6-15-2020

An inverse function theorem converse

Jimmie Lawson

Louisiana State University

AN INVERSE FUNCTION THEOREM CONVERSE

JIMMIE LAWSON

Abstract. We establish the following converse of the well-known inverse function theorem. Let $g : U \to V$ and $f : V \to U$ be inverse homeomorphisms between open subsets of Banach spaces. If g is differentiable of class C^p and f if locally Lipschitz, then the Fréchet derivative of g at each point of U is invertible and f must be differentiable of class C^p.

Primary 58C20; Secondary 46B07, 46T20, 46G05, 58C25

Key words and phrases. Inverse function theorem, Lipschitz map, Banach space, chain rule

1. Introduction

A general form of the well-known inverse function theorem asserts that if g is a differentiable function of class C^p, $p \geq 1$, between two open subsets of Banach spaces and if the Fréchet derivative of g at some point x is invertible, then locally around x, there exists a differentiable inverse map f of g that is also of class C^p. But in various settings, one may have the inverse function f readily at hand and want to know about the invertibility of the Fréchet derivative of g at x and whether f is of class C^p. Our purpose in this paper is to present a relatively elementary proof of this converse result under the general hypothesis that the inverse f is (locally) Lipschitz. Simple examples like $g(x) = x^3$ at $x = 0$ on the real line show that the assumption of continuity alone is not enough. Thus it is a bit surprising that the mild strengthening to the assumption that the inverse is locally Lipschitz suffices.

Helpful tools for the task at hand have been developed in the intense study of Lipschitz functions in the Banach space setting motivated by Rademacher's theorem concerning the existence of an abundance of points of differentiability of Lipschitz mappings in the setting of euclidean spaces. In particular we recall in the next

Date: December, 2018.

section a useful generalization of the chain rule by O. Maleva and D. Preiss [4]. In Sections 3 and 4 we present our main results and provide some follow-up illustrative material in Section 5.

For a comprehensive reference to the inverse function and implicit function theorems and related theory, we refer the reader to [1].

2. A general version of the chain rule

In this section we recall some notions of differentiability of Lipschitz functions between open subsets of a Banach spaces and a generalized chain rule from the work of Maleva and Preiss [4]. This material will be crucial to the derivation of our main results in the next section.

Suppose that Y and Z are Banach spaces, U is a nonempty open subset of Y, $f : U \to Z$, and $y \in U$, $v \in Y$. Recall that $\lim_{t\to 0^+}(f(y+tv)-f(y))/t$, if it exists, is called the *one-sided directional derivative of f at y in the direction v*. Similarly if $\lim_{t\to 0}(f(y+tv)-f(y))/t$ exists, it is called the (*bilateral*) *directional derivative of f at y in the direction v*. If the directional derivative of f at y in the direction v exists for all $v \in Y$, then the mapping from Y to Z sending v to its directional derivative is, by definition, the *Gâteaux derivative* (by some authors the mapping is also required to be a continuous linear map).

Maleva and Preiss [4] have given the following generalization of the one-sided directional derivative.

Definition 2.1. The *derived set of f at the point y in the direction of v* is defined as the set $\mathcal{D}f(y,v)$ consisting of all existing limits $\lim_{n\to\infty}(f(y+t_n v)-f(y))/t_n$, where $t_n \searrow 0$. The *δ-approximating derived set of f at y in the direction of v* is defined, for $\delta > 0$, by

$$\mathcal{D}_\delta f(y,v) = \left\{ \frac{f(y+tv)-f(y)}{t} : 0 < t < \delta \right\}.$$

Remark 2.2. It is easy to see that

$$\mathcal{D}f(y,v) = \bigcap_{\delta>0} \overline{\mathcal{D}_\delta f(y,v)} \tag{2.1}$$

In general the derived set may be empty, a single point, or multi-valued. If $\mathcal{D}f(y,v)$ is a single point, then it is the one-sided directional derivative of f at y in the direction

v, and we denote it by $f'_+(y,v)$. If the directional derivative of f at y in the direction v exists, it is denoted $f'(y,v)$.

If f is Gâteaux differentiable at y, then the Gâteaux derivative at y is equal to $f'(y,\cdot) = \mathcal{D}f(y,\cdot)$.

We recall a general version of the chain rule from [4, Corollary 2.6].

