Problem Set #8b – Solutions

Mike Greene
Math 25

Problem 1.

Let $f : [a, b] \to [c, d]$ (where $a, b, c, d \in \mathbb{R}$) be continuous and strictly increasing with $f(a) = c$ and $f(b) = d$.

Claim 1.1. **(a)** There exists a unique continuous function $f : [c, d] \to [a, b]$ satisfying

$$f \circ g = Id : [c, d] \to [c, d]$$

$$g \circ f = Id : [a, b] \to [a, b]$$

We will first prove that f is bijective. f is surjective by the intermediate value theorem: for any $y \in [c, d]$, $\exists x \in [a, b]$ s.t. $f(x) = y$. To demonstrate injectivity, we simply recall that f is strictly increasing, and thus for any $x_1 < x_2$, we have $f(x_1) < f(x_2)$.

Because f is bijective, we know that $\forall y \in [c, d]$, $f^{-1}(y)$ has exactly one element. We then see that any g satisfying the composition criteria above must also satisfy $g(y) \in f^{-1}(y)$. Because $f^{-1}(y)$ has exactly one element, such a g is unique and well defined. We note then that, by construction, if $x \in f^{-1}(y)$:

$$f(g(y)) = f(x) = y \text{ and}$$

$$g(f(x)) = g(y) = x$$

Finally, we must prove that this g is continuous, which we will do by proving that the pre-image of an open set is open. First, let $(a_0, b_0) \subset [a, b]$ be a nonempty open interval. Then by the fact that f is increasing, $f((a_0, b_0)) \subset (f(a_0), f(b_0))$, and by the intermediate value theorem $f((a_0, b_0)) \supset (f(a_0), f(b_0))$. Thus $f((a_0, b_0)) = (f(a_0), f(b_0))$ and $g^{-1}((a_0, b_0)) = (f(a_0), f(b_0))$ which is open. Now, any open $U \subset [c, d]$ can be expressed as the union of a set of open intervals: any $p \in U$ is contained in some open interval, $U_p \subset U$, and

$$g^{-1}(U) = g^{-1}\left(\cup_{p \in U} U_p\right) = \cup_{p \in U} g^{-1}(U_p)$$

which, being the union of open sets is open.

NOTE: Those of you who have studied some additional point-set topology will notice that we proved a function was continuous by proving that the pre-image of every element of the topological basis of its image space is open.

□

Claim 1.2. **(b)** $g(y)$ can be found by solving $y - f(x) = 0$ for x by bisection.

Let $y \in (c, d)$. Because we know that g is well defined, we can define $x := g(y)$. We must now show that x can be *found* by solving $y - f(x) = 0$ by bisection.

Recalling the definition of bisection in Hubbard (pp. 218-9 in my edition), we define a sequence of potential solutions, x_i, using bisection. Intuitively, if we know that x is in a particular interval, then we guess that it is at the midpoint of the interval. If our guess proves high, we know that x is in the lower have of the interval, and if our guess proves low, we know that x is in the higher half of the interval. We then repeat the process, calling the our guesses x_i. *(For those of you with a computer science background, this is analogous to a "binary search" on an* *ordered list.)* Formally, we define the sequence inductively, using two additional sequences, a_i and b_i:

Base Case: $a_0 := 0$ and $b_0 := 0$.

Induction Step: If $a_i = b_i$, then we define $x_i := a_{i+1} := b_{i+1} := a_i$. Otherwise, define $x_i := \frac{a+b}{2}$. If $f(x_i) = y$, then we define $a_{i+1} := b_{i+1} := x_i$. If $f(x_i) < y$, we define $a_{i+1} := x_i$ and $b_{i+1} = b_i$. If $f(x_i) > y$, we define $a_{i+1} = a_i$ and $b_{i+1} = x_1$. We note that $b_{i+q} - a_{i+1} \leq \frac{1}{2}(b_i - a_i)$ and

$$(a_i \leq x \leq b_i) \Rightarrow (a_{i+1} \leq x \leq b_{i+1})$$

due to the fact that f is strictly increasing.

