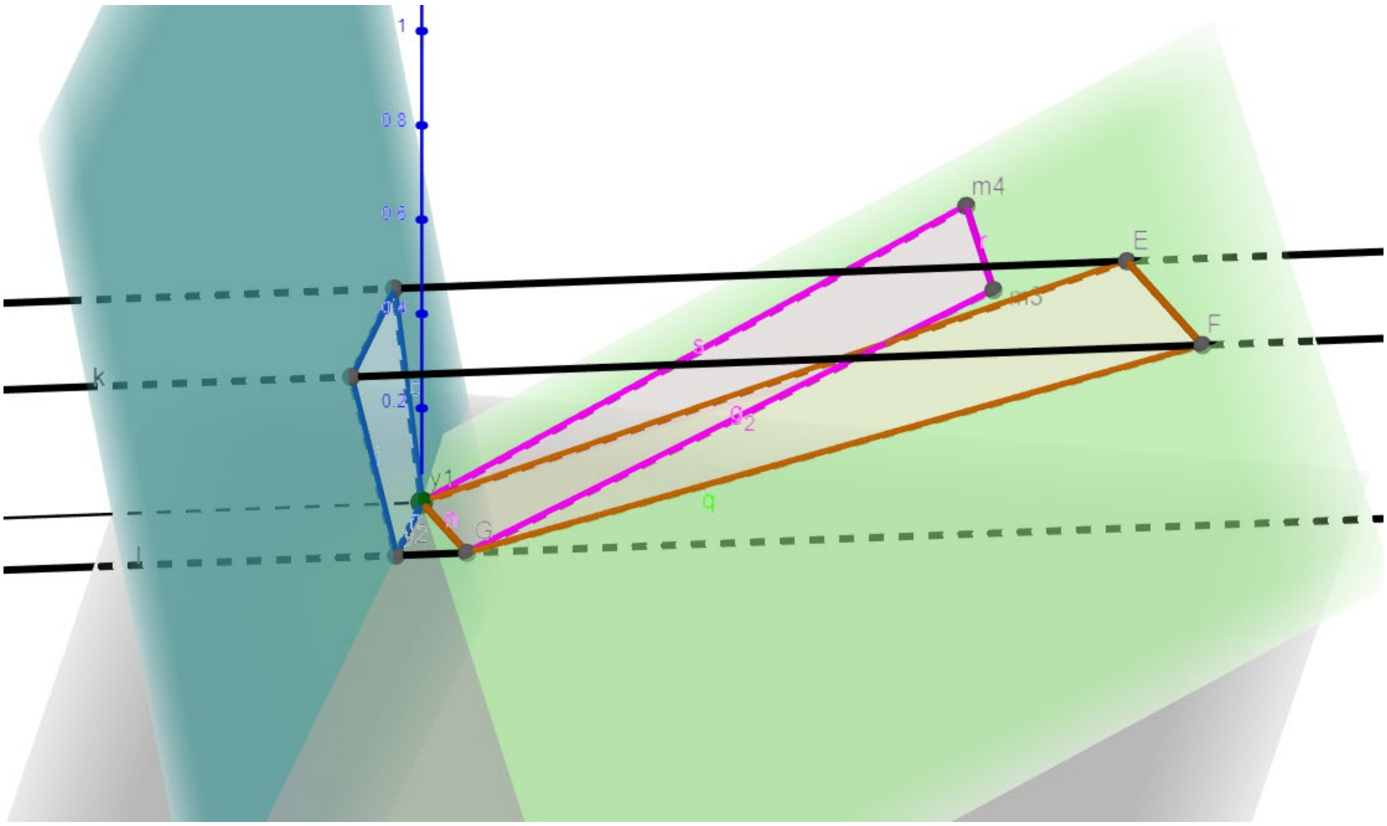
The vector from the origin to T (bottom yellow vector) needs to be multiplied by -4/3 to produce the yellow top vector. This will change the direction and stretch it to a length of .4 from a length of .3. We will add that to the Y3 coordinate vector we just calculated above:

$$\frac{-4}{3}\begin{bmatrix} .299 \\ -.021 \\ 0 \end{bmatrix} + \begin{bmatrix} .295 \\ -.091 \\ .3934 \end{bmatrix} = \begin{bmatrix} -.1037 \\ -.063 \\ .3934 \end{bmatrix}$$

So the coordinates of Y4 are (-.1037,-.063,.3934).

Projecting BlueY onto the Mountain from the Direction of the Center of Campus

We have reached our last step. We need project the 4 corners of BlueY onto the mountain so they are inline with someone at the center of BYU campus. Below is an image of what we are doing. The pink quadrilateral is approximately the corner positions of OldY, the blue polygon is the corner positions of BlueY, and the brown polygon represents the position of NewY. It took us a lot of time to find the coordinates of the corners of Blue Y, but, but now it is pretty easy to find the coordinates of NewY. Our strategy is to find the equation of the lines from campus through the vertices of BlueY and find out were these lines intersect the equation of the mountain plane. the equation of the line and the equation of the plane is enough to find the point of intersection.



