Math 300 **Introduction to Mathematical Reasoning** **Autumn 2018**

Handout 11: Inverse Functions

Please read this handout in place of Section 6.5 in the text.

Solving Equations, Revisited

When we discussed surjectivity of functions, we noted that determining whether a function f is surjective often amounts to choosing an arbitrary y in the codomain and solving the equation $f(x) = y$ for x; and determining whether it's injective amounts to determining if the solution is unique when it exists. In such cases, we were often able to derive an explicit formula for the value of x in terms of y.

For example, Exercise 4(a) on page 318 of the textbook asks you to consider the function $F\colon \mathbb{R} \to \mathbb{R}$ defined by $F(x) = 5x + 3$. As the solution in the back of the book explains, F is bijective, and for each $y \in \mathbb{R}$, the number $x = (y - 3)/5$ is the *unique* solution to $F(x) = y$. This formula actually defines a new function $G\colon \mathbb{R} \to \mathbb{R}$:

$$G(y) = \frac{y-3}{5}.$$

This function has the feature that for each $y \in \mathbb{R}$, $G(y)$ is the unique number x such that $F(x) = y$.

The following definition generalizes this.

Definition. Suppose A and B are nonempty sets and $f\colon A \to B$ is a function. A function $g\colon B \to A$ is called an ***inverse function for f*** if it satisfies the following condition:

For all $a \in A$ and $b \in B$, $f(a) = b$ if and only if $g(b) = a$.

Thus, in the example above, G is an inverse function for F.

Theorems About Inverse Functions

Theorem 1. *Let A and B be nonempty sets, and let $f\colon A \to B$ and $g\colon B \to A$ be functions. Then g is an inverse function for f if and only if*

$$\text{for every } a \in A,\ g(f(a)) = a, \text{ and} \qquad (1)$$

$$\text{for every } b \in B,\ f(g(b)) = b. \qquad (2)$$

Proof. Assume first that g is an inverse function for f. We need to show that both (1) and (2) are satisfied. Let $a \in A$ be arbitrary, and let $b = f(a)$. Then by definition of an inverse function, $f(a) = b$ implies $g(b) = a$, so we can compute

$$g(f(a)) = g(b) = a.$$

This proves (1). To prove (2), let $b \in B$ be arbitrary, and let $a = g(b)$. Once again, the fact that g is an inverse function for f tells us that $g(b) = a$ implies $f(a) = b$, and therefore,

$$f(g(b)) = f(a) = b,$$

and (2) is proved.

Conversely, assume that f and g satisfy (1) and (2). Let $a \in A$ and $b \in B$ be arbitrary; we need to show that $f(a) = b$ if and only if $g(b) = a$. On the one hand, if $f(a) = b$, then

$$\begin{aligned} g(b) &= g(f(a)) && \text{(substituting } f(a) = b\text{)}, \\ &= a && \text{(using equation (1))} \end{aligned}$$

as desired. Conversely, if $g(b) = a$, then a similar argument using equation (2) shows that $f(a) = f(g(b)) = b$, and we are done. □

There is a nice way to rephrase this result in terms of compositions. Before we do so, let us note some important facts about compositions.

If A is any nonempty set, one of the simplest functions we can write down is the ***identity function of A***, denoted by $I_A \colon A \to A$, and defined by $I_A(x) = x$ for every $x \in A$.

Lemma 2. *Suppose A, B, C, D are any nonempty sets, and $f \colon A \to B$, $g \colon B \to C$, and $h \colon C \to D$ are any functions. Then the following identities hold:*

$$f \circ I_A = I_B \circ f = f,$$
$$(h \circ g) \circ f = h \circ (g \circ f).$$

Proof. See Exercises 4 and 5 on page 331 of the textbook. □

Here is a concise reformulation of Theorem 1 in terms of compositions.

Corollary 3. *Let A and B be nonempty sets, and let $f \colon A \to B$ and $g \colon B \to A$ be functions. Then g is an inverse function for f if and only if*

$$f \circ g = I_B \text{ and } g \circ f = I_A. \tag{3}$$

Proof. You can check using the definitions of composition and identity functions that (3) is true if and only if both (1) and (2) are true, and then the result follows from Theorem 1. □

Another important consequence of Theorem 1 is that if an inverse function for f exists, it is unique. Here is the proof.

