The purpose of this document is to provide a couple examples of how to organize and write proofs in this class. There is no single "format" that you must follow; I have aimed to highlight a couple different ways you can organize your work.

A few general tips to keep in mind.

1. The first step in writing a proof is to convince yourself that statement is true. Make sure you really believe it! Play devil's advocate, and look for any possible weakness in your reasoning.
2. You can write either in English or in symbols. If you write in symbols, remember that you should still follow basic grammar. The symbol "=" is a verb; an equation like $5 = x + 2$ is a complete sentence.
3. Always explain notation before you use it. Never write $\vec{x}$ until you've told me what it is, e.g. by writing "let $\vec{x} \in \mathbb{R}^3$ be an arbitrary vector."
4. Begin by clearly stating your assumptions.
5. Every statement you make (after your assumptions) should be supported either by statements you have already proved (or which have been proved in class) or by your assumptions. When in doubt, explicitly cite the fact that supports your claim (e.g. "by Theorem 2 on page 107 of the textbook").
6. You can always bring a draft proof to me in office hours, and we can talk through it, both in content and style.

# 1 Converses, and "if and only if" statements

Consider the following proposition[1].

**Proposition 1.** *Suppose that $A$ is an invertible $n \times n$ matrix, and $\vec{x}, \vec{b}$ are vectors in $\mathbb{R}^n$. Then $A\vec{x} = \vec{b}$ if and only if $\vec{x} = A^{-1}\vec{b}$.*

The phrase "if and only if" here actually means that we are making two statements at once:

- If $A\vec{x} = \vec{b}$, then $\vec{x} = A^{-1}\vec{b}$.
- If $A^{-1}\vec{b} = \vec{x}$, then $A\vec{x} = \vec{b}$.

These statements are closely related, but they are *not equivalent.* Instead, each one is called the *converse* of the other.

**Vocabulary:** The *converse* of a statement "P implies Q" is the statement "Q implies P." A statement can be true even if its converse is false. The phrase "P if and only if Q" means that *both* "P implies Q" and "Q implies P" are true. It is common in math to abbreviate the phrase "if and only if" by simply writing "iff."

Here is one way to write a proof of Proposition 1.

[1] A "proposition" is simply a statement to be proved. It is essentially interchangeable with the word "theorem," although traditionally "theorem" is reserved for more important or powerful statements.

**Proof of Proposition 1**

(“⇒”) Suppose that $A\vec{x} = \vec{b}$. Multiplying both sides by $A^{-1}$, it follows that $A^{-1}A\vec{x} = A^{-1}\vec{b}$. Since $A^{-1}A = I$ and $I\vec{x} = \vec{x}$, this implies that $\vec{x} = A^{-1}\vec{b}$.

(“⇐”) Conversely, suppose that $x = A^{-1}\vec{b}$. It follows that $A\vec{x} = AA^{-1}\vec{b}$. Since $AA^{-1} = I$ and $I\vec{b} = \vec{b}$, this implies that $A\vec{x} = \vec{b}$. Therefore $A\vec{x} = \vec{b}$ if and only if $\vec{x} = A^{-1}\vec{b}$. □

The two symbols “⇒” and “⇐” are just a handy visual cue to the reader, informing her which implication you are proving. They also help you quickly glance over your work to check that you’ve proved both of them. However, they are not strictly necessary.

Some observations about this proof:

1. It is really two proofs in one. Each paragraph is a self-contained proof.
2. The word “conversely” signals to the reader that one statement has just been proved (in this case, “if $A\vec{x} = \vec{b}$, then $\vec{x} = A^{-1}\vec{b}$”), and that the author is now setting out the prove the converse (“if $\vec{x} = A^{-1}\vec{b}$, then $A\vec{x} = \vec{b}$”).
3. Each paragraph is a sequence of claims. Each claim in the sequence is supported by the previous claims in the paragraph, or previously proved facts. For example, the first paragraph consists of the following four statements, each of which is justified in a different way:
   (a) $A\vec{x} = \vec{b}$ (the initial assumption)
   (b) $A^{-1}A\vec{x} = A^{-1}\vec{b}$ (justified by the phrase “multiplying both sides by $A^{-1}$”)
   (c) $A^{-1}A = I$ (justified by the definition of “inverse”)
   (d) $I\vec{x} = \vec{x}$ (justified by a general fact about identity matrices)
   (e) $\vec{x} = A^{-1}\vec{b}$ (justified by the previous three statements put together)

Here is a second way to write essentially the same proof, but relying more on symbols than on words.

