The Inverse Function Theorem

This is a rather terse – but complete – proof of the inverse function theorem. It is a combination of proofs, mostly from Rudin.

Theorem: Assume that $f : \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable on $B(x_0, r)$, and $f'(x_0)$ is invertible. Then

(i) there is an s, $0 < s \le r$, such that f is one-to-one on $B(x_0, s)$, and

(ii) the image, V, of $B(x_0, s)$ under f is an open set.

In view of (i) we can define the inverse function g on V so that $g(y) \in B(x_0, s)$ and $f(g(y)) = y$. Then

(iii) g is continuously differentiable.

Proof: This proof is divided into three parts corresponding to statements (i)-(iii) in the theorem.

Part 1: This makes use of the following auxiliary function:

$$\phi(x) = x + (f'(x_0))^{-1}(y - f(x)).$$

Note that we are already using the hypothesis that $f'(x_0)$ is invertible. Think of ϕ as a function of x with y as a parameter. Note that $\phi(x) = x$ is equivalent to $f(x) = y$. Note also that $\phi'(x) = (f'(x_0))^{-1}(f'(x_0) - f'(x))$. Since $f'(x)$ is continuous at x_0, there is an $s > 0$ such that $f'(x)$ is invertible and

$$||\phi'(x)|| \le ||(f'(x_0))^{-1}||||f'(x_0) - f'(x)|| \le \frac{1}{2} \tag{1}$$

for $x \in B(x_0, s)$. Hence for $x_1, x_2 \in B(x_0, s)$ we have

$$|\phi(x_1) - \phi(x_2)| \le \frac{1}{2}|x_1 - x_2|. \tag{2}$$

Thus, for any choice of y, the equation $\phi(x) = x$ has at most one solution in $B(x_0, s)$. In other words f is one-to-one on $B(x_0, s)$, and we have proven part (i) of the theorem. Note that we cannot invoke the Contraction Mapping theorem here because we do not know that ϕ maps $B(x_0, s)$ into itself (it probably doesn't). [This proof is taken from Rudin's Principles of Mathematical Analysis.]

Part 2: We need to show that V is open. Let y_1 be a vector in V. By definition there is an $x_1 \in B(x_0, s)$ such that $f(x_1) = y_1$. Choose t so that the sphere $|x - x_1| = t$ is contained in $B(x_0, s)$. Consider $l(x) = |f(x) - y_1|$ as a function on the sphere $|x - x_1| = t$. Since $f(x_1) = y_1$, Part 1 implies that $l(x)$ is never zero on this sphere. Since $l(x)$ is continuous, we can conclude that the minimum value m of $l(x)$ on the sphere is positive. Then for $|y - y_1| < m/3$

$$|f(x) - y| \ge |f(x) - y_1| - |y - y_1| \ge \frac{2m}{3}. \tag{3}$$

Now let $h(x) = |f(x) - y|^2$ on the closed ball $\overline{B(x_1, t)}$. We know $h(x_1) = |y - y_1|^2 < m^2/9$ and $h(x) \ge 4m^2/9$ on the boundary of the ball. So h must assume its minimum value on the closed ball at an interior point z where $\nabla h = 0$. However, that means

$$0 = \frac{\partial h}{\partial x_j}(z) = 2\sum_{k=1}^{n} \frac{\partial f_k}{\partial x_j}(z)(f_k(z) - y_k),\ j = 1, ..., n.$$

This says $(f'(z))^T(f(z) - y) = 0$. If $f(z) \neq y$, this implies that $(f'(z))^T$ is not invertible, and hence $f'(z)$ is not invertible. That contradicts the invertibility of $f'(x)$ for $x \in B(x_0, s)$, so we must conclude that $f(z) = y$. Thus $B(y_1, m/3) \subset V$ and V is open, proving (2). [This proof has been circulating around the department.]

Part 3: We begin by showing that g is differentiable at all points of V. Given $y \in V$, $y + k \in V$ for $|k|$ sufficiently small. Then there will be a unique x and h such that $x \in B(x_0, s)$, $x + h \in B(x_0, s)$, $f(x) = y$ and $f(x + h) = y + k$. From the Chain Rule $g'(y)$ *ought* to be the inverse of $f'(g(y))$. Set $H = (f'(g(y)))^{-1}$ – remember that this inverse exists because $g(y) = x \in B(x_0, s)$. We have

$$g(y+k) - g(y) - Hk = x + h - x - H(f(x+h) - f(x)) = -H(f(x+h) - f(x) - f'(x)h).$$

[Note that only in the final equality do we use what H is.] Since f is differentiable at x, we see that

$$|g(y + k) - g(y) - Hk| \leq \delta(h)|h|,$$

where $\delta(h)$ goes to zero as h goes to zero. So what we need to show now is $|h| \leq C|k|$ for some C independent of h. Recall ϕ:

$$\phi(x + h) - \phi(x) = h - (f'(x_0))^{-1}k.$$

So (2) implies

$$|h - (f'(x_0))^{-1}k| \leq \frac{1}{2}|h|.$$

The triangle inequality gives

$$\frac{1}{2}|h| \leq |(f'(x_0))^{-1}k| \text{ and we have } |h| \leq 2||(f'(x_0))^{-1}|||k|$$

which is exactly what we need. Thus g is differentiable at all points of V and hence continuous. Moreover, $g'(y) = (f'(g(y)))^{-1}$ which is continuous. Thus g is continuously differentiable and we have (iii). [This argument is also from Rudin's, Principles of Mathematical Analysis.]