Math 54: Proof Workshop

February 28

Alas, there is no general recipe for mathematical proofs. (If there were, there wouldn't be so many open problems!) However, here are some tips that should be helpful for this class.

If and only if. If you are asked to prove that 'P' holds **if and only if** (sometimes written **iff**) 'Q' holds, then you are really being asked to prove two things: (1) that P implies Q (or, $P \Rightarrow Q$) and (2) that Q implies P, i.e., that $Q \Rightarrow P$. Pick one direction to tackle first. (Often one direction is easier than the other, but if it's hard to tell, just pick at random.) I recommend labeling each half of the proof on the page before you get started on it (e.g., by '$P \Rightarrow Q$,' where P and Q correspond to the statements in the problem). While you're working on one direction, just forget that you were also asked to prove other! It is not logically relevant until you complete the first half!

When you are working on, e.g., the $P \Rightarrow Q$ half, this means that you get to assume that P is true. Then you *want to show* (**WTS**) that Q is true. This aspect of the format is always the same. So I recommend writing down (after your '$P \Rightarrow Q$' label): 'Assume P. WTS: Q.' For example: 'Let $T : V \to W$ be a linear transformation, and assume that T is one-to-one. WTS: $\ker(T) = \{\mathbf{0}\}$.' The next step is usually to unpack what P and Q mean. More on this later....

Proving that something is a subspace. The definition of subspace isn't changing, so neither should your format for these proofs. Suppose you are asked to prove that V is a subspace of a vector space W. We need to check a few things:

- That V is a *subset* of W. Usually this is inherent in the way V is defined in the problem, but you should give the reason, whatever it is.
- That $\mathbf{0} \in V$. Usually V is defined by some rule, and you need to check that $\mathbf{0}$ satisfies the rule. For example, if $V = \text{Null}(A)$ where A is some given $m \times n$ matrix, then $\text{Null}(A) = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}$. So for $\mathbf{0}$ to be in $\text{Null}(A)$, we need precisely that $A\mathbf{0} = \mathbf{0}$. This is true, so we're done.
- That V is closed under addition. You should do the following. 'Assume $\mathbf{u}$ and $\mathbf{v}$ are in V. WTS: $\mathbf{u} + \mathbf{v} \in V$.' Unpack what it means for $\mathbf{u}$ and $\mathbf{v}$ to be in V — this is exactly what we need to use to show that $\mathbf{u} + \mathbf{v} \in V$. For example, let's say that $V = \text{span}\{\mathbf{b}_1, \mathbf{b}_2\}$. Then $\mathbf{u} \in V$ means that we can write $\mathbf{u} = \lambda_1\mathbf{b}_1 + \lambda_2\mathbf{b}_2$ for some scalars λ_1, λ_2. This is a good example of a sort of useful logic that may be unfamiliar. The premise that $\mathbf{u} \in \text{span}\{\mathbf{b}_1, \mathbf{b}_2\}$ means that **there exist** scalars λ_1, λ_2 such that $\mathbf{u} = \lambda_1\mathbf{b}_1 + \lambda_2\mathbf{b}_2$. We don't know what they are, but we know that, whatever $\mathbf{u}$ is, as long as it is in $\text{span}\{\mathbf{b}_1, \mathbf{b}_2\}$, I can find some scalars to make that equation hold. By saying something like 'there exist scalars λ_1, λ_2 such that $\mathbf{u} = \lambda_1\mathbf{b}_1 + \lambda_2\mathbf{b}_2$' or 'we can write $\mathbf{u} = \lambda_1\mathbf{b}_1 + \lambda_2\mathbf{b}_2$ for some scalars λ_1, λ_2,' these scalars magically come into being and we can use them for the duration of this part of the proof. Usually a proof is based on some tension: you are afforded some luxury in the hypothesis, and you need to figure out how to use this to pay for the desired result. When we get to write $\mathbf{u} = \lambda_1\mathbf{b}_1 + \lambda_2\mathbf{b}_2$ for some scalars λ_1, λ_2, inside our heads we should be saying 'yay.' We have cashed out our supposition that $\mathbf{u} \in V$ as a concrete statement involving some new quantities that we just constructed. We don't care where λ_1, λ_2 came from anymore; all we need to know is that we have $\mathbf{u} = \lambda_1\mathbf{b}_1 + \lambda_2\mathbf{b}_2$. In this example, since $\mathbf{v} \in V$ by assumption, we can similarly say that there exist scalars μ_1, μ_2 such that $\mathbf{v} = \mu_1\mathbf{b}_1 + \mu_2\mathbf{b}_2$. Try to finish the proof from here.
- That V is closed under scalar multiplication. 'Assume that $\mathbf{u} \in V$ and λ is a scalar. WTS: $\lambda\mathbf{u} \in V$.'