**Notation:** The symbol “⇒” means “implies.” The symbol ⇔ means “if and only if.” These symbols can be chained together, as in the following proof. If you prove that $P \Rightarrow Q$ and $Q \Rightarrow P$, then you have proved that $P \Leftrightarrow Q$.

Another proof of Proposition 1 (more symbols, fewer words)

First, observe that

$$\begin{aligned} A\vec{x} = \vec{b} &\Rightarrow A^{-1}A\vec{x} = A^{-1}\vec{b} \\ &\Rightarrow I\vec{x} = A^{-1}\vec{b} \\ &\Rightarrow \vec{x} = A^{-1}\vec{b}. \end{aligned}$$

Conversely,

$$\begin{aligned} \vec{x} = A^{-1}\vec{b} &\Rightarrow A\vec{x} = AA^{-1}\vec{b} \\ &\Rightarrow A\vec{x} = I\vec{b} \\ &\Rightarrow A\vec{x} = \vec{b}. \end{aligned}$$

Therefore $A\vec{x} = \vec{b} \Leftrightarrow \vec{x} = A^{-1}\vec{b}$. □

## 2 Proving uniqueness

The following two propositions are both examples of proving that something is unique. There are two main ways to prove that something is unique:

1. Give an explicit formula for it, and prove that this formula is correct.
2. Begin by assuming that you have two such things. Prove that they are equal to each other.

Here's an example of a proof that proves that something is unique using the first method.

**Proposition 2.** *If $A$ is an invertible $n \times n$ matrix, and $\vec{b}$ is a vector in $\mathbb{R}^n$, then the matrix equation $A\vec{x} = \vec{b}$ always has a unique solution $\vec{x}$.*

Proof of Proposition 2

*Existence:* Proposition 1 implies that $A^{-1}\vec{b}$ is a solution.

*Uniqueness:* Proposition 1 implies that if $\vec{x}$ is a solution, then $\vec{x} = A^{-1}\vec{b}$. So no other solution exists. □

Comments:

1. Proposition 2 makes two claims: that a solution exists, and that it is unique. That is why the proof is separated into two parts, which are labeled to aid the reader.
2. Because we recently proved a proposition that is very useful (Proposition 1), it was convenient to simply cite it and use its conclusions. There is no need to reinvent the wheel.

Here's another statement, with the second sort of uniqueness proof.

**Proposition 3.** *Suppose that $A$ is a square matrix. If $A$ is invertible, then the inverse matrix is unique.*

Proof of Proposition 3

Suppose that $B_1$ and $B_2$ are both inverses of $A$. Then both $B_1A$ and $B_2A$ are equal to the identity matrix $I$. In particular, they are equal to each other:

$$B_1A = B_2A.$$

Therefore, multiplying both sides of this equation by $B_2$ on the right, and using the fact that $B_2$ is an inverse, we obtain the following equations.

$$\begin{aligned} B_1AB_2 &= B_2AB_2 \\ \Leftrightarrow B_1(AB_2) &= B_2(AB_2) \\ \Leftrightarrow B_1I &= B_2I \\ \Leftrightarrow B_1 &= B_2 \end{aligned}$$

(All the equations above are equivalent, since each rewrites the previous one in a different way, using associativity and the definition of identities and inverses.) Hence any two inverses of $A$ must in fact be equal to each other. □

Observe that, read literally, this proof demonstrated this statement: "If $B_1$ is an inverse of $A$, and $B_2$ is an inverse of $A$, then $B_1 = B_2$." The point is that this is just a cumbersome way to say what we actually care about: "the inverse of $A$ is unique."

