MATH 436 Notes: Functions and Inverses.

Jonathan Pakianathan

September 12, 2003

1 Functions

Definition 1.1. *Formally, a function* $f : A \to B$ *is a subset* f *of* $A \times B$ *with the property that for every* $a \in A$*, there is a unique element* $b \in B$ *such that* $(a, b) \in f$*. The set* A *is called the domain of* f *and the set* B *is the codomain of* f*.*

While the above definition provides a definition of a function purely in terms of set theory, it is usually not a useful picture to work with. However it does emphasize the important point that the domain and codomain of a function are an intrinsic part of any function f.

Less formally, we usually think of a function $f : A \to B$ as a rule of assignment which assigns a unique output $f(a) \in B$ for every input $a \in A$. The graph of f, denoted $Graph(f) = \{(a, f(a))|a \in A\} \subset A \times B$ then recovers the more formal set theoretic definition of the function.

Definition 1.2. *Let* $f : A \to B$ *be a function.*

Given $S \subset A$ *we define* $f(S) = \{f(s)|s \in S\}$*. Note that* $f(S) \subseteq B$*.* $f(S)$ *is called the image of the set* S *under* f*.*

$f(A)$ *is called the image of* f*, and is denoted* $Im(f)$*.*

Given $T \subset B$ *we define* $f^{-1}(T) = \{a \in A|f(a) \in T\}$*. Note that* $f^{-1}(T) \subseteq A$*.* $f^{-1}(T)$ *is called the preimage of the set* T *under* f*.*

Fix a function $f : A \to B$, then it is easy to see that for all $S \subset A$, $S \subset f^{-1}(f(S))$ and for all $T \subseteq B$, we have $f(f^{-1}(T)) \subseteq T$. The next example shows that these inclusions do not have to be equalities in general.

Example 1.3. *Define* $f : \mathbb{Z} \to \mathbb{Z}$ *via* $f(n) = n^2$ *for all* $n \in \mathbb{Z}$. *Then* $Im(f) = \{0, 1, 4, 9, 16, \dots\}$, $f(\{2\}) = \{4\}$ *and* $f^{-1}(\{0, 1, 2\}) = \{0, -1, 1\}$. *It follows that* $f(f^{-1}(\{0, 1, 2\})) = \{0, 1\}$ *and* $f^{-1}(f(\{2\})) = \{-2, 2\}$.

The following are among the most important concepts involving functions, we shall see why shortly.

Definition 1.4. *Given a function* $f : A \to B$ *we say that:*
(a) f *is surjective (equivalently "onto") if* $Im(f) = B$.
(b) f *is injective (equivalently "one-to-one") if for all* $a, a' \in A$,
$(f(a) = f(a')) \implies (a = a')$.
Equivalently, f *is injective if for all* $a, a' \in A$, $(a \neq a') \implies (f(a) \neq f(a'))$.
Thus an injective function is one that takes distinct inputs to distinct outputs.
(c) f *is bijective (equivalently a "one-to-one correspondence") if it has both properties (a) and (b) above.*

These properties are important as they allow f to be inverted in a certain sense to be made clear soon.

First let us recall the definition of the composition of functions:

Definition 1.5. *If we have two functions* $f : A \to B$ *and* $g : B \to C$ *then we may form the composition* $g \circ f : A \to C$ *defined as* $(g \circ f)(a) = g(f(a))$ *for all* $a \in A$.

It is fundamental that the composition of functions is associative:

Proposition 1.6 (Associativity of composition). *Let* $f : A \to B, g : B \to C, h : C \to D$ *be functions. Then* $(h \circ g) \circ f = h \circ (g \circ f)$.

Proof. The two expressions give functions from A to C. To show they are equal we only need to show they give the same output for every input $a \in A$. Computing we find:

$$((h \circ g) \circ f)(a) = (h \circ g)(f(a)) = h(g(f(a))).$$

$$(h \circ (g \circ f))(a) = h((g \circ f)(a)) = h(g(f(a))).$$

Since the two expressions agree, this completes the proof. □

Definition 1.7. *For any set* S, *the identity function on* S, *denoted by* $1_S : S \to S$ *is the function defined by* $1_S(s) = s$ *for all* $s \in S$.

It is easy to check that if we have a function $f : A \to B$ *then*

$$f \circ 1_A = f = 1_B \circ f.$$

Now we are ready to introduce the concept of inverses:

Definition 1.8. *Suppose we have a pair of functions $f : A \to B$ and $g : B \to A$ such that $g \circ f = 1_A$. Then we say that f is a right inverse for g and equivalently that g is a left inverse for f.*

The following is fundamental:

Theorem 1.9. *If $f : A \to B$ and $g : B \to A$ are two functions such that $g \circ f = 1_A$ then f is injective and g is surjective. Hence a function with a left inverse must be injective and a function with a right inverse must be surjective.*

Proof. $g \circ f = 1_A$ is equivalent to $g(f(a)) = a$ for all $a \in A$.