Proposition 2.3. *Suppose X, Y are Banach spaces, $x \in U$, an open subset of X, and V is an open subset of Y. If $g : U \to V$ is continuous and has a one-sided directional derivative at x in the direction of v and $f : V \to Z$ is Lipschitz, then*

$$\mathcal{D}(f \circ g)(x,v) = \mathcal{D}f(g(x), g'_+(x,v)).$$

Proof. Let $a \in \mathcal{D}(f \circ g)(x,v)$. Then there exists a sequence $t_n \to 0^+$ such that

$$a = \lim_n \frac{f \circ g(x + t_n v) - f \circ g(x)}{t_n}.$$

Let κ be a Lipschitz constant for f on V. Let $\varepsilon > 0$ and choose N such that $\|(g(x+t_n v) - g(x))/t_n - g'_+(x,v)\| < \varepsilon/\kappa$ for $n \geq N$. We then note for $n \geq N$

$$\begin{aligned}
\left\| \frac{f \circ g(x + t_n v) - f \circ g(x)}{t_n} \right. & - \left. \frac{f(g(x) + t_n g'_+(x,v)) - f(g(x))}{t_n} \right\| \\
&= \left\| \frac{f(g(x+t_n v)) - f(g(x) + t_n g'_+(x,v))}{t_n} \right\| \\
&\leq \frac{\kappa}{t_n} \|g(x+t_n v) - g(x) - t_n g'_+(x,v)\| \\
&= \kappa \left\| \frac{g(x+t_n v) - g(x)}{t_n} - g'_+(x,v) \right\| \leq \kappa(\varepsilon/\kappa) = \varepsilon.
\end{aligned}$$

It thus follows that the sequence $[f(g(x) + t_n g'_+(x,v)) - f(g(x))]/t_n$ also converges to a, so $a \in \mathcal{D}f(g(x), g'_+(x,v))$. The argument is reversible so the equality claimed in the proposition holds.

□

3. A Converse of the Inverse Function Theorem

The inverse function theorem asserts the existence of a local inverse if the derivative is invertible. In this section we derive a converse result: a Lipschitz continuous local inverse implies an invertible derivative.

In this section we work in the following setting. Let X, Y be Banach spaces, let U and V be nonempty open subsets of X and Y resp., each equipped with the restricted metric from the containing Banach space, and let $g : U \to V$ and $f : V \to U$ be inverse homeomorphisms.

Definition 3.1. For $g : U \to V$, the *Gâteaux derivative of g at $x \in U$* is defined by $d_x^G g(v) = g'_+(x, v)$ for all $v \in X$, provided such one-sided directional derivatives exist for all $v \in X$ and the resulting map $d_x^G g : E \to F$ is a continuous linear map.

Lemma 3.2. *Let $x \in U$, $y = f(x) \in V$. Assume that g has a Gâteaux derivative $d_x^G g : E \to F$ at x and assume that f is Lipschitz. Then $\mathcal{D}f(g(x), \cdot) \circ d_x^G g$ is the identity map on X and the Gâteaux derivative $d_x^G g$ is injective.*

Proof. Since $f \circ g$ on U is the identity map, it follows directly from the definition of $\mathcal{D}$ that $\mathcal{D}(f \circ g)(x, \cdot)$ is the identity map on X. Hence by Proposition 2.3 for any $v \in X$,

$$v = \mathcal{D}(f \circ g)(x, v) = \mathcal{D}f(g(x), g'_+(x, v)) = \mathcal{D}f(g(x), d_x^G(v)).$$

This equality of the left-hand and right-hand sides of the equation yields the two concluding assertions. □

Lemma 3.2 yields the invertibility of the the Gâteaux derivative in the finite dimensional setting.

Corollary 3.3. *If X, Y are both finite dimensional and g is Gâteaux differentiable at $x \in U$, then d_x^G is invertible.*

Proof. Since g and f are homeomorphisms, X and Y must have the same dimension. Hence the linear map d_x^G is a linear isomorphism if and only if it injective, which is the case by Lemma 3.2. □

The infinite dimensional case requires more work and stronger hypotheses.

Lemma 3.4. *Let $x \in U$, $y = g(x) \in V$. Assume that g is Fréchet differentiable at x and that f is Lipschitz on V. Then the image of X under the Fréchet derivative $dg_x : X \to Y$ is dense in Y.*

Proof. Let M_f be a Lipschitz constant for f on V. Let $w \in Y$. Pick $\tau > 0$ small enough so that $y + tw \in V$ for $0 < t < \tau$. We set $z_t = \frac{f(y+tw)-f(y)}{t}$ and note from the Lipschitz condition that

$$\|z_t\| = \left\|\frac{f(y+tw)-f(y)}{t}\right\| \leq \frac{1}{t}M_f\|(y+tw)-y\| = M_f\|w\|. \tag{3.2}$$

We divide the remainder of the proof into steps.
Step 1: $(g(x+tz_t)-g(x))/t = w$ for $0 < t < \tau$.