## 3 The contrapositive

I mentioned earlier that the converse of a statement is not equivalent to the statement. However, there is another variation that *is* equivalent to the original statement.

**Vocabulary:** The *contrapositive* of a statement "P implies Q" is the statement "If Q is false, then P is false."
**Key fact:** A statement is true if and only if its contrapositive is true.

For example, first consider the following statement, which follows quickly from an earlier proposition.

**Proposition 4.** *If $A$ is invertible, then the only solution $\vec{x}$ to the matrix equation $A\vec{x} = \vec{0}$ is the trivial solution $\vec{x} = \vec{0}$.*

Proof of Proposition 4

It follows from Proposition 1 that the only solution to $A\vec{x} = \vec{b}$ is $\vec{x} = A^{-1}\vec{0} = \vec{0}$. □

Once we know this proposition, we can use its contrapositive to deduce the follow corollary[2].

**Corollary 5.** *If $A$ is a matrix, and there exists a nonzero vector $\vec{v}$ such that $A\vec{v} = \vec{0}$, then $A$ is not invertible.*

[2]The word "corollary," like "proposition," just means a statement to be proved. It is usually reserved for statements that can be easily deduced from a previously proved statement.

Proof of Corollary 5

Suppose that $A\vec{v} = \vec{0}$. Proposition 4 says that if $A$ is invertible, then $\vec{v} = \vec{0}$. By the contrapositive, if $\vec{v} \neq \vec{0}$, then $A$ is not invertible, as desired. □

## 4 Equality of sets

It is frequently convenient to express certain if and only if statements as equation of sets. For example, consider the following proposition.

**Proposition 6.** *Let $A$ be an $m \times n$ matrix, and $\vec{b} \in \mathbb{R}^m$ a vector. Suppose that the matrix equation $A\vec{x} = \vec{b}$ has at least one specific solution, $\vec{x} = \vec{v}$ (that is, $\vec{v} \in \mathbb{R}^n$ satisfies $A\vec{v} = \vec{b}$). Then the set of all solutions to $A\vec{x} = \vec{b}$ can be described in terms of the set of solutions to a homogeneous matrix equation, as follows.*

$$\left\{\vec{x} \in \mathbb{R}^n : \ A\vec{x} = \vec{b}\right\} = \left\{\vec{v} + \vec{w} : \ \vec{w} \in \mathbb{R}^n \text{ and } A\vec{w} = \vec{0}\right\}$$

*Comment 1:* The curly bracket notation above is called "**set builder notation**." If you've never seen it before, don't worry! With a bit of practice you'll become fluent with it. Here's how you read it: the expression $\left\{\vec{x} \in \mathbb{R}^n : \ A\vec{x} = \vec{b}\right\}$ would be read, in English, as "the set of all vectors $\vec{x}$ in $\mathbb{R}^n$ satisfying the equation $A\vec{x} = \vec{b}$." The text on the left of the colon gives a formula for the elements in the set, while the text on the right of the colon specifies the criteria that must be met in order to include this in the set. The formula on the right means: find all the choices of $\vec{w}$ solving the homogeneous equation, and add each one to $\vec{v}$; bundle all those sums together into a set.

*Comment 2:* It can be dizzying to keep the many variables in this statement straight, so take your time reading it! It may help to repeat to yourself a brief description of each variable when you read it, as in "$\vec{v}$, the specific solution to the inhomogeneous equation, plus $\vec{w}$, a solution to the homogeneous equation." The more you practice using the terms in complete sentences, the easier a time you'll have keeping straight which one is which.

There are many ways to prove that two sets are equal, but one of the most reliable and common in a course like this is a "double-containment proof." In this type of proof, you show two claims: first, assume you have an element of the set on the left. Prove that it is also an element of the set on the right. This proves that the left set is contained in the right set. Next, assume that you have an element of the set on the right. Prove that it is an element of the set on the left. This shows that reverse containment. Together, these two containments mean that the sets are equal. For example, here's a double-containment proof of Proposition 6.

