The Inverse Function Theorem

The Inverse Function Theorem says that if X and Y are Banach spaces and $U \subseteq X$, $F : U \to Y$ is C^1 and $DF(x_0)$ is invertible, then in a small neighborhood of x_0, the function F is one-to-one, so that an inverse function G exists and that inverse function is also C^1 on some neighborhood of $F(x_0)$.

The following two lemmas use the operator norm to guarantee the invertibility of an operator-valued function. Roughly, the first one says that if a bounded linear operator is close enough to an invertible operator, then it is also invertible. The second lemma states that if a continuous operator-valued function has values that are invertible (operators), then the map to the inverses is also continuous.

Lemma 1. *Let $T, S : X \to Y$ be linear maps between Banach spaces X to Y. If T is (boundedly) invertible and $\|T - S\| < \frac{1}{\|T^{-1}\|}$, then S is invertible.*

Proof. Let $G = T^{-1} + \sum_{n=1}^{\infty} T^{-1}((T - S)T^{-1})^n$ then by assumption the series on the right converges in operator norm, because $\|T^{-1}(T - S)\| \leq \|T^{-1}\|\|T - S\| < 1$. We show that the bounded linear operator G is the inverse of S.

With G as defined, it can be verified that

$$G = T^{-1} + T^{-1}(T - S)G\,.$$

Consequently, left multiplying by T gives

$$TG = I + (T - S)G$$

from which we get $0 = I - SG$, so G is a right inverse to S. Similarly, $G = T^{-1} + G(T - S)T^{-1}$ yields that G is a left inverse of S. □

Lemma 2. *Let $T : U \to B(X, Y)$, be a map from an open subset $U \subset X$ of a Banach space X to $B(X, Y)$, the space of bounded operators mapping X to a Banach space Y such that T is continuous with respect to the norms on X and $B(X, Y)$, and for each $u \in U$, $T(u)$ has a bounded inverse $(T(u))^{-1}$, then $S : U \to B(Y, X), S(u) = (T(u))^{-1}$ is continuous.*

Proof. We know from the preceding lemma that for each $x, u \in U$ with $\|x - u\| \leq \frac{1}{2\|(T(u))^{-1}\|}$,

$$S(x) = (T(u))^{-1} + \sum_{n=1}^{\infty} (T(u))^{-1}((T(u) - T(x))(T(u))^{-1})^n$$

and the convergence is uniform in $B_r(u)$ with $r = \frac{1}{2\|(T(u))^{-1}\|}$, so S is continuous at each $u \in U$. $\square$

Theorem 3 (Inverse Function Theorem)**.** *Let $U \subseteq X$ be an open subset of a Banach space X, let $x_0 \in U$, and let $F : U \to Y$ be C^1 on U. If $DF(x_0)$ is invertible, then*

a. there exist open sets V and W in U and Y, respectively, such that the restriction $F|_V$ is a homeomorphism from V to W,

b. if $G = F^{-1}$ denotes the inverse function, then G is C^1 on W and $G'(y) = (F')^{-1}(G(y))$.

Proof. It is enough to show this for the special case $x_0 = F(x_0) = 0$, otherwise we shift the domain and values of F by appropriate constants.

Let for $x_0 = 0$, $L = DF(0)$ and fix a number λ, $0 < \lambda < 1$.

Step 1. Compare F with its linearization. From F being continuously differentiable, we can choose $\rho > 0$ such that for all $x \in B_\rho(0)$,

$$\|DF(x) - L\| \leq \frac{\lambda}{\|L^{-1}\|}.$$

Let $H(x) = F(x) - Lx$, then for $x \in B_\rho(0)$, by a consequence of the mean-value theorem and the assumption $F(0) = 0 = L0$,

$$\|H(x)\| = \|F(x) - Lx\| \leq \frac{\lambda}{\|L^{-1}\|}\|x\|.$$

More generally, we have for $x_1, x_2 \in B_\rho(0)$,

$$\|H(x_2) - H(x_1)\| \leq \frac{\lambda}{\|L^{-1}\|}\|x_2 - x_1\|.$$

Step 2. Invert F using a contraction mapping. Next, we claim that if $\|y\| < r \equiv \frac{(1-\lambda)\rho}{\|L^{-1}\|}$, then there is a unique $x \in \overline{B}_\rho(0)$ with $F(x) = y$. This will be established with the help of the Contraction Mapping Theorem. To this end we define a function $\Phi_y : \overline{B}_\rho(0) \to X$ by

$$\Phi_y(x) = L^{-1}(y - H(x)) = x + L^{-1}(y - F(x)).$$

From the definition of Φ_y, we have $\Phi_y(x) = x$ if and only if $F(x) = y$.

