(*November 20, 2014*)

Inverse function theorems

Paul Garrett garrett@math.umn.edu http://www.math.umn.edu/˜garrett/

[This document is
http://www.math.umn.edu/˜garrett/m/complex/notes_2014-15/05c_inverse_function.pdf]

1. Fixed-point lemma
2. Smooth inverse function theorem
3. Holomorphic inverse function theorem
4. Perturbations $f(z) + h \cdot g(z)$

---

1. *Fixed-point lemma*

[1.0.1] Lemma: Let X be a complete metric space with distance function d. Let $f : X \to X$ be a continuous map *uniformly contractive* in the sense that there is $0 < c < 1$ so that $d(fx, fy) \le c \cdot d(x, y)$ for all $x, y \in X$. Then f has a unique *fixed point*: there is a unique $x \in X$ with $f(x) = x$. *Further*, $\lim_{n\to\infty} f^n y = x$ for any $y \in X$.

Proof: First, for any $y \in X$, by repeated application of the triangle inequality,

$$d(y, f^n(y)) \le d(y, f(y)) + d(f(y), f^2(y)) + d(f^2(y), f^3(y)) \ldots + d(f^{n-1}(y), f^n(y))$$

$$\le (1 + c + c^2 + \ldots + c^{n-1}) \cdot d(y, f(y)) < \frac{d(y, f(y))}{1 - c}$$

Next, claim that for any $y \in X$ the sequence $y, f(y), f^2(y), \ldots$ is Cauchy. Indeed, for $n_o \le m \le n$, using the previous inequality,

$$d(f^m(y), f^n(y)) \le c^{n_o} \cdot d(f^{m-n_o}(y), f^{n-n_o}(y)) \le c^{n_o} \cdot c^{m-n_o} \cdot d(y, f^{n-m}(y)) \le c^{n_o} \cdot \frac{d(y, f(y))}{1 - c}$$

This goes to 0 as $n_o \to +\infty$, so the sequence is Cauchy.

Similarly, for any y, z in X, with $m \le n$

$$d(f^m(y), f^n(z)) \le c^m \cdot d(y, f^{n-m}(z)) \le c^m \cdot \Big(d(y, z) + d(z, f^{n-m}(z))\Big) \le c^m \cdot \Big(d(y, z) + \frac{d(z, f(z))}{1 - c}\Big)$$

which goes to 0 as $m \to +\infty$. The limit is the same for y unchanged but z arbitrary. Thus, $z, f(z), f^2(z), \ldots$ has limit x for all $z \in X$. Further, taking $z = x$, $f^n(x) \to x$. Given $\varepsilon > 0$, take n_o large enough so that $d(f^n(x), x) < \varepsilon$ for $n \ge n_o$. For $n \ge n_o$,

$$d(f(x), x) \le d(f(x), f^{n+1}(x)) + d(f^{n+1}(x), x) < c \cdot d(x, f^n(x)) + \varepsilon < (c + 1) \cdot \varepsilon$$

This holds for all $\varepsilon > 0$, so $f(x) = x$. ///

2. *Smooth inverse function theorem*

The *derivative* γ' of a smooth function $\gamma : [a,b] \to U \subset \mathbb{R}^n$ is the usual

$$\gamma'(t) \;=\; \lim_{h\to 0} \frac{\gamma(t+h)-\gamma(t)}{h}$$

For this section, the *derivative* f' of an $\mathbb{R}^n$-valued function on an open $U \subset \mathbb{R}^n$ is the n-by-n-matrix-valued function so that for every smooth path $\gamma : [a,b] \to U$

$$(f\circ\gamma)'(t) \;=\; f'(\gamma(t))\cdot\gamma'(t) \qquad\qquad \text{(matrix multiplication)}$$

Equivalently, for small real h, $x_o \in U$, and $v \in \mathbb{R}^n$, as $h \to 0$, using Landau's little-oh notation,[1]

$$f(x_o + h\cdot v) \;=\; f(x) + h\cdot f'(x_o)\cdot v + o(h) \qquad\qquad \text{(matrix multiplication)}$$

[2.0.1] Theorem: Let U be an open subset of $\mathbb{R}^n$ and $f : U \to \mathbb{R}^n$ a continuously differentiable function. For $x_0 \in U$ such that $f'(x_0) : \mathbb{R}^n \to \mathbb{R}^n$ is a linear isomorphism, there is a neighborhood $V \subset U$ of x_0 so that $f|_V$ has a continuously differentiable inverse on $f(V)$.

Proof: Let $x \to |x|$ be the usual norm on $\mathbb{R}^n$, and $|T|$ the operator norm[2] on n-by-n real matrices. Without loss of generality, $x_o = 0$, $f(x_o) = 0$, and $f'(x_o) = f'(0) = 1_n$. Let $F(x) = x - f(x)$, so that $F'(0) = 0$. By continuity, there is $\delta > 0$ so that $|F'(x)| < \frac{1}{2}$ for $|x| < \delta$.

