11 Inverse functions

In this short chapter we digress back to the material of the fall semester. There, we defined functions, and showed how new functions could be formed from old, by addition, subtraction, multiplication by a constant, multiplication, division, and, most importantly, composition. In the chapter on integration, we introduced another technique to obtain a new function F from an old function f: setting $F(x) = \int_a^x f$. Before exploiting the full power of that technique, we need one more way of forming new functions from old: inverting.

11.1 Definition and basic properties

First, some background: recall the notation

$$f : A \to B,$$

standing for "f is a function with domain A, co-domain B". Precisely, this means that f is a set of pairs, with each element of A — the *domain* of f — occurring *exactly once* as a first entry of a pair, and with the set of second entries being a subset of B — a *co-domain* of f.

Here are some special kinds of functions, that come up frequently:

Injective functions $f : A \to B$ is *injective* or *one-to-one* if: no element of B appears more than once as a second entry; or, equivalently,

if $x, y \in A$ are different, then $f(x), f(y) \in B$ are different.

Such an f is also called *an injection* or *an injective map.*

Surjective functions $f : A \to B$ is *surjective* or *onto* if: every element of B appears at least once as a second entry; or, equivalently, if

for every $y \in B$ there is a (not necessarily unique) $x \in A$ with $f(x) = y$.

Another way to say this is that B is not just a co-domain for f, it is in fact the *range* — the exact set of second entries of the pairs that comprise f. Such an f is also called *a surjection* or *a surjective map.*

Bijective functions $f : A \to B$ is *bijective* if: f is both injective and surjective; or, equivalently, if

for every $y \in B$ there is a *unique* $x \in A$ with $f(x) = y$.

Such an f is also called *a bijection* or *a bijective map.*

Note that if

$$f : A \to B \tag{10}$$

is an injective function, then there is naturally associated with f a bijective function, namely

$$f : A \to R \tag{11}$$

where $R \subseteq B$ is the range of f. The f's in (10) and (11) are *the same function* — they are comprised of the same set of pairs. The only difference between them is in the notion with which they are presented.[180]

For an injective, or a bijective, function f, we can form a new function g that we can think of as "undoing the action" of f, by simply reversing all the pairs that make up f. For example, if $f(1) = 2$, i.e., $(1, 2) \in f$, then we put the pair $(2, 1)$ in g, i.e., set $g(2) = 1$. Is this really a function? It is a set of ordered pairs, certainly. Suppose that some number b appears twice in the set of ordered pairs, say as (b, a_1) and (b, a_2), with $a_1 \neq a_2$. Then that means that (a_1, b) and (a_2, b) are both in f (that's how (b, a_1) and (b, a_2) got into g). The presence of (a_1, b) and (a_2, b) in f then contradicts that f is injective.

We've just argued that if f is injective, then g is a function. On the other hand, if f is *not* injective, then the process we have described does *not* produce a new function g. Indeed, let $(a_1, b), (a_2, b) \in f$ with $a_1 \neq a_2$ witness the failure of injectivity of f. We have $(b, a_1), (b, a_2) \in g$, witnessing the failure of g to be a function.

This all shows that the process of reversing all the pairs that make up a function f, to form a new function g, makes sense if and only f is injective. We now note some further properties.

- if f is injective, then so is g. Indeed, suppose g is not injective, and let $(b_1, a), (b_2, a) \in g$ with $b_1 \neq b_2$ witness failure of injectivity of g. Then $(a, b_1), (a, b_2) \in f$ witness the failure of f to be a function, a contradiction;

- the range of f is the domain of g; the domain of f is the range of g (this is obvious);

- $f \circ g = \text{id}$, where $\text{id} : \text{Domain}(g) \to \text{Domaing}$ is the identity function, consisting of pairs whose first and second coordinates are the same; and $g \circ f = \text{id}$, where $\text{id} : \text{Domain}(f) \to \text{Domainf}$ is the identity function[181] (this should also be obvious); and

- if the operation that is used to produce g from f is applied to g (this makes sense, since g is injective), then the result is f (this should also be obvious).