Make it clear what you are doing in each part!

Proving linear independence. Often the most useful way to approach such a proof is to use the definition directly. Remember: a set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly independent if the only

scalars $\lambda_1, \ldots, \lambda_n$ that make $\lambda_1\mathbf{v}_1 + \ldots + \lambda_n\mathbf{v}_n = \mathbf{0}$ hold are $\lambda_1 = \cdots = \lambda_n = 0$. This suggests the following strategy. (Think about why!)

'Let $\lambda_1, \ldots, \lambda_n$ be scalars such that $\lambda_1\mathbf{v}_1 + \ldots + \lambda_n\mathbf{v}_n = \mathbf{0}$. WTS: $\lambda_1 = \cdots = \lambda_n = 0$.'

Once we show this, we are done. Once again, we have cashed out our definition as a concrete statement involving some new quantities $\lambda_1, \ldots, \lambda_n$ that we just constructed, and we have a very clear goal. Yay!

Using linear dependence in a proof. Let's say that we know that a set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly dependent. Often the most useful thing that we can get out of this knowledge comes right from the definition. We know that there exist scalars $\lambda_1, \ldots, \lambda_n$ **not all zero** such that $\lambda_1\mathbf{v}_1 + \ldots + \lambda_n\mathbf{v}_n = \mathbf{0}$. (An equivalent fact is that we know that we can write one vector as a linear combination of the others, but do not underestimate the original definition!) This means that we can do the following: let $\lambda_1, \ldots, \lambda_n$ be scalars **not all zero** such that $\lambda_1\mathbf{v}_1 + \ldots + \lambda_n\mathbf{v}_n = \mathbf{0}$. Once again we have cashed out our definition as a concrete statement involving some new quantities. Somehow in the future we will need to use this equation, as well as the fact that the scalars are not all zero.

Proving linear dependence. Let's say we want to show that a set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly dependent. This means that we need to ***find*** scalars $\lambda_1, \ldots, \lambda_n$ **not all zero** such that $\lambda_1\mathbf{v}_1 + \ldots + \lambda_n\mathbf{v}_n = \mathbf{0}$. Inside our heads we are saying 'boo.' We know that we have to find scalars to make this equation hold, but we don't know what they are yet. We will have to construct them *in terms of previously constructed objects*, and this usually involves some cleverness.

PRACTICE: Let $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ where V is a vector space. Suppose that the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}\}$ is linearly dependent. Show that the set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly dependent. [Note that you have to use a linear dependence as a premise and prove a linear dependence!]

Using linear independence in a proof. This is a bit more annoying than using linear dependence in a proof. Unfortunately, if we know that a set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ of vectors is linearly independent, we cannot immediately cash this knowledge out and then forget about it. We have to remain on the lookout for how to use the linear independence. To wit, we need to engineer a situation in which we have $\lambda_1\mathbf{v}_1 + \ldots + \lambda_n\mathbf{v}_n = \mathbf{0}$ for some scalars. We don't know what the scalars are going to be at the beginning of the proof, but once we have an equation like that, we can finally deploy linear independence to conclude that $\lambda_1 = \cdots = \lambda_n = 0$. Then we have to trust that this fact will be useful for completing the proof.