Using the bound for $\|H(x)\|$ and the triangle inequality, for $x \in B_\rho(0)$ and $y \in B_r(0)$

$$\|\Phi_y(x)\| \leq \lambda\|x\| + \|L^{-1}\|\|y\| \leq \lambda\|x\| + \|L^{-1}\|r.$$

Choosing $\lambda < 1/2$ and ρ as above and $r \leq \rho/(2\|L^{-1}\|)$ then implies that Φ_y maps $\overline{B}_\rho$ to $\overline{B}_\rho$.

Moreover, given $x_1, x_2 \in \overline{B}_\rho$, integrating $\varphi'(t) = DF(x_1+t(x_2-x_1))(x_2-x_1)$ we have by the Fundamental Theorem of Calculus

$$F(x_2) - F(x_1) = \int_0^1 DF(x_1 + t(x_2 - x_1))(x_2 - x_1)dt$$

which implies

$$\begin{aligned}\|\Phi_y(x_2) - \Phi_y(x_1)\| &\leq \|L^{-1}\| \left\| L(x_2 - x_1) - \int_0^1 \varphi'(t)dt \right\| \\ &\leq \|L^{-1}\| \int_0^1 \|(\varphi'(t) - L(x_2 - x_1)\|dt\end{aligned}$$

and again by the continuity of DF,

$$\|\Phi_y(x_2) - \Phi_y(x_1)\| \leq \lambda \|x_2 - x_1\|.$$

Thus, Φ_y is a contraction mapping that recovers $x = F^{-1}(y)$ as its unique fixed point.

From this, we conclude that the map $G : y \mapsto F^{-1}(y)$ is one-to-one when we restrict its domain to $B_r(0)$.

Step 3. Show F restricts to a homeomorphism. Next, we show G is continous, which proves that $F : V \to W$ is a homeomorphism with $V = F^{-1}(W)$ and $W = B_r(0)$. By definition, given $y_1, y_2 \in B_r(0)$, we have unique x_i with $F(x_i) = y_i$ and

$$\begin{aligned}\|G(y_2) - G(y_1)\| &= \|x_2 - x_1\| \\ &= \|\Phi_{y_2}(x_2) - \Phi_{y_1}(x_1)\| \\ &\leq \|L^{-1}\|\|y_2 - H(x_2) - y_1 + H(x_1)\| \\ &\leq \|L^{-1}\|\|y_2 - y_1\| + \lambda\|x_2 - x_1\| \\ &= \|L^{-1}\|\|y_2 - y_1\| + \lambda\|G(y_2) - G(y_1)\|\end{aligned}$$

from which we get Lipschitz continuity,

$$\|G(y_2) - G(y_1)\| \leq \frac{1}{1-\lambda}\|L^{-1}\|\|y_2 - y_1\|\,.$$

Finally, we show G is C^1. We know for $y, y+h \in B_r(0)$,

$$\begin{aligned}h &= F(G(y+h)) - F(G(y)) \\ &= \int_0^1 DF(G(y) + t(G(y+h) - G(y)))dt(G(y+h) - G(y))\,.\end{aligned}$$

Let $L_{G(y)} = DF(G(y))$. Applying $L_{G(y)}^{-1}$ to both sides yields

$$G(y+h) - G(y) = L_{G(y)}^{-1}h + R$$

with

$$R = L_{G(y)}^{-1} \int_0^1 \left(L_{G(y)} - L_{G(y)+t(G(y+h)-G(y))}\right) dt (G(y+h) - G(y)) \, .$$

By the continuity of G and F being C^1, the integrand can be made arbitrarily small for sufficiently small $\|h\|$. Hence, we have

$$\frac{\|G(y+h) - G(y) - L_{G(y)}^{-1}h\|}{\|h\|} \to 0$$

which shows that G is differentiable, and that the derivative of G is $DG(y) = L_{G(y)}^{-1} = (DF(G(y)))^{-1}$.

Now applying the preceding lemma to the function $T : u \mapsto L_u$ gives that $S : u \mapsto (L_u)^{-1}$ is continuous, and so is $DG : y \mapsto L_{G(y)}^{-1}$. □