[180] Which begs the question: why not just present all functions in the form $f : A \to R$, where A is the domain and R the range? The issue is that it is often very difficult to pin down the exact range of a function, so it is often convenient to simply present a definitely valid co-domain, such as $\mathbb{R}$. Try, for example, finding the exact range of $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^6 + x^5 + 1$.

[181] But note that the two identity functions here are not necessarily the same — there is no reason why $\text{Domain}(g)$ should equal $\text{Domain}(f)$.

We formalize all this, in a definition and a theorem. The theorem has already been proven; it is just the above discussion, reframed in the language of the definition.

Definition of inverse function Let $f = \{(a, b) : a \in \text{Domain} f\}$ be an injective function. The *inverse* of f, denoted f^{-1}, is defined by

$$f^{-1} = \{(b, a) : (a, b \in f)\}.$$

If f has an inverse, it is said to be *invertible*[182].

Theorem 11.1. *Let f be an injective function, with domain D and range R.*

- *f^{-1} is a function, with domain R and range D.*
- *f^{-1} is injective, and $(f^{-1})^{-1} = f$.*
- *$f \circ f^{-1}$ is the identity function on R (that is, for all $x \in R$, $f(f^{-1}(x)) = x$).*
- *$f^{-1} \circ f$ is the identity function on D.*

There's a very easy way to construct the graph of f^{-1} from the graph of f: the set of points of the form (b, a) (that comprises the graph of f^{-1}) is the reflection across the line $x = y$ of the set of points of the form (a, b) (that comprises the graph of f). Because vertical lines in the plane are mapped to horizontal lines by reflection across the line $x = y$, this leads to an easy visual test for when a function is invertible: f is invertible if it's graph passes the

Horizontal line test : every horizontal line in the plane crosses the graph of f at most once.

Which functions are invertible?

- Certainly, if f is increasing[183] on it's domain, then it is invertible (and it is an easy check that the inverse f^{-1} in this case is also increasing). On the other hand, if f is only weakly increasing, then it may not necessarily be invertible (think of the constant function).
- Similarly if f is decreasing, it's invertible, and f^{-1} is decreasing too.
- There are certainly examples of invertible functions that are *not* monotone (increasing or decreasing); consider, for example, $f : [0, 1] \to [0, 1]$ given by

$$f(x) = \begin{cases} 1 & \text{if } x = 0 \\ 0 & \text{if } x = 1 \\ x & \text{if } 0 < x < 1. \end{cases}$$

[182]For practical purposes, we can think of "invertible" and "injective" as synonymous.

[183]Recall: $f : A \to \mathbb{R}$ is *increasing* on A if $x < y \in A$ implies $f(x) < f(y)$; we sometimes say *strictly* increasing, but our convention is that without any qualification, "increasing" is the same as "strictly increasing"; we use *weakly increasing* to indicate $x < y \in A$ implies $f(x) \leq f(y)$.

- Even adding the assumption of continuity, there are still non-monotone invertible functions; consider, for example, $f : (0,1) \cup (1,2) \to \mathbb{R}$ given by

$$f(x) = \begin{cases} x & \text{if } 0 < x < 1 \\ 3 - x & \text{if } 1 < x < 2. \end{cases}$$

If f is continuous, however, *and* defined on a single interval, then it seems reasonable to expect that invertibility forces monotonicity. This is the content of our first significant theorem on invertibility.

Theorem 11.2. *Suppose that $f : I \to \mathbb{R}$ is continuous on the interval I. If f is invertible, then it is monotone (either increasing or decreasing).*

Proof: We prove the contrapositive. Suppose that f is *not* monotone on I. That means that

- there is $x_1 < x_2$ with $f(x_1) \leq f(x_2)$ (witnessing that f is not decreasing), and
- there is $y_1 < y_2$ with $f(y_1) \geq f(y_2)$ (witnessing that f is not increasing).

If either $f(x_1) = f(x_2)$ or $f(y_1) = f(y_2)$ then f is not invertible. So we may assume that in fact $f(x_1) < f(x_2)$ and $f(y_1) > f(y_2)$.