In order to prove that $y - f(x) = 0$ is solvable by bisection we must show that $x_i \to x$. By the above, we know that for any give i, $x, x_i \in [a_i, b_i]$ and thus

$$|x - x_i| \leq b_i - a_i \leq 2^{-i}(b - a)$$

We thus see that for any $\epsilon > 0$ we can select $N := \lceil \log_2 \frac{(b-a)}{\epsilon} \rceil$ and then be sure that whenever $i > N$, we will have $|x_i - x| \leq \epsilon$

□

Claim 1.3. **(c)** If f is differentiable at $x \in (a, b)$, and $f'(0) \neq 0$, then g is differentiable at $f(x)$, and $g'(f(x)) = \frac{1}{f'(x)}$.

It is clear that $g \circ f$ is differentiable (in particular, its derivative is everywhere 1). Consequently, we employ the formal definition of a derivative, and then derive like mad (employing the fact that $f'(x) \neq 0$ in the third line):

$$\begin{aligned}
(g \circ f)'(x) &= \lim_{h \to 0} \frac{g(f(x+h)) - g(f(x))}{h} \\
1 &= \lim_{h \to 0} \frac{g\left(f(x) + h\frac{f(x+h)-f(x)}{h}\right) - g(f(x))}{h} \\
\frac{1}{f'(x)} &= \left(\frac{1}{f'(x)}\right) \cdot \lim_{h \to 0} \frac{g\left(f(x) + hf'(x)\right) - g(f(x))}{h} \\
&= \lim_{h \to 0} \frac{g\left(f(x) + hf'(x)\right) - g(f(x))}{hf'(x)} \\
&= \lim_{h \to 0} \frac{g\left(f(x) + h\right) - g(f(x))}{h} \\
&= g'(f(x))
\end{aligned}$$

□

Problem 2.

Claim 2.1. "2×2 matrices near the identity have square roots." Formally, the function $f : \mathrm{Mat}(2,2) \to \mathrm{Mat}(2,2)$ given by $f : M \mapsto MM$ is invertible in a neighborhood of I.

Our proof will be an application of the inverse function theorem. We begin by calculating $Df(I)$. We see that

$$f : \begin{bmatrix} a & b \\ c & d \end{bmatrix} \mapsto \begin{bmatrix} a^2 + bc & ab + bd \\ ac + cd & bc + d^2 \end{bmatrix}$$

Because all of the entries are polynomials, we know that f is continuous and

$$Df\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \mathfrak{J}f\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} \frac{\partial}{\partial a}\left(a^2+bc\right) & \frac{\partial}{\partial b}\left(a^2+bc\right) & \frac{\partial}{\partial c}\left(a^2+bc\right) & \frac{\partial}{\partial d}\left(a^2+bc\right) \\ \frac{\partial}{\partial a}\left(ab+bd\right) & \frac{\partial}{\partial b}\left(ab+bd\right) & \frac{\partial}{\partial c}\left(ab+bd\right) & \frac{\partial}{\partial d}\left(ab+bd\right) \\ \frac{\partial}{\partial a}\left(ac+cd\right) & \frac{\partial}{\partial b}\left(ac+cd\right) & \frac{\partial}{\partial c}\left(ac+cd\right) & \frac{\partial}{\partial d}\left(ac+cd\right) \\ \frac{\partial}{\partial a}\left(bc+d^2\right) & \frac{\partial}{\partial b}\left(bc+d^2\right) & \frac{\partial}{\partial c}\left(bc+d^2\right) & \frac{\partial}{\partial d}\left(bc+d^2\right) \end{bmatrix}$$

$$= \begin{bmatrix} 2a & c & b & 0 \\ b & a+d & 0 & b \\ c & 0 & a+d & c \\ 0 & c & b & 2d \end{bmatrix}$$

and thus

$$Df\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

which is clearly invertible (e.g. its determinant is 16). Consequently, f satisfies the criteria for the inverse function theorem at I and is therefore invertible around I, implying that matrices "near" I have square roots.

NOTE: When we calculate its Jacobian, we are thinking of f as a map $\mathbb{R}^4 \to \mathbb{R}^4$, thus we should expect that $Df(I)$ is a 4×4 matrix.

□