

The Principle of Mathematical Induction

Mathematical induction is a proof technique, not unlike direct proof or proof by contradiction. However, while both direct proof and proof by contradiction are somewhat natural styles of proof - they arise from the way we think of the statements we want to prove - proof by mathematical induction is different and a little bit strange. Let's start with an example.

1 Stamps

Here is a question about making different amounts of postage:

You need to mail a package, but don't yet know how much postage you will need. You have a large supply of 8-cent stamps and 5-cent stamps. Which amounts of postage can you make exactly using these stamps? Which amounts are impossible to make?

Try this problem on your own first. Perhaps in investigating the problem you picked some amounts of postage, and then figured out whether you could make that amount using just 8-cent and 5-cent stamps. Perhaps you did this in order - can you make 1 cent of postage? Can you make 2 cents? 3 cents? And so on. If this is what you did, you were actually answering a *sequence* of questions. We have methods for dealing with sequences. Let's see if that helps.

Actually, we will not make a sequence of questions, but rather a sequence of statements. Let $P(n)$ be the statement "you can make n cents of postage using just 8-cent and 5-cent stamps." Since each $P(n)$ is a statement, it is either true or false. So if we form the sequence of statements

$$P(1), P(2), P(3), P(4), \dots$$

the sequence will consist of T 's (for true) and F 's (for false). In our particular case the sequence starts

$$F, F, F, F, T, F, F, T, F, F, T, F, F, T, \dots$$

because $P(1), P(2), P(3), P(4)$ are all false (you cannot make 1, 2, 3, or 4 cents of postage) but $P(5)$ is true (use one 5-cent stamp), and so on.

Let's think a bit about how we could find the value of $P(n)$ for some specific value of n (the "value" will be either T or F). How did we find the value of the n th term of a sequence of numbers? How did we find a_n ? There were two ways we could do this: either there was a closed formula for a_n , so we could plug in n into the formula and get our output value, or we had a recursive definition for the sequence, so we could use the previous terms of the sequence to compute the n th term. When dealing with sequences of statements, we could use either of these techniques as well. Maybe there is a way to use n itself to determine whether we can make n cents of postage. That would be something like a closed formula. Or instead we could use the previous terms in the sequence (of statements) to determine whether we can make n cents of postage. That is, if we know the value of $P(n-1)$, can we get from that to the value of $P(n)$? That would be something like a recursive definition for the sequence. Remember, finding recursive definitions for sequences was often easier than finding closed formulas. The same is true here.

Suppose I told you that $P(43)$ was true (it is). Can you determine from this fact the value of $P(44)$ (whether it true or false)? Yes you can. Even if we don't know how exactly we made 43 cents out of the 5-cent and 8-cent stamps, we do know that there was some way to do it. What if that way used at least three 5-cent stamps (making 15 cents)? We could replace those three 5-cent stamps with two 8-cent stamps (making 16 cents). The total postage has gone up by 1, so we have a way to make 44 cents, so $P(44)$ is true. Of course, we assumed that we had at least three 5-cent stamps. What if we didn't? Then we must have at least three 8-cent stamps (making 24 cents). If we replace those three 8-cent stamps with five 5-cent stamps (making 25 cents) then again we have bumped up our total by 1 cent so we can make 44 cents, so $P(44)$ is true.

Notice that we have not said how to make 44 cents, just that we can, on the basis that we can make 43 cents. How do we know we can make 43 cents? Perhaps because we know we can make 42 cents, which we know we can do because we know we can make 41 cents, and so on. It's a recursion! As with a recursive definition of a numerical sequence, we must specify our initial value. In this case, the initial value is " $P(1)$ is false." That's not good, since our recurrence relation just says that $P(k+1)$ is true *if* $P(k)$ is also true. We need to start the process with a true $P(k)$. So instead, we might want to use " $P(31)$ is true" as the initial condition.