There are twelve possibilities for the relative order of x_1, x_2, y_1, y_2:

- $y_1 < y_2 < x_1 < x_2$
- $x_1 < y_1 < y_2 < x_2$
- $x_1 < x_2 < y_1 < y_2$
- $y_1 < x_1 < y_2 < x_2$
- $y_1 < x_1 < x_2 < y_2$
- $x_1 < y_1 < x_2 < y_2$
- $y_1 < x_1 < x_2 = y_2$
- $x_1 < y_1 < x_2 = y_2$
- $y_1 < x_1 = y_2 < x_2$
- $y_1 = x_1 < y_2 < x_2$
- $y_1 = x_1 < x_2 < y_2$
- $x_1 < x_2 = y_1 < y_2$

In each of these twelve cases, it is possible to find $x < y < z$ with either

- $f(y) > f(x), f(z)$

or

- $f(y) < f(x), f(z)$

For example, if $y_1 = x_1 < y_2 < x_2$, then we may take $y = y_2$, $x = x_1 = y_1$ and $z = x_2$ to get $f(y) < f(x), f(z)$. For a more involved example, consider $y_1 < y_2 < x_1 < x_2$. We have $f(y_2) < f(y_1)$. If $f(x_2) < f(y_2)$ then we may take $y = x_1$, $x = y_2$ or y_1 and $z = x_2$ to get $f(y) < f(x), f(z)$, while if $f(x_2) > f(y_2)$ we may take $y = y_2$, $x = y_1$ and $z = x_2$ to again get $f(y) < f(x), f(z)$ (note that we won't have $f(x_2) = f(y_2)$, as this automatically implies non-invertibility).

Suppose $f(y) > f(x), f(z)$. Let $m = \max\{f(x), f(z)\}$. By the intermediate value theorem applied to the interval $[x, y]$, f takes on the value $(f(y) + m)/2$ in (x, y). But by the intermediate value theorem applied to the interval $[y, z]$, f takes on the value $(f(y) + m)/2$ in (y, z). Since (x, y) and (y, z) don't overlap, this shows that f takes on the same value at least two different times, so is not invertible.

Here's an alternate, direct, proof, that uses a shorter case analysis. Suppose that f is invertible. Then it is injective, so $x \neq y$ implies $f(x) \neq f(y)$. Fix $y \in I$ that is not an endpoint; let I_1 be $\{x \in I : X < y\}$ and I_2 be $\{z \in I : z > y\}$.

It cannot be the case that there is some $x \in I_1$ with $f(x) > f(y)$, and some $x' \in I_1$ with $f(x') < f(y)$; for then we could easily find $x' < y' < z'$ with either $f(y') > f(x'), f(z')$ or $f(y') < f(x'), f(z')$, and the IVT argument from above gives a contradiction. So either $f(x) > f(y)$ for all $x \in I_1$, or $f(x) < f(y)$ for all $x \in I_1$. Similarly, either $f(z) > f(y)$ for all $z \in I_2$, or $f(z) < f(y)$ for all $z \in I_2$.

If either

- $f(x) > f(y)$ for all $x \in I_1$ and $f(z) > f(y)$ for all $z \in I_2$

or

- $f(x) < f(y)$ for all $x \in I_1$ and $f(z) < f(y)$ for all $z \in I_2$

then we could easily find $x' < y' < z'$ with either $f(y') > f(x'), f(z')$ or $f(y') < f(x'), f(z')$, for a contradiction.

If $f(x) > f(y)$ for all $x \in I_1$ and $f(z) < f(y)$ for all $z \in I_2$, then we claim that f is monotone decreasing. Indeed, consider $a < b \in I$. If one of a, b is y, we immediately have $f(a) > f(b)$. If $a < y < b$, we immediately have $f(a) > f(y) > f(b)$, so $f(a) > f(b)$. If $a, b < y$, and $f(a) < f(b)$, then we $x' < y' < z'$ with $f(y') > f(x'), f(z')$, a contradiction, so $f(a) > f(b)$ in this case. Similarly, if $a, b > y$ we also get $f(a) > f(b)$.