With $g(t) = F(tx)$ for $t \in [0,1]$, the Mean Value Theorem in one variable gives

$$F(x) \;=\; g(1) \;=\; g(0) + g'(t)(1-0) \;=\; F(0) + F'(tx)(x) \;=\; F'(tx)(x) \qquad\qquad \text{(for some } 0 \le t \le 1)$$

so

$$|F(x)| \;\le\; |F'(tx)|\cdot|x| \;\le\; \tfrac{1}{2}\cdot|x| \;\le\; \frac{\delta}{2} \qquad\qquad \text{(for } |x| < \delta)$$

Thus, F maps the closed ball B_δ of radius δ to the closed ball $B_{\delta/2}$.

We claim that $f(B_\delta) \supset B_{\delta/2}$, and that f is *injective* on $f^{-1}(B_{\delta/2}$. To this end, take $y \in B_{\delta/2}$, and let $\Phi_y(x) = y + F(x) = y + x - f(x)$. For $|y| \le \delta/2$ and $|x| \le \delta$, $|\Phi_y(x)| \le \delta$, so Φ_y is a continuous map of the complete metric space B_δ to itself. A similar estimate shows that Φ_y is *contractive*: letting $g(t) = F((1-t)x_1 + tx_2)$,

$$|\Phi_y(x_2) - \Phi_y(x_1)| \;=\; |F(x_2) - F(x_1)| \;=\; |g(1) - g(0)| \;=\; |g'(t)|\cdot|1-0|$$

$$=\; |F'((1-t)x_1 + tx_2)|\cdot|x_2 - x_1| \;\le\; \tfrac{1}{2}\cdot|x_1 - x_2| \qquad\qquad \text{(for given } x_1, x_2 \in B_\delta, \text{ for some } 0 \le t \le 1)$$

By the fixed-point lemma, Φ_y has a unique fixed point x_o, that is,

$$x_o \;=\; \Phi_y(x_o) \;=\; y + x_o - f(x_o)$$

so x_o is the unique solution in B_δ to the equation $f(x_o) = y$. This proves $f(B_\delta) \supset B_{\delta/2}$ as well as the injectivity on $f^{-1}(B_{\delta/2})$.

[1] When $f(x)/g(x) \to 0$ as $x \to x_o$, write $f(x) = o(g(x))$.

[2] The usual *operator norm* is $|T| = \sup_{|x|\le 1} |Tx|$.

To prove differentiability of the inverse map $\varphi = f^{-1}$, take $x_1, x_2 \in B_\delta$. Continuity of φ follows from

$$|x_1 - x_2| \;\le\; |f(x_1) - f(x_2)| + \big|\big(x_1 - f(x_1)\big) - \big(x_2 - f(x_2)\big)\big|$$

$$\le\; |f(x_1) - f(x_2)| + |F(x_1) - F(x_2)| \;\le\; |f(x_1) - f(x_2)| + \tfrac{1}{2}|x_1 - x_2|$$

by the inequality $|F(x_1) - F(x_2)| < \frac{1}{2}|x_1 - x_2|$ from above. Subtracting $\frac{1}{2}|x_1 - x_2|$ from both sides,

$$\tfrac{1}{2}|x_1 - x_2| \;\le\; |f(x_1) - f(x_2)|$$

giving continuity of the inverse.

For differentiability, let $y_1 = f(x_1)$ and $y_2 = f(x_2)$ with y_1, y_2 in the interior of $B_{\delta/2}$. Then

$$\varphi(y_1) - \varphi(y_2) - f'(x_2)^{-1}(y_1 - y_2) \;=\; x_1 - x_2 - f'(x_2)^{-1}(f(x_1) - f(x_2))$$

$$=\; x_1 - x_2 - f'(x_2)^{-1}\Big(f'(x_2)(x_1 - x_2) + o(x_1 - x_2)\Big) \qquad \text{(as } x_1 \to x_2\text{)}$$

$$=\; x_1 - x_2 - (x_1 - x_2) + o(x_1 - x_2) \;=\; o(x_1 - x_2) \qquad \text{(as } x_1 \to x_2\text{)}$$

By the already-established continuity, this is $o(y_1 - y_2)$. Thus, the inverse φ is differentiable at $y_2 = f(x_2)$, and its derivative is $\varphi'(y_2) = f'(x_2)^{-1}$, for $|y| < \delta/2$. ///

[2.0.2] Remark: An elaboration of this discussion proves higher-order continuous differentiability in the real-variables sense, but we do not need this for application to the holomorphic inverse function theorem below.