Theorem 4. *Let A and B be nonempty sets, and let $f \colon A \to B$ be a function. If $g_1 \colon B \to A$ and $g_2 \colon B \to A$ are inverse functions for f, then $g_1 = g_2$.*

Proof. Let $f \colon A \to B$, and assume $g_1, g_2 \colon B \to A$ are both inverse functions for f. By Corollary 3, they satisfy

$$f \circ g_1 = I_B, \qquad g_1 \circ f = I_A,$$
$$f \circ g_2 = I_B, \qquad g_2 \circ f = I_A.$$

Using these formulas together with the results of Lemma 2, we compute as follows:

$$\begin{aligned} g_1 &= g_1 \circ I_B \\ &= g_1 \circ (f \circ g_2) \\ &= (g_1 \circ f) \circ g_2 \\ &= I_A \circ g_2 \\ &= g_2, \end{aligned}$$

which is what we wanted to prove. □

We now make the following definitions:

Definition. Let A and B be nonempty sets. A function $f\colon A \to B$ is said to be ***invertible*** if it has an inverse function.
Notation: If $f\colon A \to B$ is invertible, we denote the (unique) inverse function by $f^{-1}\colon B \to A$.

Using this notation, we can rephrase some of our previous results as follows.

Corollary 5. *Suppose $f\colon A \to B$ is an invertible function. Then*

$$f^{-1}(f(a)) = a \text{ for every } a \in A; \tag{4}$$

$$f(f^{-1}(b)) = b \text{ for every } b \in B; \tag{5}$$

$$f \circ f^{-1} = I_B \text{ and } f^{-1} \circ f = I_A. \tag{6}$$

Proof. These are just the results of Theorem 1 and Corollary 3 with g replaced by f^{-1}. □

It is important to be careful with the notation f^{-1}: The superscript in this case does *not* represent a multiplicative inverse; instead it represents a different function, one that satisfies (4), (5), and (6). Also, *the notation f^{-1} should never be used to represent an inverse function unless you have verified that f is invertible.*

Here is a simple criterion for deciding which functions are invertible.

Theorem 6. *A function is invertible if and only if it is bijective.*

Proof. Let $f\colon A \to B$ be a function, and assume first that f is invertible. Then it has a unique inverse function $f^{-1}\colon B \to A$. To show that f is surjective, let $b \in B$ be arbitrary, and let $a = f^{-1}(b)$. Then (5) implies $f(a) = f(f^{-1}(b)) = b$.

To show that f is injective, let $a_1, a_2 \in A$ and suppose $f(a_1) = f(a_2)$. Applying the function f^{-1} to both sides yields $f^{-1}(f(a_1)) = f^{-1}(f(a_2))$, and then (4) shows that $a_1 = a_2$.

Conversely, assume f is bijective. We define a function $g\colon B \to A$ as follows: Given $b \in B$, because f is surjective there is an element $a \in A$ such that $f(a) = b$, and because f is injective, a is the unique such element. Thus we can unambiguously define $g(b)$ to be this particular value of a. By our definition of g, we see that $g(b) = a$ if and only if $f(a) = b$. Thus by the definition of an inverse function, g is an inverse function of f, so f is invertible. □

These theorems yield a streamlined method that can often be used for proving that a function is bijective and thus invertible. Given a function $f\colon A \to B$, if we can (by any convenient means) come up with a function $g\colon B \to A$ and prove that it satisfies both $f \circ g = I_B$ and $g \circ f = I_A$, then Corollary 3 implies g is an inverse function for f, and thus Theorem 6 implies that f is bijective. Moreover, since the inverse is unique, we can conclude that $g = f^{-1}$.

Consider for example the function $F\colon \mathbb{R} \to \mathbb{R}$ given by $F(x) = 5x + 3$, which we studied above. Preliminary computations suggest that $G\colon \mathbb{R} \to \mathbb{R}$ given by $G(y) = (y-3)/5$ ought to be an inverse function for it. We can verify this by showing that $F \circ G = I_{\mathbb{R}}$ and $G \circ F = I_{\mathbb{R}}$, which is just a couple of computations:

$$F \circ G(y) = F\left(\frac{y-3}{5}\right) = 5\left(\frac{y-3}{5}\right) + 3 = (y-3) + 3 = y,$$

$$G \circ F(x) = G(5x+3) = \frac{(5x+3)-3}{5} = \frac{5x}{5} = x.$$

Once G is defined, this computation is all that is needed to prove that F is bijective and invertible, and that $G = F^{-1}$. (Of course, in general, you have to take care when defining G to ensure that it gives a well-defined element of A for every element of B, or in other words that it is indeed a function.)

Next we have a formula for the inverse of an inverse.