Finally, in the case If $f(x) < f(y)$ for all $x \in I_1$ and $f(z) > f(y)$ for all $z \in I_2$, a similar agument gives that f is monotone increasing. $\square$

A consequence of the first proof above is the following useful fact:

If $f : I \to \mathbb{R}$ is not monotone, then there $x < y < z \in I$ with either $f(y) > f(x), f(z)$ or $f(y) < f(x), f(z)$.

Note that this fact does not need continuity of f.

The proof of Theorem 11.2 uses IVT, which raises a nice challenge: in $\mathbb{Q}$-world, find an example of a function f defined on an interval, that is continuous and invertible, but not monotone. (Such an example would show that the completeness axiom is necessary to prove Theorem 11.2).

Suppose that $f : I \to \mathbb{R}$ is continuous and invertible, and so, by Theorem 11.2, monotone. We can easily determine the range of f, and so the domain of f^{-1}. The verification of all of these are left as exercises.

- if $I = [a, b]$ and f is increasing, then Range$(f) = [f(a), f(b)]$, while if f decreasing, then Range$(f) = [f(b), f(a)]$;
- if $I = (a, b)$ or $(-\infty, b)$ or (a, ∞) or $(-\infty, \infty)$, then, whether f is increasing or decreasing, we have Range$(f) = (\inf\{f(x) : x \in I\}, \sup\{f(x) : x \in I\})$ (where here we allow $\inf\{f(x) : x \in I\}$ to possibly take the value $-\infty$, and $\sup\{f(x) : x \in I\}$ to possibly take the value $-\infty$);
- and if I is a mixed interval (open at one end, closed at the other end), then we do the obvious thing: for example, if $I = [a, b)$ and f is decreasing, then Range$(f) = (\inf\{f(x) : x \in I\}, f(a)]$.

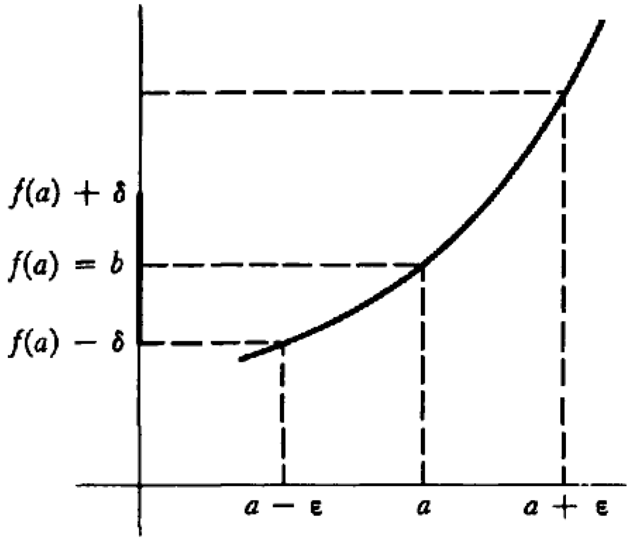
11.2 The inverse, continuity and differentiability

The inverse function behaves well with respect to continuity and differentiability, as we now show.

Theorem 11.3. *If I is an interval, and $f : I \to \mathbb{R}$ is continuous on I, and invertible, then f^{-1} is also continuous on its whole domain.*

Proof: By Theorem 11.2, we know that f is either increasing or decreasing on I. We can assume that f is increasing; if it is decreasing, we obtain the result by considering (increasing) $-f$.

Given $b \in$ Domain(f^{-1}), we want to show that $\lim_{x \to b} f^{-1}(x) = f^{-1}(b)$. Now there is $a \in$ Domain(f) with $f(a) = b$, so $f^{-1}(b) = a$. Given $\varepsilon > 0$, we want to find a $\delta > 0$ such that $f(a) - \delta < x < f(a) + \delta$ implies $a - \varepsilon < f^{-1}(x) < a + \varepsilon$. The picture below, taken from Spivak, should both make the choice of notation clear, and suggest how to proceed:



Let δ be small enough that

$$f(a-\varepsilon) < f(a) - \delta < b = f(a) < f(a) + \delta < f(a+\varepsilon)$$

(since $f(a+\varepsilon) - f(a) > 0$ and $f(a) - f(a-\varepsilon) > 0$, such a δ can be found — just take δ to be anything smaller than the minimum of $f(a+\varepsilon) - f(a)$ and $f(a) - f(a-\varepsilon)$).