$$g(x+tz_t)-g(x) = g\left(f(y)+t\left(\frac{f(y+tw)-f(y)}{t}\right)\right) - y = gf(y+tw) - y = tw,$$

so $(g(x+tz_t)-g(x))/t = w$.
Step 2: *For $\varepsilon > 0$ there exists $t < \tau$ such that $\|(g(x+tz_t)-g(x))/t - dg_x(z_t)\| < \varepsilon$.*
From the Fréchet differentiability of g at x, the Fréchet derivative dg_x satisfies

$$\lim_{u\to 0}\frac{\|g(x+u)-g(x)-dg_x(u)\|}{\|u\|} = 0.$$

For $\varepsilon > 0$ pick $\delta > 0$ such that

$$\frac{\|g(x+u)-g(x)-dg_x(u)\|}{\|u\|} < \frac{\varepsilon}{M_f\|w\|} \text{ whenever } 0 < \|u\| < \delta.$$

Pick $t > 0$ such that $y + tw \in V$ and $tM_f\|w\| < \delta$. We conclude from inequality (3.2) and the preceding that

$$\begin{aligned}\left\|\frac{g(x+tz_t)-g(x))-dg_x(tz_t)}{t}\right\| &= \|z_t\|\left\|\frac{g(x+tz_t)-g(x))-dg_x(tz_t)}{\|tz_t\|}\right\| \\ &< M_f\|w\|\frac{\varepsilon}{M_f\|w\|} = \varepsilon\end{aligned}$$

Using the linearity of of the Fréchet derivative dg_x, we obtain $dg_x(tz_t) = tdg_x(z_t)$, which allows us to rewrite the first entry in the preceding string to obtain

$$\left\|\frac{g(x+tz_t)-g(x)}{t} - dg_x(z_t)\right\| < \varepsilon,$$

which establishes the Step 2.

We note that combining Steps 1 and 2 yields

$$\|w - dg_x(tz_t)\| \leq \left\|w - \frac{g(x+tz_t)-g(x)}{t}\right\| + \left\|\frac{g(x+tz_t)-g(x)}{t} - dg_x(z_t)\right\| \leq \varepsilon.$$

Since $w \in Y$ and $\varepsilon > 0$ were chosen arbitrarily, this completes the proof. ∎

We come now to a central result of the paper, what we are calling a converse of the inverse function theorem.

Theorem 3.5. *Let X and Y be Banach spaces, let U and V be nonempty open subsets of X and Y resp., and let $g : U \to V$ and $f : V \to U$ be inverse homeomorphisms. Let $x \in U$ and $y = g(x) \in V$. Let g be Fréchet differentiable at x and let f be Lipschitz continuous on V. Then $dg_x(\cdot) : X \to Y$ is an isomorphism.*

Proof. By Lemma 3.2 the Fréchet derivative dg_x is injective, hence a linear isomorphism onto its image $Z = g(X)$, a subspace of Y, and has inverse $\mathcal{D}f(y,\cdot) : Z \to X$. (In particular in this setting for $w \in Z$ it must be the case that $\mathcal{D}f(y,w)$ is a singleton, which we could write alternatively as $f'_+(y,w)$.)

Let M_f be the Lipschitz constant for f on V. By equation (3.2) every member of $\mathcal{D}_\delta f(y,w)$ is bounded in norm by $M_f\|w\|$, and hence the same is true for $\mathcal{D}f(y,w) = \bigcap_{\delta>0} \overline{\mathcal{D}_\delta f(y,w)}$. We conclude that $\mathcal{D}f(y,\cdot) : Z \to X$ is Lipschitz for the Lipschitz constant M_f.

By Lemma 3.4 Z is dense in Y. Thus the linear Lipschitz map $\mathcal{D}f(y,\cdot) : Z \to X$ extends uniquely to a linear Lipschitz map (hence a bounded linear operator) from Y to X. Label $dg_x = \Gamma_X$, $\mathcal{D}f(y,\cdot) = \Gamma_Z$ and its extension to Y by Γ_Y. We note for $0 \neq z \in Z$,

$$\|z\| = \|\Gamma_X(\Gamma_Z(z)\| \leq \|\Gamma_X\|\|\Gamma_Z(z)\|,$$

so $(1/\|\Gamma_X\|)\|z\| \leq \|\Gamma_Z(z)\|$. By continuity of the norm and density of Z in Y this inequality carries over from Γ_Z to its extension Γ_Y. But this extended inequality implies Γ_Y has a trivial kernel, which means that Γ_Y is injective. But if Z is proper in Y, then this is impossible, since $\Gamma_Z(Z) = X$. Hence $Z = Y$ and $\Gamma_Z = \Gamma_Y$ is an inverse for dg_x. □

4. A Global Result

We can use Theorem 3.5 to derive useful results for studying differentiable functions with locally Lipschitz inverses, in particular for deriving differentiability properties of the inverse. In the following C^p means have continuous derivatives through order p for p a positive integer, have continuous derivatives of all order for $p = \infty$, and being analytic (having locally power series expansions) for $p = \omega$.