---

3. *Holomorphic inverse function theorem*

Now we return to *complex* differentiability.

[3.0.1] Theorem: For f holomorphic on a neighborhood U of z_o and $f'(z_o) \neq 0$, there is a holomorphic inverse function g on a neighborhood of $f(z_o)$, that is, such that $(g \circ f)(z) = z$ and $(f \circ g)(z) = z$.

Proof: The idea is to consider f as a real-differentiable map $f : \mathbb{R}^2 \to \mathbb{R}^2$, obtain a real-differentiable inverse g and then observe that *complex* differentiability of f implies that of g.

The complex differentiability of f can be expressed as

$$f(z_o + hw) \;=\; f(x_o) + hf'(z_o)\cdot w + o(h) \qquad \text{(small real } h\text{, complex } w\text{)}$$

where $f'(z_o)\cdot w$ denotes multiplication in $\mathbb{C}$. Separate real and imaginary parts: let $f'(z_o) = a + bi$ with $a, b \in \mathbb{R}$, and $w = u + iv$ with $u, v \in \mathbb{R}$, giving

$$f(z_o + hw) \;=\; f(x_o) + h(a + bi)\cdot(u + iv) + o(h) \;=\; f(x_o) + h\big((au - bv) + i(av + bu)\big) + o(h)$$

The multiplication in $\mathbb{C}$ is achieved by matrix multiplication of real and imaginary parts:

$$\begin{pmatrix} au - bv \\ av + bu \end{pmatrix} \;=\; \begin{pmatrix} a & -b \\ b & a \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix}$$

so the real-variable derivative of f at z_o is

$$\begin{pmatrix} \mathrm{Re}\, f'(z_o) & -\mathrm{Im}\, f'(z_o) \\ \mathrm{Im}\, f'(z_o) & \mathrm{Re}\, f'(z_o) \end{pmatrix}$$

The real-variable derivative has determinant $|f'(z_o)|^2$, so is invertible for $f'(z_o) \neq 0$. Let $\alpha = f'(z_o)$. Thus, there exists a real-differentiable inverse g, with real-variable derivative at $f(z_o)$ given by

$$\begin{pmatrix} \mathrm{Re}\,\alpha & -\mathrm{Im}\,\alpha \\ \mathrm{Im}\,\alpha & \mathrm{Re}\,\alpha \end{pmatrix}^{-1} \;=\; \frac{1}{|\alpha|^2} \begin{pmatrix} \mathrm{Re}\,\alpha & \mathrm{Im}\,\alpha \\ -\mathrm{Im}\,\alpha & \mathrm{Re}\,\alpha \end{pmatrix} \;=\; \begin{pmatrix} \mathrm{Re}\,(\alpha^{-1}) & -\mathrm{Im}\,(\alpha^{-1}) \\ \mathrm{Im}\,(\alpha^{-1}) & \mathrm{Re}\,(\alpha^{-1}) \end{pmatrix}$$

That is, with $w_o = f(z_o)$,

$$g(w_o + h(u+iv)) \;=\; g(w_o) + h(1\ i) \begin{pmatrix} \mathrm{Re}\,(\alpha^{-1}) & -\mathrm{Im}\,(\alpha^{-1}) \\ \mathrm{Im}\,(\alpha^{-1}) & \mathrm{Re}\,(\alpha^{-1}) \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} + o(h) \;=\; g(w_o) + h\alpha^{-1}(u+iv) + o(h)$$

This holds for all real u, v, so g is *complex*-differentiable at $f(z_o)$, with complex derivative $1/f'(z_o)$. ///

---

4. *Perturbations* $f(z) + h \cdot g(z)$

[4.0.1] Corollary: For f, g holomorphic near z_o, with z_o a simple zero of $f(z_o)$, for all $\varepsilon > 0$ there is $\delta > 0$ such that $f - h \cdot g$ has a zero z_h with $|z_o - z_h| < \varepsilon$, and z_h is a holomorphic function of h.

Proof: In the anomalous case that $g(z_o) = 0$, then $z_h = z_o$ suffices.

For $g(z_o) \neq 0$, solve $f(z) + h \cdot g(z) = 0$ for h:

$$h \;=\; \frac{-f(z)}{g(z)}$$

and then

$$h' \;=\; \frac{-f'(z)}{g(z)} - \frac{f(z) \cdot g'(z)}{g(z)^2}$$

and

$$h'(z_o) \;=\; \frac{-f'(z_o)}{g(z_o)} - \frac{0 \cdot g'(z)}{g(z_o)^2} \;=\; \frac{-f'(z_o)}{g(z_o)} \;\neq\; 0$$

Apply the holomorphic inverse function theorem to obtain the holomorphic inverse $F(h) = z$ such that $F(0) = z_o$. ///