Putting this all together we arrive at the following fact: it is possible to (exactly) make any amount of postage greater than 27 cents using just 5-cent and 8-cent stamps.¹ In other words, $P(k)$ is true for any $k \geq 28$. To prove this, we could do the following:

1. Demonstrate that $P(28)$ is true.
2. Prove that if $P(k)$ is true, then $P(k+1)$ is true (for any $k \geq 28$).

Suppose we have done this. Then we know that the 28th term of the sequence above is a T (using (1) - the initial condition or *base case*), and that every term after the 28th is T also (using (2) - the recursive part or *inductive case*). Here is what the proof would actually look like.

Proof. Let $P(n)$ be the statement "it is possible to make exactly n cents of postage using 5-cent and 8-cent stamps." We will show $P(n)$ is true for all $n \geq 28$.

First, we show that $P(28)$ is true: $28 = 4 \cdot 5 + 1 \cdot 8$, so we can make 28 cents using four 5-cent stamps and one 8-cent stamp.

Now suppose $P(k)$ is true for some arbitrary $k \geq 28$. Then it is possible to make k cents using 5-cent and 8-cent stamps. Note that since $k \geq 28$, it cannot be that we use less than three 5-cent stamps *and* less than three 8-cent stamps: using two of each would give only 26 cents. Now if we have made k cents using at least three 5-cent stamps, replace three 5-cent stamps by two 8-cent stamps. This replaces 15 cents of postage with 16 cents, moving from a total of k cents to $k+1$ cents - so $P(k+1)$ is true. On the other hand, if we have made k cents using at least three 8-cent stamps, then we can replace three 8-cent stamps with five 5-cent stamps, moving from 24 cents to 25 cents, giving a total of $k+1$ cents of postage. So in this case as well $P(k+1)$ is true.

Therefore, by the principle of mathematical induction, $P(n)$ is true for all $n \geq 28$. QED

¹This is not claiming that there are no amounts less than 27 cents which can also be made - there are.

2 Formalizing proofs

What we did in the stamp example above works for many types of problems. Proof by induction is useful when trying to prove statements about all natural numbers, or all natural numbers greater than some fixed first case (like 28 in the example above), and in some other situations besides. In particular, induction should be used when there is some way to go from one case to the next - when you can see how to always “do one more.”

This is a big idea. Thinking about a problem *inductively* can give new insight into the problem. For example, to really understand the stamp problem, you should think about how any amount of postage (greater than 28 cents) can be made (this in non-inductive reasoning) and also how the ways in which postage can be made *changes* as the amount increases (inductive reasoning). When you are asked to provide a proof by induction, you are being asked to think about the problem *dynamically*; how does increasing n change the problem?

But there is another side to proofs by induction as well. In mathematics, it is not enough to understand a problem, you must also be able to communicate the problem to others. Like any discipline, mathematics has standard language and style, allowing mathematicians to share their ideas efficiently. Proofs by induction have a certain formal style, and being able to write in this style is important. It allows us to keep our ideas organized and might even help us with formulating a proof.

Here is the general structure of a proof by mathematical induction:

Induction Proof Structure

Start by saying what the statement is which you want to prove: “Let $P(n)$ be the statement...” To prove that $P(n)$ is true for all $n \geq 0$, you must prove two facts:

1. Base case: Prove that $P(0)$ is true. You do this directly. This is often easy.
2. Inductive case: Prove that $P(k) \rightarrow P(k+1)$ for all $k \geq 0$. That is, prove that for any $k \geq 0$ if $P(k)$ is true, then $P(k+1)$ is true as well. This is the proof of an if...then...statement, so you can assume $P(k)$ is true ($P(k)$ is called the *inductive hypothesis*). You must then explain why $P(k+1)$ is also true, given that assumption.

Assuming you are successful on both parts above, you can conclude, “Therefore by the principle of mathematical induction, $P(n)$ is true for all $n \geq 0$.”

Sometimes the statement $P(n)$ will only be true for values of $n \geq 4$, for example, or some other value. In such cases, replace all the 0’s above with 4’s (or the other value).