A note on equations of lines in R^3: We have worked with the equation of a line in a situation earlier. We wrote down the equation as a system of equations where each equation represented a plane. The intersection of the two planes is a line (as long as the planes are not parallel, otherwise they might never intersect or be the same plane where they intersect at more than just a line). We wrote down the equation of a line in equations illustrated below.

General equation of a line in R3: $\begin{Bmatrix} Ax + By + Cz = D \\ Ex + Fy + Gz = H \end{Bmatrix}$

A specific example of the line at the intersection of planeB and planeM: $\begin{Bmatrix} .4x + 5.58y + z = 0 \\ z = 0 \end{Bmatrix}$

The general form of a line (above) is nice for some uses, but if you have two points in R3 and are trying to find the line between them, there are other equations that are better. One easy way if you know two points (or a point and a direction vector), is to use one of the points then add to it all possible multiples of the vector between the two points. In the case of having a point and a direction vector already, you can write down the equation very easily. If we know a line goes through (1,1,0) and moves in the direction of $\begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix}$, then we can represent all of the points on the line as: $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix}$. Sometimes the left side is just written as **x** (in bold) to represent a vector, so it looks like: $\boldsymbol{x} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix}$, where **x** represents the vector of coordinates. The variable t is called a parameter.

Any value of t from negative infinity to infinity will generate a unique point on the line. One way to think about the equation of the line in this form is that the line is being traced out by a particle that is moving along the line, and t is the time that the that particle is at a particular point. This form of the equation of a line is called the ***vector form***. It is easy to generate from points on the line. An arbitrary vector form of a line can be written as:

$$\boldsymbol{x} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} + t \begin{bmatrix} d \\ e \\ f \end{bmatrix} \text{ or } \boldsymbol{x} = \boldsymbol{p} + t\boldsymbol{u}$$

The symbols $\boldsymbol{p}$ and $\boldsymbol{u}$ are names for the vectors in the first equation instead of writing them out. This form works in any dimension. For example, the line $\boldsymbol{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix} + t \begin{bmatrix} 1 \\ 4 \end{bmatrix}$ represents a line in the x-y plane that goes through the point (2,3) and moves in the direction of vector $\begin{bmatrix} 1 \\ 4 \end{bmatrix}$, or in other words, has a slope of 4 (a change of 1 in the x-direction corresponds to a change of 4 in the y direction).

Since we are finding equations of lines through a point at the center of campus and then through the corner points of BlueY, then vector form is an easy way for us to write down the equation of these lines. For example, the line going through campus (coordinates (-1.68, -23.70, -4.25)) through the upper right corner (Y3 has coordinates of .295, -.091, .3934) can by found by doing the following:

1) Find the direction vector, $\begin{bmatrix} .295 \\ -.091 \\ .3934 \end{bmatrix} - \begin{bmatrix} -1.68 \\ -23.70 \\ -4.25 \end{bmatrix} = \begin{bmatrix} 1.975 \\ 23.61 \\ 4.6434 \end{bmatrix}$
2) Write down in vector form using either point for **p**, $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} .295 \\ -.091 \\ .3934 \end{bmatrix} + t \begin{bmatrix} 1.975 \\ 23.61 \\ 4.6434 \end{bmatrix}$

Now we have the equation of one of the lines, and the plane (planeM) and we need to find out where they intersect. Finding out the instersection depends on how the line and plane are represented. If both were in general form we would have three equations in three variables (one from the plane, two from the line) and we could solve the system of equations to find the point (x,y,z) that satisfied all three equations. But we can't do that because our line is in general form. The connections that we have between our vector form of the line and the general form of plane is that they each have an x,y, and z. Maybe we can use substitution. If we rewrite the vector form into three separate equations (really a system of equations) then we get an expression for x, y, an z all in terms of t. We could substitute these x, y, and z's into the equation of the plane and then solve for t.
Let's try it.

> The equation of the line written as three separate equations (to the left) is called the **parameterized form** of the equation of a line.