Showing f is injective: Suppose $a, a' \in A$ and $f(a) = f(a') \in B$. Then we may apply g to both sides of this last equation and use that $g \circ f = 1_A$ to conclude that $a = a'$. Thus f is injective.

Showing g is surjective: Let $a \in A$. Then $f(a) \in B$ and $g(f(a)) = a$. Thus $a \in g(B) = Im(g)$ for all $a \in A$ showing $A = Im(g)$ so g is surjective. □

The following example shows that left (right) inverses need not be unique:

Example 1.10. *Let $A = \{1, 2\}$ and $B = \{a, b, c\}$. Define $f_1, f_2 : A \to B$ by $f_1(1) = a, f_1(2) = c, f_2(1) = b, f_2(2) = c$. Define $g_1, g_2 : B \to A$ by $g_1(a) = g_1(b) = 1, g_1(c) = 2$ and $g_2(a) = 1, g_2(b) = g_2(c) = 2$. Then it is an easy exercise to show that $g_1 \circ f_1 = 1_A = g_1 \circ f_2$ so that g_1 has two distinct right inverses f_1 and f_2. Furthermore since g_1 is not injective, it has no left inverse.*

Similarly one can compute, $g_2 \circ f_1 = 1_A$ so that f_1 has two distinct left inverses g_1 and g_2. Furthermore since f_1 is not surjective, it has no right inverse.

From this example we see that even when they exist, one-sided inverses need not be unique.

However we will now see that when a function has both a left inverse and a right inverse, then all inverses for the function must agree:

Lemma 1.11. *Let $f : A \to B$ be a function with a left inverse $h : B \to A$ and a right inverse $g : B \to A$. Then $h = g$ and in fact any other left or right inverse for f also equals h.*

Proof. We have that $h \circ f = 1_A$ and $f \circ g = 1_B$ by assumption. Using associativity of function composition we have:

$$h = h \circ 1_B = h \circ (f \circ g) = (h \circ f) \circ g = 1_A \circ g = g.$$

So h equals g. Since this argument holds for any right inverse g of f, they all must equal h. Since this argument holds for any left inverse h of f, they all must equal g and hence h. So all inverses for f are equal. □

We finish this section with complete characterizations of when a function has a left, right or two-sided inverse.

Proposition 1.12. *A function $f : A \to B$ has a left inverse if and only if it is injective.*

Proof. $\implies$: Follows from Theorem 1.9. $\impliedby$: If $f : A \to B$ is injective then we can construct a left inverse $g : B \to A$ as follows. Fix some $a_0 \in A$ and define

$$g(b) = \begin{cases} a \text{ if } b \in Im(f) \text{ and } f(a) = b \\ a_0 \text{ otherwise} \end{cases}$$

Note this defines a function only because there is at most one a with $f(a) = b$. It is an easy computation now to show $g \circ f = 1_A$ and so g is a left inverse for f.

□

Proposition 1.13. *A function $f : A \to B$ has a right inverse if and only if it is surjective.*

Proof. $\implies$: Follows from Theorem 1.9. $\impliedby$: Suppose $f : A \to B$ is surjective. Then for each $b \in B$, $f^{-1}(\{b\})$ is a nonempty subset of A. Thus by the Axiom of Choice we may construct a "choice" function $g : B \to A$ such that $g(b)$ is a choice of element from the nonempty set $f^{-1}(\{b\})$ for all $b \in B$. It is easy now to compute that $f \circ g = 1_B$ and so g is a right inverse for f. □

Proposition 1.14. *A function $f : A \to B$ has a two-sided inverse if and only if it is bijective. In this case, the two-sided inverse will be unique and is usually denoted by $f^{-1} : B \to A$.*

Proof. First note that a two sided inverse is a function $g : B \to A$ such that $f \circ g = 1_B$ and $g \circ f = 1_A$. $\Longrightarrow$: Theorem 1.9 shows that if f has a two-sided inverse, it is both surjective and injective and hence bijective. $\Longleftarrow$: Now suppose f is bijective. From the previous two propositions, we may conclude that f has a left inverse and a right inverse. By Lemma 1.11 we may conclude that these two inverses agree and are a two-sided inverse for f which is unique. Alternatively we may construct the two-sided inverse directly via $f^{-1}(b) = a$ whenever $f(a) = b$. □