The other advantage of formalizing inductive proofs is it allows us to verify that the logic behind this style of argument is valid. Why does induction work? Think of a row of dominoes set up standing on their edges. We want to argue that in a minute, all the dominoes will have fallen down. For this to happen, you will need to push the first domino - that is the base case. It will also have to be that the dominoes are close enough together that when

any particular domino falls, it will cause the next domino to fall - that is the inductive case. If both of these conditions are met - you push the first domino over and each domino will cause the next to fall, then all the dominoes will fall.

Induction is powerful! Think how much easier it is to knock over dominoes when you don't have to push over each domino yourself. You just start the chain reaction, and then rely on the relative nearness of the dominoes to take care of the rest.

Think about our study of sequences. It is easier to find recursive definitions for sequences than closed formulas - going from one case to the next is easier than going directly to a particular case. That is what is so great about induction. Instead of going directly to the (arbitrary) case for n , we just need to say how to get from one case to the next.

When you are asked to prove a statement by mathematical induction, you should first think about *why* the statement is true, using inductive reasoning. Explain why induction is the right thing to do, and roughly why the inductive case will work. Then, sit down and write out a careful, formal proof using the structure above.

3 Examples

Here are some examples of proof by mathematical induction.

Example: Prove for each natural number $n \geq 1$ that $1+2+3+\cdots+n = \frac{n(n+1)}{2}$.

Solution: First, let's think inductively about this equation. In fact, we know this is true for other reasons (reverse and add comes to mind). But why might induction be applicable? The left hand side is adding up the numbers from 1 to n . If we know how to do that, adding just one more term ($n+1$) would not be that hard. For example, if $n = 100$, suppose we know that the sum of the first 100 numbers is 5050 (so $1+2+3+\cdots+100 = 5050$, which is true). Now to find the sum of the first 101 numbers, it makes more sense to just add 101 to 5050, instead of computing the entire sum again. We would have $1+2+3+\cdots+100+101 = 5050+101 = 5151$. In fact, it would always be easy to add just one more term. This is why we should use induction.

Now the formal proof.

Proof. Let $P(n)$ be the statement $1+2+3+\cdots+n = \frac{n(n+1)}{2}$. We will show that $P(n)$ is true for all natural numbers $n \geq 1$.

Base case: $P(1)$ is the statement $1 = \frac{1(1+1)}{2}$ which is clearly true.

Inductive case: Let $k \geq 1$ be a natural number. Assume (for induction) that $P(k)$ is true. That means $1+2+3+\cdots+k = \frac{k(k+1)}{2}$. We will prove that $P(k+1)$ is true as well. That is, we must prove that $1+2+3+\cdots+k+(k+1) = \frac{(k+1)(k+2)}{2}$. To prove this equation, start with the left hand side:

$$1+2+3+\cdots+k+(k+1) = \frac{k(k+1)}{2} + (k+1)$$

by the induction hypothesis. Now simplifying:

$$\frac{k(k+1)}{2} + k + 1 = \frac{k(k+1)}{2} + \frac{2(k+1)}{2} = \frac{k(k+1) + 2(k+1)}{2} = \frac{(k+2)(k+1)}{2}$$

Thus $P(k+1)$ is true, so by the principle of mathematical induction $P(n)$ is true for all natural numbers $n \geq 1$. QED

Note that in the part of the proof in which we proved $P(k+1)$ from $P(k)$, we used the equation $P(k)$ - this was the inductive hypothesis. This is very easy to do with proving a fact about a sum like this. Sometimes it is not as easy to see where to use the inductive hypothesis.

Example: Prove that for all $n \in \mathbb{N}$, $6^n - 1$ is a multiple of 5.

Solution: Again, start by understanding the dynamics of the problem. What does increasing n do? Let's try with a few examples. If $n = 1$, then yes, $6^1 - 1 = 5$ is a multiple of 5. What does incrementing n to 2 look like? We get $6^2 - 1 = 35$, which again is a multiple of 5. Next, $n = 3$: but instead of just finding $6^3 - 1$, what did the increase in n do? We will still subtract 1, but now we are multiplying by another 6 first. Viewed another way, we are multiplying a number which is one more than a multiple of 5 by 6 (because $6^2 - 1$ is a multiple of 5, so 6^2 is one more than a multiple of 5). What do numbers which are one more than a multiple of 5 look like? They must have last digit 1 or 6. What happens when you multiply such a number by 6? Depends on the number, but in any case, the last digit of the new number must be a 6. And then if you subtract 1, you get last digit 5, so a multiple of 5.

