MAT 372
Spring, 2005

NOTES ON THE INVERSE FUNCTION THEOREM

We will follow the text's approach in using the contraction mapping theorem to prove the inverse function theorem. However, we will use the operator norm to make the necessary estimates. In addition, we will find it useful to have an analogue of the alternate characterization of differentiability in the case of a continuously differentiable function. We present this analogue first. We state it for real-valued functions. For functions taking values in $\mathbf{R}^m$ the same result holds — the function ϕ defined in the lemma will then be $\mathbf{R}^m$-valued also.

Lemma. Let $U \subseteq \mathbf{R}^n$ be an open set, and let $f : U \to \mathbf{R}$ be a continuously differentiable function. Define $\phi : U \times U \to \mathbf{R}$ by

$$\phi(x, y) = \begin{cases} \frac{f(y)-f(x)-f'(x)(y-x)}{|y-x|}, & \text{if } x \neq y, \\ 0, & \text{if } x = y. \end{cases}$$

Then ϕ is continuous on $U \times U$.

Proof. It is clear that ϕ is continuous at points of the form (x, y) with $x \neq y$. Let's consider continuity at (x_0, x_0) for some $x_0 \in U$. Let $r > 0$ with $B_r(x_0) \subseteq U$. For x, $y \in B_r(x_0)$ with $x \neq y$, the mean-value theorem gives a point z on the segment between x and y such that $f(y) - f(x) = f'(z)(y - x)$. Then

$$\begin{aligned} \phi(x, y) &= \frac{f(y) - f(x) - f'(x)(y - x)}{|y - x|} \\ &= \frac{f'(z)(y - x) - f'(x)(y - x)}{|y - x|} \\ &= \big(f'(z) - f'(x)\big)\frac{y - x}{|y - x|} \\ &\to 0, \quad \text{as } x,\ y \to x_0, \end{aligned}$$

since f' is continuous and $(y - x)/|y - x|$ is bounded. ∎

We now turn to the proof of the inverse function theorem. The point of the theorem is that if the derivative (of a continuously differentiable function) at a certain point x_0 is invertible, then *near* x_0 the function itself is invertible — in other words, the derivative controls the behavior of the function locally. For a function on $\mathbf{R}$, this is easy to prove — it comes down to the fact that the continuous image of an interval is again an interval, and hence that a small interval around x_0 is mapped onto an interval around $f(x_0)$. For a function on $\mathbf{R}^n$ for $n > 1$, the analogous fact is not so easy — why should the image under f of a ball centered at x_0 contain a ball centered at $f(x_0)$? We will use the modification of Newton's method described in class (and the text) to prove that this happens.

Theorem. (Inverse function theorem) Let $U \subseteq \mathbf{R}^n$ be an open set, and let $f : U \to \mathbf{R}^n$ be a continuously differentiable function. Let $a \in U$ and suppose that $f'(a)$ is an invertible linear map. Let $b = f(a)$. Then there are r, $s > 0$ such that for every $y \in B_s(b)$ there exists a unique $x \in B_r(a)$ such that $f(x) = y$. Define $g : B_s(b) \to B_r(a)$ by the equation $f\big(g(y)\big) = y$. Then g is continuously differentiable on $B_s(b)$, and $g'(y) = f'\big(g(y)\big)^{-1}$.

Proof. Fix k with $0 < k < 1$ (k will be the contraction constant). Choose $r > 0$ so that

$$\text{(i)} \quad \overline{B}_r(a) \subseteq U$$

$$\text{(ii)} \quad \|f'(z) - f'(w)\| < \frac{k}{2\|f'(a)^{-1}\|}, \quad z,\, w \in \overline{B}_r(a)$$

$$\text{(iii)} \quad \|f'(z)^{-1}\| < 2\|f'(a)^{-1}\|, \quad z \in \overline{B}_r(a)$$

$$\text{(iv)} \quad \big|\phi(z, w)\big| < \frac{k}{2\|f'(a)^{-1}\|}, \quad z,\, w \in \overline{B}_r(a).$$