Equation of Line: $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} .295 \\ -.091 \\ .3934 \end{bmatrix} + t \begin{bmatrix} 1.975 \\ 23.61 \\ 4.6434 \end{bmatrix} \rightarrow \begin{Bmatrix} x = .295 + 1.975t \\ y = -.091 + 23.61t \\ z = .3934 + 4.6434t \end{Bmatrix}$

Equation of Mountain Plane: $\mathrm{x} - 3\mathrm{y} + 6.08\mathrm{z} = 0$

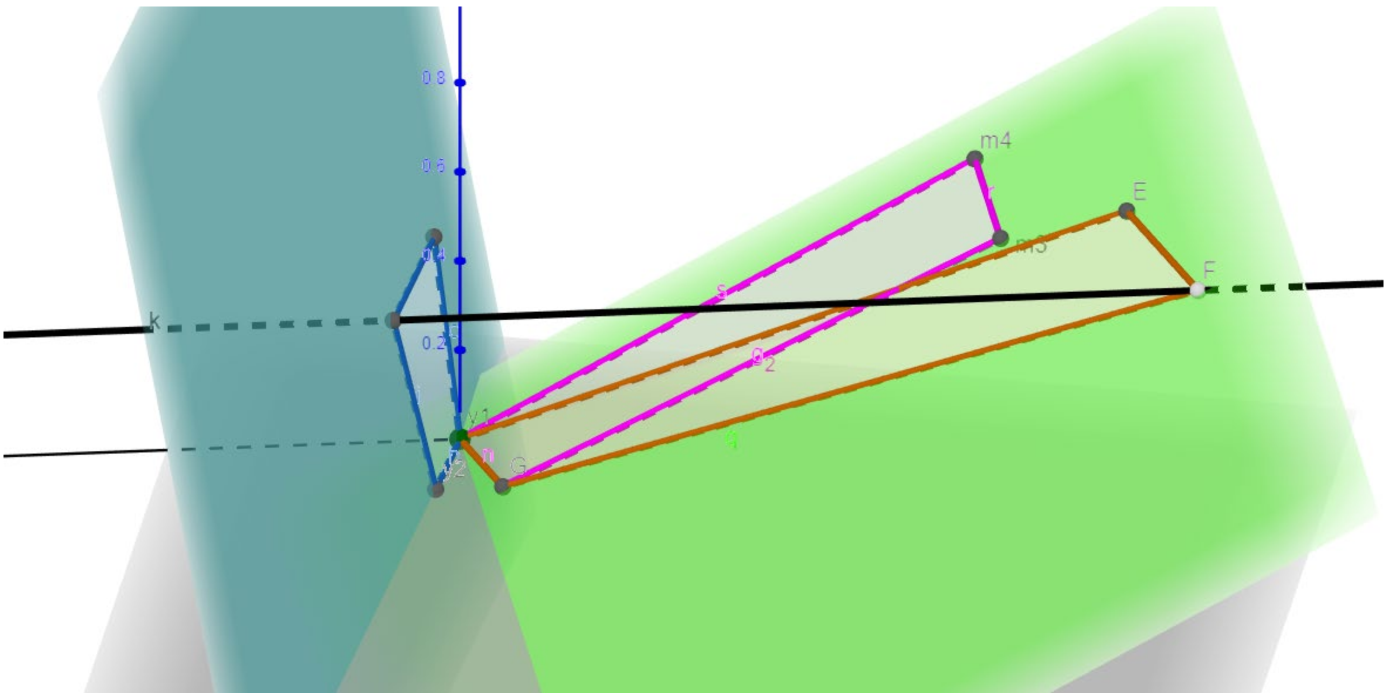
Substitute x,y, and z from the equation of line into the equation of the plane:

$$(.295 + 1.975t) - 3(-.091 + 23.61t) + 6.08(.3934 + 4.6434t) = 0$$

This is nice. It is just a linear equation in t. Solving for t give us $t \approx 0.07286$. Now to turn this information into a coordinate, we plug this back into our vector form (below). [Side note, one way of thinking about what we did in solving for t is to think about the line as being traced out by a particle and the vector equation connects t (a time) and the location of the particle on the line. We found the time that the particle intersects the mountain plane.

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} .295 \\ -.091 \\ .3934 \end{bmatrix} + .07286 \begin{bmatrix} 1.975 \\ 23.61 \\ 4.6434 \end{bmatrix} = \begin{bmatrix} 0.439 \\ 1.629 \\ 0.732 \end{bmatrix}$$

The vector that we just found has the coordinates of the point F in the image below. The black line that goes through the upper right corner of the blue and brown quadrilaterals is the line we had written in vector form. The green plane represents the mountain plane. The intersection of the two is the point F.



Now we could translate our coordinates of F to a latitude, longitude, and elevation and someone could use GPS device or GPS app, climb the mountain, and find the exact point of the upper right corner of NewY. We could go through the same process to find the four other corners (and any other important points that we need for layout).

Conclusion

I hope this example has helped to illustrate some of the tools and the power of these tools to work through practical problems in 3-dimensional geometry. One thing we didn't do is use our knowledge of dot products to find an angle. We didn't need to find any angles in this example, but we could easily find some if needed. Another common use in geometry is to do projections. We learned in the last chapter to project a vector onto another vector/line. Sometimes we project vectors (or images) onto planes (or hyperplanes when dealing in higher dimensions). Our example may help you to see how the use of these vector tools (and others we haven't got to yet) would be useful in 3D computer animations, or programming drones, or programming 5-axis mills to cut out surfaces, or things like that.