The point is, every time we multiply by just one more six, we still get a number with last digit 6, so subtracting 1 gives us a multiple of 5. Now the formal proof:

Proof. Let $P(n)$ be the statement, " $6^n - 1$ is a multiple of 5." We will prove that $P(n)$ is true for all $n \in \mathbb{N}$.

Base case: $P(0)$ is true: $6^0 - 1 = 0$ which is a multiple of 5.

Inductive case: Let k be an arbitrary natural number. Assume, for induction, that $P(k)$ is true. That is, $6^k - 1$ is a multiple of 5. Then $6^k - 1 = 5j$ for some integer j . This means that $6^k = 5j + 1$. Multiply both sides by 6:

$$6^{k+1} = 6(5j + 1) = 30j + 6$$

But we want to know about $6^{k+1} - 1$, so subtract 1 from both sides:

$$6^{k+1} - 1 = 30j + 5$$

Of course $30j + 5 = 5(6j + 1)$, so is a multiple of 5.

Therefore $6^{k+1} - 1$ is a multiple of 5, or in other words, $P(k+1)$ is true. Thus, by the principle of mathematical induction $P(n)$ is true for all $n \in \mathbb{N}$. QED

We had to be a little bit clever (i.e., use some algebra) to locate the $6^k - 1$ inside of $6^{k+1} - 1$ before we could apply the inductive hypothesis. This is what can make inductive proofs challenging.

In the two examples above, we started with $n = 1$ or $n = 0$. We can start later if we need to.

Example: Prove that $n^2 < 2^n$ for all integers $n \geq 5$.

Solution: First, the idea of the argument. What happens when we increase n by 1? On the left hand side, we increase the base of the square - we go to the next square number. On the right hand side, we increase the power of 2. This means we double the number. So the question is, how does doubling a number relate to increasing to the next square? Think about what the difference of two consecutive squares looks like. We have $(n + 1)^2 - n^2$. This factors:

$$(n + 1)^2 - n^2 = (n + 1 - n)(n + 1 + n) = 2n + 1$$

(this should not be a surprise - we already know that the sum of consecutive odd numbers gives the perfect squares). But doubling the right hand side increases it by 2^n , since $2^{n+1} = 2^n + 2^n$. Of course when n is large enough, $2^n > 2n + 1$.

What we are saying here is that each time n increases, the left hand side grows by less than the right hand side. So if the left hand side starts smaller (as it does when $n = 5$), it will never catch up. Now the formal proof:

Proof. Let $P(n)$ be the statement $n^2 < 2^n$.² We will prove $P(n)$ is true for all integers $n \geq 5$.

Base case: $P(5)$ is the statement $5^2 < 2^5$. Since $5^2 = 25$ and $2^5 = 32$, we see that $P(5)$ is indeed true.

Inductive case: Let $k \geq 5$ be an arbitrary integer. Assume, for induction, that $P(k)$ is true. That is, assume $k^2 < 2^k$. We will prove that $P(k + 1)$ is true, i.e., $(k + 1)^2 < 2^{k+1}$. To prove such an inequality, start with the left hand side and work towards the right hand side:

$$\begin{aligned} (k + 1)^2 &= k^2 + 2k + 1 \\ &< 2^k + 2k + 1 && \dots \text{by the inductive hypothesis} \\ &< 2^k + 2^k && \dots \text{since } 2k + 1 < 2^k \text{ for } k \geq 5 \\ &= 2^{k+1} \end{aligned}$$

Following the equalities and inequalities through, we get $(k + 1)^2 < 2^{k+1}$, in other words, $P(k + 1)$. Therefore by the principle of mathematical induction, $P(n)$ is true for all $n \geq 5$. QED

² $P(n)$ is not the statement " $n^2 < 2^n$ for all integers $n \geq 5$ " - you do not include the quantifier on n in the statement of $P(n)$.