Theorem 7. *Let A and B be nonempty sets, and suppose $f\colon A \to B$ is invertible. Then $f^{-1}\colon B \to A$ is also invertible, and $(f^{-1})^{-1} = f$.*

Proof. By Corollary 3, f^{-1} is invertible if there is a function $g\colon A \to B$ that satisfies $g \circ f^{-1} = I_B$ and $f^{-1} \circ g = I_A$; and in that case the function g is the unique inverse of f^{-1}. Since $g = f$ is such a function, it follows that f^{-1} is invertible and f is its inverse. □

Now we consider inverses of composite functions. Suppose we have two functions $f\colon A \to B$ and $g\colon B \to C$:

$$A \xrightarrow{f} B \xrightarrow{g} C.$$

Theorem 6.20 in the textbook showed that if both functions f and g are bijective, then so is the composite function $g \circ f\colon A \to C$. Here is a nice formula for its inverse.

Theorem 8. *Let A, B, C be nonempty sets, and let $f\colon A \to B$ and $g\colon B \to C$ be functions. If f and g are invertible, then $g \circ f\colon A \to C$ is invertible and $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.*

Proof. Assume f and g are invertible. Then they are both bijective by Theorem 6 above, so $g \circ f$ is bijective by Theorem 6.20 and thus invertible.

To show that the inverse is equal to $f^{-1} \circ g^{-1}$, by Corollary 3 we just need to show that the following two compositions are identity maps:

$$\begin{aligned}(g \circ f) \circ (f^{-1} \circ g^{-1}) &= I_C,\\ (f^{-1} \circ g^{-1}) \circ (g \circ f) &= I_A.\end{aligned}$$

To prove the first identity, we compute as follows using the results of Lemma 2:

$$\begin{aligned}(g \circ f) \circ (f^{-1} \circ g^{-1}) &= g \circ \big(f \circ (f^{-1} \circ g^{-1})\big)\\ &= g \circ \big((f \circ f^{-1}) \circ g^{-1})\big)\\ &= g \circ (I_B \circ g^{-1})\\ &= g \circ g^{-1}\\ &= I_C.\end{aligned}$$

The proof of the other identity is similar:

$$\begin{aligned}(f^{-1} \circ g^{-1}) \circ (g \circ f) &= f^{-1} \circ \big(g^{-1} \circ (g \circ f)\big)\\ &= f^{-1} \circ \big((g^{-1} \circ g) \circ f\big)\\ &= f^{-1} \circ (I_B \circ f)\\ &= f^{-1} \circ f\\ &= I_A,\end{aligned}$$

which completes the proof. □

Functions as Sets of Ordered Pairs

We end this section with a brief discussion of a way of looking at functions as certain kinds of sets.

Recall that we defined a function as a pair of sets (the domain A and the codomain B), together with a "rule" that associates with each element of A one and only one element of B. You might have noticed that we were a little vague about exactly what a "rule" is supposed to be. Does it have to be a formula? Or can it be a set of formulas for different parts of the domain? Or can it be some other type of algorithm? Does such an algorithm need to produce an exact value after finitely many steps, or is something like an infinite series acceptable? Can a function assign "random" values (whatever that might mean)?

These are questions that vexed mathematicians for centuries as the foundations of calculus were being worked out. Shortly after calculus was invented, most mathematicians believed a function had to be defined by an explicit formula. Gradually, the requirements were relaxed to allow more things to be considered functions. Finally, in the mid-twentieth century, mathematicians settled on a definition that is general enough to accommodate anything anybody wanted to call a "function." It is, ultimately, a definition of a function as a kind of *set*, like just about everything else in modern mathematics. Like the definition of ordered pairs as sets (see Exercise 10 on p. 263 of the textbook), it is not very practical to work with, but it is important to know that there is such a definition, because it ensures that the theory of functions rests on the same solid logical foundation as set theory.

The key idea is that once the domain A and codomain B have been chosen, whatever techniques we might use to specify a function $f\colon A \to B$, the function is completely determined by the pairing of each element of A with a specific element of B. In other words, if we know all of the ordered pairs $(a, b) \in A \times B$, where a ranges over all elements of A and $b = f(a)$ for each such a, then we know the function. Thus a function determines a certain subset of $A \times B$, which we denote by Γ_f:

$$\Gamma_f = \{(a, b) \in A \times B : a \in A \text{ and } b = f(a)\}.$$

For example, if $f\colon \mathbb{R} \to \mathbb{R}$ is the function $f(x) = x^2$, then Γ_f is the subset $\{(x, y) \in \mathbb{R}^2 : y = x^2\}$; it is the familiar parabola that you drew when you were first studying quadratic equations.

By analogy with examples like this, for any function $f\colon A \to B$, one often calls the set Γ_f the ***graph of f***. As we have seen, every function determines a graph, which is a subset of the Cartesian product of its domain with its codomain.