For $f(a) - \delta < x < f(a) + \delta$ we have $f(a-\varepsilon) < x < f(a+\varepsilon)$ and so (using that f^{-1} is increasing) we that get $a - \varepsilon < f^{-1}(x) < a + \varepsilon$, as required. □

What about differentiability and the inverse? By considering the graph of an increasing, continuous, f, at a point $(a, f(a))$, where f is differentiable, and by then considering the reflection of the graph across $x = y$, it is fairly easy to form the hypothesis that $(f^{-1})'(f(a))$ is well-defined — unless $f'(a) = 0$, when it appears that the tangent line through $(f(a), f^{-1}(f(a)))$ is vertical.[184] In other words, it appears that for b in the domain of f^{-1}, we have that $(f^{-1})'(b)$ is well defined unless $f'(f^{-1}(b)) = 0$.

Differentiating both sides of $(f \circ f^{-1})(x) = x$ we get $f'(f^{-1}(x))(f^{-1})'(x) = 1$, that is,

$$(f^{-1})'(x) = \frac{1}{(f'(f^{-1}(x))},$$

suggesting what the derivative of f^{-1} should be at b (as long as $f'(f^{-1}(b)) \neq 0$) ("suggesting" because we don't a priori know that f^{-1} is differentiable at b).

All this can be made precise.

Theorem 11.4. *Suppose I is an interval, and that $f : I \to \mathbb{R}$ is continuous on I, and invertible. Suppose further that for some b in the domain of f^{-1}, f is differentiable at $f^{-1}(b)$.*

- *If $f'(f^{-1}(b)) = 0$ then f^{-1} is* not *differentiable at b.*

[184] Draw some graphs! Convince yourself.

- *If* $f'(f^{-1}(b)) \neq 0$ *then* f^{-1} is *differentiable at* b*, with derivative* $1/(f'(f^{-1}(b)))$.

Proof: We consider the first bullet point first: Suppose (for a contradiction) that f^{-1} *is* differentiable at b. We apply the chain rule to conclude

$$f'(f^{-1}(b))(f^{-1})'(b) = 1$$

but this is impossible since $f'(f^{-1}(b)) = 0$.

We now move on to the second bullet point. Let a be such that $f(a) = b$. We have

$$\frac{f^{-1}(b+h) - f^{-1}(b)}{h} = \frac{f^{-1}(b+h) - a}{h} = \frac{k}{h}$$

where $k = k(h)$ is such that $f^{-1}(b+h) = a + k$. We also have $f(a) + h = b + h = f(a+k)$, so the we have

$$\frac{k}{h} = \frac{k}{f(a+k) - f(a)} = \frac{1}{\left(\frac{f(a+k)-f(a)}{k}\right)}.$$

Because f^{-1} is continuous, we have $\lim_{h \to 0} f^{-1}(b+h) = f^{-1}(b) = a$, so$\lim_{h \to 0} k = 0$. So as h approaches 0, so does k, and $(f(a+k) - f(a))/k$ approaches $f'(a) = f'(f^{-1}(b))$ (which exists by hypothesis). Since (also be hypothesis) $f'(f^{-1}(b))$ is non-zero, we can put everything together to get

$$(f^{-1})'(b) = \lim_{h \to 0} \frac{f^{-1}(b+h) - f^{-1}(b)}{h} = \frac{1}{f'(f^{-1}(b))}.$$

□

What about the interaction between invertibility and integrability? Certainly, if $f : [a, b] \to \mathbb{R}$ is *continuous* and invertible (say, for simplicity, increasing), then since f^{-1} is continuous on $[f^{-1}(a), f^{-1}(b)]$, it is integrable over that range. In all the applications that are coming up, that is all that we will need. What happens to f^{-1} vis a vis integrability, when f is only assumed to be *integrable* and invertible, will be explored in an exercise.