The previous example might remind you of the *racetrack principle* from calculus, which says that if $f(a) < g(a)$, and $f'(x) < g'(x)$ for $x > a$, then $f(x) < g(x)$ for $x > a$. Same idea: the larger function is increasing at a faster rate than the smaller function, so the larger function will stay larger. In discrete math, we don't have derivatives, so we look at differences - when n increase by 1. Thus induction is the way to go.

Warning:

With great power, comes great responsibility. Induction isn't magic. It seems very powerful to be able to assume $P(k)$ is true - after all, we are trying to prove $P(n)$ is true (and the only difference is in the variable: k vs. n). Are we assuming that what we want to prove is true? Not really. We assume $P(k)$ is true only for the sake of proving that $P(k+1)$ is true.

Still you might start to believe that you can prove anything with induction. Consider this incorrect "proof" that every Canadian has the same eye color: Let $P(n)$ be the statement that any n Canadians have the same eye color. $P(1)$ is true, since everyone has the same eye color as themselves. Now assume $P(k)$ is true. That is, assume that in any group of k Canadians, everyone has the same eye color. Now consider an arbitrary group of $k+1$ Canadians. The first k of these must all have the same eye color, since $P(k)$ is true. Also, the last k of these must have the same eye color, since $P(k)$ is true. So in fact, everyone the group must have the same eye color. Thus $P(k+1)$ is true. So by the principle of mathematical induction, $P(n)$ is true for all n .

Clearly something went wrong. The problem is that the proof that $P(k)$ implies $P(k+1)$ assumes that $k \geq 2$. We have only shown $P(1)$ is true. In fact, $P(2)$ is false.

4 Strong Induction

Consider the following classic puzzle:

You have 9 coins which all look identical. However, you know that one of the coins is counterfeit, and weighs slightly less than the other coins (which all have the same weight). The weight difference is not great enough to tell by holding the coins, but luckily you have an old-fashioned balance scale (like the one held by Lady Justice). What is the smallest number of weighings needed to find the counterfeit coin?

If you have never tried to solve this puzzle, you should try now. Don't read on until you have given it some thought.

Need a hint? The answer is 2. Why?

Here's what you do. Put 3 coins on each side of the balance scale, leaving 3 off. Now if the scales are balanced, then the light coin must be one of the 3 you didn't weigh. If the scale tips, then the light coin is one of the three on the higher side. Now you have eliminated all but 3 coins using just one weighing. For the second weighing, put one of the three suspect coins on each side of the scales (leaving one off). Same idea: if the scales balance, then the unweighed coin is the counterfeit. Otherwise, the side of the scale that tipped up contains the light coin. Neat huh?

Let's generalize? If you had more than 9 coins to start with, say you had n coins, how many weighings would you need? A reasonable guess would be that each time you weigh the coins, you can eliminate $2/3$ of them. So using 3 weighings, you could find the counterfeit coin among up to 27 coins. A fourth weighing would be required for 28 to 81 coins.

Conjecture. *The minimum number of weighings required to find the counterfeit coin among n coins is $\log_3(n)$, rounded up. In other words, if x is the smallest integer such that $n \leq 3^x$, then the minimum number of weighings is x .*

How do we know this is true? If you have n coins, divide them into 3 equal piles (if this is not possible, make two piles the same size, and the third pile one more or less as needed). Put one pile on each side of the scale, leaving one pile off (if the piles were not equal, weigh the piles with an equal number of coins against each other). If the scales balance, the counterfeit coin is in the unweighed pile. If the scales tip, the counterfeit coin is in the light pile. In any event, we have now reduced the number of coins to roughly $n/3$.

What now? Well obviously we do this again. And again, until we find our bad coin. Here is where induction comes in: instead of going through the process over and over again, we can just say that we have reduced the problem to an earlier case, for which we assume our conjecture holds. For example if $n = 81$, we do the weighing as described, leaving us with $n/3 = 27$ coins. By induction, we already know how to complete the problem for 27 coins, so we are done.