Choose $s > 0$ so that

$$\text{(v)} \quad s < \frac{r}{2\|f'(a)^{-1}\|}.$$

Now fix $y \in B_s(b)$. We will prove that $y = f(x)$ for some $x \in B_r(a)$. For this we define a mapping $\lambda : \overline{B}_r(a) \to \overline{B}_r(a)$ by

$$\lambda(z) = z - f'(a)^{-1}\big(f(z) - y\big).$$

Note that $f(x) = y$ if and only if $\lambda(x) = x$. Thus the point x we are searching for must be a fixed point of λ. To prove that such a point x exists, we must show that λ is a contraction mapping of the compact (hence complete) set $\overline{B}_r(a)$. There are two parts to the proof: first we show that λ maps this set into itself, and then we show that it is a contraction mapping with contraction constant k.

Claim 1. $\lambda\big(\overline{B}_r(a)\big) \subseteq \overline{B}_r(a)$.

Proof of claim 1: Let $z \in \overline{B}_r(a)$. We have

$$\begin{aligned}
\big|\lambda(z) - a\big| &= \big|z - a - f'(a)^{-1}\big(f(z) - y\big)\big| \\
&\le \Big|f'(a)^{-1}\Big(f'(a)(z - a) - f(z) + y\Big)\Big| \\
&\le \|f'(a)^{-1}\|\Big|f'(a)(z - a) - f(z) + f(a) + y - b\Big| \\
&= \|f'(a)^{-1}\|\Big|-\phi(a, z)|z - a| + y - b\Big| \\
&\le \|f'(a)^{-1}\|\Big(\big|\phi(a, z)\big||z - a| + \big|y - b\big|\Big) \\
&< \|f'(a)^{-1}\|\left(\frac{k|z - a|}{2\|f'(a)^{-1}\|} + s\right) \\
&< \|f'(a)^{-1}\|\left(\frac{1 \cdot r}{2\|f'(a)^{-1}\|} + \frac{r}{2\|f'(a)^{-1}\|}\right) \\
&= r.
\end{aligned}$$

Note also that we have actually proved that $\lambda\big(\overline{B}_r(a)\big) \subseteq B_r(a)$. Therefore any fixed point of λ in $\overline{B}_r(a)$ must actually lie in $B_r(a)$.

Claim 2. For z, $w \in \overline{B}_r(a)$ we have $\big|\lambda(z) - \lambda(w)\big| \le k|z-w|$.

Proof of claim 2: Let z, $w \in \overline{B}_r(a)$. Then

$$\begin{aligned}
\big|\lambda(z) - \lambda(w)\big| &= \Big|\Big(z - f'(a)^{-1}\big(f(z)-y\big)\Big) - \Big(w - f(a)^{-1}\big(f(w)-y\big)\Big)\Big| \\
&= \big|(z-w) - f'(a)^{-1}\big(f(z)-f(w)\big)\big| \\
&\le \|f'(a)^{-1}\|\Big|f'(a)(z-w) - \big(f(z)-f(w)\big)\Big| \\
&= \|f'(a)^{-1}\|\Big|f'(a)(z-w) - f'(w)(z-w) - \phi(w,z)|w-z|\Big| \\
&\le \|f'(a)^{-1}\|\left(\Big|\big(f'(a)-f'(w)\big)(z-w)\Big| + \big|\phi(w,z)\big||z-w|\right) \\
&\le \|f'(a)^{-1}\|\Big(\big\|f'(a)-f'(w)\big\| + \big|\phi(w,z)\big|\Big)|z-w| \\
&< \|f'(a)^{-1}\|\left(\frac{k}{2\|f'(a)^{-1}\|} + \frac{k}{2\|f'(a)^{-1}\|}\right)|z-w| \\
&= k|z-w|.
\end{aligned}$$

Therefore λ is a contraction mapping of $\overline{B}_r(a)$, and so it has a unique fixed point, which lies in $B_r(a)$, as we already observed. Therefore there is a unique point $g(y) \in B_r(a)$ such that $f\big(g(y)\big) = y$. This defines the function $g : B_s(b) \to B_r(a)$. To finish the proof we must show that g is continuously differentiable.