PRACTICE: Let $\mathbf{u}, \mathbf{v} \in V$ where V is a vector space. Show that $\mathbf{u}$ and $\mathbf{v}$ are linearly independent if and only if $\mathbf{u} + \mathbf{v}$ and $\mathbf{u} - \mathbf{v}$ are linearly independent.

Proving that something is unique. In general, in math, if someone asks you to prove that there is at most one X that satisfies property P, the strategy is to do the following. 'Assume that X and Y both satisfy property P. WTS: $X = Y$.' (Think about why this is enough!) In this setting, we do not necessarily know that there exists X satisfying property P, or what it is if it does. For example, consider the definition of 'one-to-one.' A linear transformation $T : V \to W$ is one-to-one if for every $\mathbf{w} \in W$, there is at most one $\mathbf{v} \in V$ such that $T(\mathbf{v}) = \mathbf{w}$. Thus to show that T is one-to-one directly using the definition, we have to do the following. Let $\mathbf{w} \in W$ and assume that $\mathbf{v}_1, \mathbf{v}_2 \in V$ such that $T(\mathbf{v}_1) = \mathbf{w}$ and $T(\mathbf{v}_2) = \mathbf{w}$. WTS: $\mathbf{v}_1 = \mathbf{v}_2$. We have seen (and we will revisit below) the fact that T is one-to-one if and only if $\ker(T) = \{\mathbf{0}\}$. This equivalent condition is easier to check, so in the future if asked to prove that a linear transformation is one-to-one, you can prove that its kernel consists only of the zero vector.

In this class, a lot of statements look like 'the trivial solution is the unique solution' or 'the zero vector is the unique vector satisfying ____.' This is a more specific kind of uniqueness statement:

we know exactly what the thing is that is supposed to be unique. In general, if we are asked to prove that X is the unique object satisfying P, our strategy is the following. Assume that Y is an object satisfying P. WTS: $Y = X$. (We should also show that X actually satisfies P.)

Let's spell this out in a specific example. Suppose someone asks you to prove that $\mathbf{0}$ is the unique solution of the homogeneous system $A\mathbf{x} = \mathbf{0}$. 'Assume $\mathbf{x}$ is a solution of the homogeneous system. WTS: $\mathbf{x} = \mathbf{0}$.'

PRACTICE: Let $T : V \to W$ be a linear transformation where V and W are vector spaces. Prove that T is one-to-one if and only if $\ker(T) = \{\mathbf{0}\}$. [We did this in class, but it's a very good one to do again. Note that in proving the forward direction, i.e., that one-to-one implies $\ker(T) = \{\mathbf{0}\}$, we are being asked to prove that $\mathbf{0}$ is the unique vector $\mathbf{x}$ satisfying $T(\mathbf{x}) = \mathbf{0}$. So, for this direction, assume $T(\mathbf{x}) = 0$. WTS: $\mathbf{x} = \mathbf{0}$.]

Disproving things. Disproving a statement is the same thing as finding a counterexample! For some true/false questions, if it is not obvious why it is true, you might save yourself a lot of time by trying to find a simple counterexample.

Contradiction. One way to approach a proof is to try to prove it by contradiction. In general, it is a good habit to ask yourself: what would go wrong if this *weren't* true? More formally, to prove something by contradiction, assume that the thing that you are trying to prove is false, and then try to derive a contradiction. A lot of the time, you will find that you didn't actually need to do the proof by contradiction, but it is a valid approach and it can be a useful way to approach a problem.

Exposition. Your proofs should be easy to read! They are not supposed to look arcane! At the end of the day a proof is just a complete and correct argument that something is true, and you should write it in whatever way makes the argument most clear. Usually this corresponds roughly to how you would explain it verbally.