This “reducing to a simpler case” is a type of induction. But not exactly the same sort of induction we have encountered so far. In ordinary induction, we assume our result holds for k , and prove that it holds for the next higher case: $k + 1$. There is not much difference in showing that the result holds for the case k assuming that it holds for the previous case ($k - 1$). But what we are doing here, is showing that the result holds for the case k , assuming it works for *all* smaller cases. Is this okay? In fact, it is. Think of the dominoes: to prove that they all fall down, it is enough to prove that the first one falls, and that if all the dominoes up to the k th one have fallen, then the k th one will fall too.

The advantage is that we now have a stronger inductive hypothesis. We can assume that $P(1), P(2), P(3), \dots, P(k)$ is true, just to show that $P(k + 1)$ is true. Previously, we just assumed $P(k)$ for this purpose.

It is slightly easier to change our variables for strong induction. Here is what the formal proof would look like:

Strong Induction Proof Structure

Again, start by saying what you want to prove: “Let $P(n)$ be the statement...” Then establish two facts:

1. Base case: Prove that $P(0)$ is true.
2. Inductive case: Assume $P(k)$ is true for all $k < n$. Prove that $P(n)$ is true.

Conclude, “therefore, by strong induction, $P(n)$ is true for all $n > 0$.”

Of course, it is acceptable to replace 0 with a larger base case if needed.³

Now let’s prove our conjecture about the coin weighing puzzle.

Proof. Let $P(n)$ be the statement, “if x is the smallest natural number such that $n \leq 3^x$, then it is possible to find the light coin in no more than x weighings.”

Base case: Consider $P(2)$. Here x is 1, since $2 \leq 3^1$. It is possible to find the heavy coin using just 1 weighing - put the coins on the scale.

Inductive case: Suppose we have n coins, and assume $P(k)$ is true for all $k < n$. Now divide the n coins into three piles - equal if possible, otherwise two equal with the third either one more or one less than the other two. Put two equal piles on each side of the scales. If the scales balance, then the counterfeit coin is in the unweighed pile. If the scales tip, the coin is in the pile on the lighter side of the scale. In any event, we now know the counterfeit coin is one of $k = n/3$ or $k = n/3 \pm 1$ coins. What’s more, if x is the smallest number such that $n \leq 3^x$, then $k \leq 3^{x-1}$. By the inductive hypothesis, it is now possible to find the counterfeit coin in $x - 1$ more weighings. Thus $P(n)$ is true.

Therefore, by strong induction, $P(n)$ is true for all $n \geq 2$.

QED

Here is a more mathematically relevant example.

Example: Prove that any natural number greater than 1 is either prime or can be written as the product of primes.

Solution: First, the idea: if we take some number n , maybe it is prime. If so, we are done. If not, then it is composite, so it is the product of two smaller numbers. Each of these factors is smaller than n (but at least 2), so we can repeat the argument with these numbers. We have reduced to a smaller case.

Now the formal proof:

Proof. Let $P(n)$ be the statement, “ n is either prime or can be written as the product of primes.” We will prove $P(n)$ is true for all $n \geq 2$.

Base case: $P(2)$ is true because 2 is indeed prime.

³Technically, strong induction does not require you to prove a separate base case. This is because when proving the inductive case, you must show that $P(0)$ is true, assuming $P(k)$ is true for all $k < 0$. But this is not any help - so you end up proving $P(0)$ anyway. To be on the safe side, we will always include the base case separately.

Inductive case: assume $P(k)$ is true for all $k < n$. We want to show that $P(n)$ is true. That is, we want to show that n is either prime or is the product of primes. If n is prime, we are done. If not, then n has more than 2 divisors, so we can write $n = m_1 \cdot m_2$, with m_1 and m_2 less than n (and greater than 1). By the inductive hypothesis, m_1 and m_2 are each either prime or can be written as the product of primes. In either case, we have that n is written as the product of primes.

Thus by the strong induction, $P(n)$ is true for all $n \geq 2$.

QED