Claim 3. g is continuous.

Proof of claim 3: Let y_1, $y_2 \in B_s(b)$. We have

$$\begin{aligned}
y_2 - y_1 &= f\big(g(y_2)\big) - f\big(g(y_1)\big) \\
&= f'\big(g(y_1)\big)\big(g(y_2)-g(y_1)\big) + \phi\big(g(y_1),g(y_2)\big)\big|g(y_2)-g(y_1)\big|.
\end{aligned}$$

Applying $f'\big(g(y_1)\big)^{-1}$

to both sides gives

$$\begin{aligned}
(*)\quad f'\big(g(y_1)\big)^{-1}(y_2-y_1) &= g(y_2) - g(y_1) \\
&\quad + f'\big(g(y_1)\big)^{-1}\Big(\phi\big(g(y_1),g(y_2)\big)\Big)\big|g(y_2)-g(y_1)\big|, \\
\big|f'\big(g(y_1)\big)^{-1}(y_2-y_1)\big| &\ge \big|g(y_2)-g(y_1)\big| \\
&\quad - \big\|f'\big(g(y_1)\big)^{-1}\big\| \cdot \big|\phi\big(g(y_1),g(y_2)\big)\big| \cdot \big|g(y_2)-g(y_1)\big| \\
&\ge \big|g(y_2)-g(y_1)\big| \\
&\quad - 2\big\|f'(a)^{-1}\big\| \cdot \frac{k}{2\|f'(a)^{-1}\|} \cdot \big|g(y_2)-g(y_1)\big| \\
(**)\quad &= (1-k)\big|g(y_2)-g(y_1)\big\|.
\end{aligned}$$

It follows that as $y_2 \to y_1$ we get $g(y_2) \to g(y_1)$, so g is continuous.

We now use the continuity of g in a bootstrapping argument to prove that g is differentiable. In fact, if g is differentiable, we can appy the chain rule to the equation $f\big(g(y)\big) = y$ to get $f'\big(g(y)\big) \cdot g'(y) = I$, so that $g'(y) = f'\big(g(y)\big)^{-1}$. Therefore we use this linear map to verify the definition of the derivative.

Claim 4. g is continuously differentiable.

Proof of claim 4: We first use the inequality $(**)$ from the proof of claim 3:

$$\begin{aligned} \big|g(y_2) - g(y_1)\big| &\le \frac{1}{1-k}\Big|f'\big(g(y_1)\big)^{-1}(y_2 - y_1)\Big| \\ &\le \frac{1}{1-k}\Big\|f'\big(g(y_1)\big)^{-1}\Big\| \cdot |y_2 - y_1|. \end{aligned} \tag{$***$}$$

We now use equation $(*)$ from the proof of claim 3. Rearranging that equation gives:

$$g(y_2) - g(y_1) - f'\big(g(y_1)\big)^{-1}(y_2 - y_1) = -f'\big(g(y_1)\big)^{-1}\Big(\phi\big(g(y_1), g(y_2)\big)\Big)\big|g(y_2) - g(y_1)\big|.$$

Therefore we have

$$\begin{aligned} &\frac{\Big|g(y_2) - g(y_1) - f'\big(g(y_1)\big)^{-1}(y_2 - y_1)\Big|}{|y_2 - y_1|} \\ &\qquad \le \frac{\Big\|f'\big(g(y_1)\big)^{-1}\Big\| \cdot \big|\phi\big(g(y_1), g(y_2)\big)\big| \cdot \big|g(y_2) - g(y_1)\big|}{|y_2 - y_1|} \\ &\qquad \le \frac{\Big\|f'\big(g(y_1)\big)^{-1}\Big\|^2}{1-k}\big|\phi\big(g(y_1), g(y_2)\big)\big|, \text{ by } (***) \\ &\qquad \to 0 \text{ as } y_2 \to y_1, \end{aligned}$$

since g and ϕ are continuous. ∎