Theorem 4.1. *Let X and Y be Banach spaces, let U and V be nonempty open subsets of X and Y resp., and let $g : U \to V$ and $f : V \to U$ be inverse homeomorphisms. Assume further that g is of class C^p on U for some $p \geq 1$ and that f is locally Lipschitz on V. Then g and f are inverse diffeomorphisms of class C^p.*

Proof. We fix $x \in U$ and $y = f(x) \in V$. We apply Theorem 3.5 to small enough neighborhoods of x and y so that f is Lipschitz to see that the hypotheses of the standard inverse function theorem (in the Banach space setting) are satisfied. The conclusions of this theorem then follow locally from the inverse function theorem. (See, for example, [5, Theorem 1.23] for the C^ω-case.) In this way we obtain that the conclusions of the theorem hold locally for each $x \in X$ and hence hold globally. □

We remark that the preceding results can be generalized to the setting of Banach manifolds. In order to obtain the preceding results in this setting one needs Lipschitz charts (between the manifold metric and the Banach space metric) at each point and then the preceding results readily extend to this more general setting.

5. An Application

We sketch in this section one setting in which our general converse of the inverse function theorem can be fruitfully applied and briefly consider a specific example.

Suppose we are given some equation of the form $F(x, y) = 0$, where x, y belong to some given open subset of a Banach space. In some cases it might be possible to solve the equation for y in terms of x, i.e., $y = g(x)$, which can seen to be of class C^p. Let's suppose further that g has an inverse given by $x = f(y)$, but no corresponding explicit description of this function. If it can be shown, however, that f is (locally) Lipschitz, then we use the results of the preceding section to show that f is also of class C^p.

Example 5.1. Let $\mathcal{B}(H)$ be the C^*-algebra of bounded linear operators on the Hilbert space H. We consider the Banach space $\mathbb{S}$ of hermitian operators and the open cone of positive invertible hermitian operators $\mathbb{P}$. In recent years a useful notion of a multivariable geometric mean Λ on $\mathbb{P}$ has arisen [2], [3], generally called the Karcher mean. One useful characterization of this n-variable mean $\Lambda(A_1, \ldots, A_n)$ for $A_1, \ldots, A_n \in \mathbb{P}$

is that it is the unique solution X of the equation

$$\log(X^{-1/2}A_1X^{-1/2}) + \cdots + \log(X^{-1/2}A_nX^{-1/2}) = 0.$$

If we fix $A_1, \ldots, A_{n-1}$, let $Y = A_n$, and the left-hand side of the equation be $F(X, Y)$, then we are in the general setting of the previous paragraph with $U = V = \mathbb{P}$. One can rather easily solve $F(X, Y)$ for Y and see it is an analytic function g of X, but not conversely. However the inverse f is given by $f(Y) = \Lambda(A_1, \ldots, A_{n-1}, Y)$. The (local) Lipschitz property is a basic property of Λ, so the earlier results yield X, the geometric mean, as an analytic function of Y, or alternatively we can say Λ is an analytic function of each of its variables. This turns out to be an important property of Λ. This type of analysis can be applied to other important operator means.

References

[1] S. Krantz and H. Parks, The Implicit Function Theorem. History, Theory, and Applications. Reprint of the 2003 edition. Modern Birkhäuser Classics. Birkhäuser/Springer, New York, 2013.

[2] J. Lawson and Y. Lim, Weighted means and Karcher equations of positive operators, Proc. Natl. Acad. Sci. USA. **110** (2013), 15626-15632.

[3] J. Lawson and Y. Lim, Karcher means and Karcher equations of positive operators, Trans. Amer. Math. Soc. Series B **1** (2014), 1-22.

[4] O. Maleva and D. Preiss, Directional one-sided derivatives and the chain rule formula for locally Lipschitz functions on Banach spaces, Trans. Amer. Math. Soc. **368** (2016), 4685-4730.

[5] H. Upmeier, Symmetric Banach Manifolds and Jordan C^*-Algebras, Mathematical Studies 104, North Holland, Amsterdam, 1985.

Department of Mathematics, Louisiana State University, Baton Rouge, LA 70803, USA

E-mail address: `lawson@math.lsu.edu`