The graph of a function $f\colon A \to B$ is not just an *arbitrary* subset of $A \times B$, though. From the definition of a function that we've been working with, we see that it satisfies a couple of conditions: First, for every $a \in A$, there must be an element b such that $(a, b) \in \Gamma_f$ (namely, $b = f(a)$). And second, that element must be *uniquely determined by* a; to put this a little more precisely, if b_1, b_2 are elements of B such that (a, b_1) and (a, b_2) are both in Γ_f (meaning that $f(a) = b_1$ and $f(a) = b_2$), then $b_1 = b_2$. Thus the subset $\Gamma_f \subseteq A \times B$ always satisfies these conditions:

(i) For every $a \in A$, there exists $b \in B$ such that $(a, b) \in \Gamma_f$;
(ii) For all $a \in A$ and $b_1, b_2 \in B$, if (a, b_1) and (a, b_2) are in Γ_f, then $b_1 = b_2$.

Condition (i) is sometimes summarized by saying that f is ***everywhere defined*** (meaning that it produces an output for every element of the domain), and the second by saying that f is ***uniquely defined*** (meaning that it produces only one output for each input). We say that f is ***well-defined*** if it satisfies both of these conditions. Saying that f is well-defined means nothing more nor less than "f is a function."

Be careful to note the difference between conditions (i) and (ii) above and the conditions for the function f to be surjective or injective:

(iii) f is ***surjective*** if for every $b \in B$, there exists $a \in A$ such that $(a, b) \in \Gamma_f$;
(iv) f is ***injective*** if for all $a_1, a_2 \in A$ and $b \in B$, if (a_1, b) and (a_2, b) are in Γ_f, then $a_1 = a_2$.

Every function satisfies (i) and (ii), but only some functions satisfy (iii) and/or (iv).

Now let's try to go back the other way—suppose we are given two sets A and B, and some subset $\Gamma \subseteq A \times B$. Can Γ be the graph of a function from A to B? Clearly it has to satisfy conditions (i) and (ii) above (with Γ_f replaced by Γ). But the remarkable thing is that nothing more is needed: If Γ is *any* subset of $A \times B$ that satisfies these two conditions, then we get a well-defined function $f\colon A \to B$ by declaring that for every $a \in A$, $f(a)$ is the unique element $b \in B$ such that $(a, b) \in \Gamma$.

Based on this observation, mathematicians in the mid-twentieth century reached consensus that a function should be formally *defined* as a subset of the Cartesian product satisfying these two properties. Thus here is the official definition:

Set-theoretic definition of a function: Let A and B be nonempty sets. A ***function from A to B*** is a subset $\Gamma \subseteq A \times B$ that satisfies the following properties:

(i) For every $a \in A$, there exists $b \in B$ such that $(a, b) \in \Gamma$;
(ii) For all $a \in A$ and $b_1, b_2 \in B$, if (a, b_1) and (a, b_2) are in Γ, then $b_1 = b_2$.

Because this definition is simply a description of a certain kind of set, it does not depend on any particular method for specifying how $f(a)$ is determined from a, thus resolving the ambiguities about what kinds of rules are legitimate ways to define functions.

Since a function by this definition *is* the subset $\Gamma \subseteq A \times B$, it would make sense to give the subset the same name as the function itself, and many authors do so, speaking of "a function $f \subseteq A \times B$." But in practice, mathematicians generally work with functions using the usual functional notation ($b = f(a)$), and think about functions in terms of rules for producing an output value from an input value, rather than as sets of ordered pairs. When it is useful to work directly with the set of ordered pairs, it is usually referred to as *the graph of the function* as we did above, not as the function itself.

The fine print: If a function $f : A \to B$ is invertible, then its graph has a very simple description in terms of the graph of f. You can check easily that in that case the graph of the inverse function f^{-1} is the subset of $B \times A$ given by

$$\Gamma_{f^{-1}} = \{(b, a) \in B \times A : (a, b) \in \Gamma_f\}. \tag{7}$$

In other words, the graph of f^{-1} is just the set of ordered pairs obtained by switching the first and second coordinates in all the ordered pairs in the graph of f.

Now it happens that the subset of $B \times A$ defined in (7) makes sense even if f is not invertible. This leads some authors (including Sundstrom, the author of our textbook) to define the ***inverse relation of f*** to be that subset of $B \times A$, whether f is invertible or not, and to call that set "f^{-1}." This is a dangerous practice, because that set is typically *not* a function, and does not behave in the way the actual inverse function behaves when f is invertible; so in this class (as in most mathematics books) we will never refer to the inverse relation, and we will use the notation f^{-1} *only* to refer to the inverse of an invertible function. (Full disclosure: In Section 6.6, we will see another common meaning for the symbol f^{-1} when applied to a *set*, and we will use it in that context; but we will never use it to refer to the inverse relation.)