Proof of Proposition 6

("$\subseteq$") Suppose that $\vec{x} \in \left\{\vec{x} \in \mathbb{R}^n : \ A\vec{x} = \vec{b}\right\}$. Then $A\vec{x} = \vec{b}$ and also $A\vec{v} = \vec{b}$ (since we are assuming that $\vec{v}$ is a specific solution). Therefore by distributivity, $A(\vec{x} - \vec{v}) = A\vec{x} - A\vec{v} = \vec{b} - \vec{b} = \vec{0}$. Let $\vec{w} = \vec{x} - \vec{b}$. Then we have shown that $A\vec{w} = \vec{0}$, so

$$\vec{x} = \vec{v} + \vec{w} \in \left\{\vec{v} + \vec{w} : \ \vec{w} \in \mathbb{R}^n \text{ and } A\vec{w} = \vec{0}\right\}.$$

("$\supseteq$") Suppose that $\vec{x} \in \left\{\vec{v} + \vec{w} : \ \vec{w} \in \mathbb{R}^n \text{ and } A\vec{w} = \vec{0}\right\}$. Then there exists $\vec{w} \in \mathbb{R}^n$ such that $A\vec{w} = \vec{0}$ and $\vec{x} = \vec{v} + \vec{w}$. It follows that

$$\begin{aligned} A\vec{x} &= A(\vec{v} + \vec{w}) \\ &= A\vec{v} + A\vec{w} \\ &= \vec{b} + \vec{0} \\ &= \vec{b}, \end{aligned}$$

hence $\vec{x} \in \left\{\vec{x} \in \mathbb{R}^n : \ A\vec{x} = \vec{b}\right\}$, as desired. □

The symbols $\subseteq, \supseteq$ mean "is a subset of" and "contains as a subset," respectively. Including them at the beginning of the two parts of the proof is optional, but can be a helpful visual cue to the reader.

Double-containment proofs can be understood as a specialized type of "if and only if" proof. You're proving that "for all $x$, $x$ is in the left set if and only if $x$ is in the right set." The two containments are merely one implication and its converse.

## 5 Casework

Sometimes in a proof, you will desperately want to make some simplifying assumption, but cannot because it is not necessarily true. For example, suppose we wish the prove the following statement (mentioned in class earlier).

**Proposition 7.** *Suppose that* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *is a* $2 \times 2$ *matrix, and* $ad - bc = 0$. *Then* $A$ *is not invertible.*

Here is a first attempt at a proof, which unfortunately contains a flaw.

Flawed proof of Proposition 7

Let $\vec{v} = \begin{pmatrix} d \\ -c \end{pmatrix}$. Then $A\vec{v} = \begin{pmatrix} ad - bc \\ cd - dc \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. By Corollary 5, $A$ is not invertible. □

**Before reading on, try to identify the flaw in this proof.**

The flaw in the proof is this: in order to apply Corollary 5, we need to assume that $\vec{v} \neq \vec{0}$. But this might not be right! For example, $A$ could be $\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$. This matrix has $ad - bc = 0$, and the statement we want to prove is true (it isn't invertible), but our proof doesn't work. We need a different proof for matrices like this one.

The key technique for situations like this is to do some casework. If you would love to make a simplifying assumption (in this case, the assumption that $\vec{v} = \vec{0}$), do it. The catch is that afterwards, you should make the opposite assumption, and prove that the statement you want is still true in that case.

More generally, you can break your work up into several distinct cases, each with a different assumption, and prove the statement in each case separately. As long as you can be sure that *at least one case* holds in every situation, you will have a complete proof. Here's an example, showing how to fix the flaw in the first proof attempt by working two separate cases.

**Corrected proof of Proposition 7**

Let $\vec{v} = \begin{pmatrix} d \\ -c \end{pmatrix}$. Observe that $A\vec{v} = \begin{pmatrix} ad - bc \\ cd - dc \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

*Case 1:* $\vec{v} \neq \vec{0}$. In this case, $\vec{v}$ constitutes a nonzero solution to the matrix equation $A\vec{x} = \vec{0}$. By Corollary 5, this implies that $A$ is not invertible.

*Case 2:* $\vec{v} = \vec{0}$. In this case, it follows that both $d$ and $c$ are equal to 0, so $A = \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix}$, i.e. the second row of $A$ is zero. But in this case, the matrix equation $A\vec{x} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ has no solution, since multiplying $A$ by any vector will always give 0 in the second coordinate. But Proposition 2 asserts that if $A$ is invertible, then every equation $A\vec{x} = \vec{b}$ must have a solution. By the contrapositive, it follows that $A$ is not invertible. □

If you have done any computer programming, this kind of casework should remind you of **checking edge cases.** In both cases, you first do the work of writing a proof (or writing some code) that works in a typical situation, and then you stop and think about any non-typical situations where the logic breaks down.

# 6 Proof by contradiction

Closely related to the concept of contrapositive statements is the technique of "proof by contradiction." It works like this: if you want to prove P, you can begin by *assuming that P is false,* and then deduce something absurd (a contradiction). This contradiction will show that it was impossible for P to be false after all, i.e. that P must be true.

As an example, here's another proof of Corollary 5.

**Another proof of Corollary 5, by contradiction**

Suppose that $A$ is a matrix and $\vec{v}$ is a nonzero vector such that $A\vec{v} = \vec{0}$.

↯ Assume, for the sake of contradiction, that $A$ is invertible. Then multiplying both sides by $A^{-1}$ yields $\vec{v} = A^{-1}\vec{0}$. The zero vector gives $\vec{0}$ when multiplied by any matrix, so $\vec{v} = \vec{0}$. This is a contradiction, since we assumed that $\vec{v}$ is nonzero ↯. Therefore our assumption cannot be true: $A$ is not invertible. □

The symbols ↯ and ↯ are optional (and idiosyncratic), but I like them. They are visual cues to label the place where you may an assumption to be contradicted, and the place where you obtain a contradiction. Together, they make it easy to see the logical outline of the proof, and also to quickly identify what it is that is proved once you get your contradiction.

As a final example, involving both some casework and a proof by contradiction, I will give a proof of the following fact that I stated without proof in class.

**Proposition 8.** *If $A$ is a nonsquare matrix, then $A$ is not invertible.*

The key to the proof I write below is to recognize that nonsquare matrices come in two flavors: either they have more columns than rows, or they have more rows than columns. It turns out that these two cases fail to be invertible for different reasons.

**Proof of Proposition 8**

Let $m, n$ the number of rows and columns in $A$, respectively. So $A$ is an $m \times n$ matrix.

*Case 1: $m < n$.* In this case, $A$ has more columns than rows. The equation $A\vec{x} = \vec{0}$ can be solved by setting up an augmented matrix with $A$ on the left of the partition, and all 0's on the right. After row-reducing this matrix, there is at most one pivot in each of the $m$ rows. Since $m < n$, this means that there is at least one column among the first $n$ columns that does not have a pivot in it. This means that there is at least one free variable in the general solution of $A\vec{x} = \vec{b}$. By selecting this free variable to be nonzero, we obtain a nonzero solution to $A\vec{x} = \vec{0}$. By Corollary 5, it follows that $A$ is not invertible.

*Case 2: $m > n$.* Assume for the sake of contradiction that $A$ is invertible. Since $A^{-1}A$ must be an identity matrix, which is a square matrix, it follows that $A^{-1}$ is an $n \times m$ matrix. So $A^{-1}$ has more columns than rows. By the proof in Case 1, $A^{-1}$ is not invertible. This is a contradiction, because $A$ is the inverse matrix of $A^{-1}$. So our assumption must be false: $A$ is not invertible after all. □