06. Implicit and inverse functions theorems

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

[This file is http://www.math.umn.edu/~garrett/m/complex/notes_2020-21/06_fp_inverse_implicit.pdf]

1. Contractive-map fixed-point lemma
2. Inverse function theorem
3. Implicit function theorem
4. Real-differentiability versus complex-differentiability
5. Holomorphic inverse function theorem
6. Holomorphic implicit function theorem

1. Contractive-map fixed-point lemma

This is central to many existence and uniqueness results.

Let $X$ be a complete metric space with distance function $d$. A continuous $f : X \to X$ is uniformly contractive when there is $0 < c < 1$ so that

$$d(f(x), f(y)) \leq c \cdot d(x, y) \quad (\text{for all } x, y \in X)$$

A fixed point of a map $f : X \to X$ is $x \in X$ such that $f(x) = x$. A point in $X$ is an attractor for $f$ when

$$\lim_{n \to \infty} f^n y = x \quad (\text{for all } y \in X)$$

Visibly, if an attractor exists, it is unique.

[1.1] Lemma: A uniformly contractive map $f : X \to X$ has a unique fixed point $x \in X$, and $x$ is the (unique) attractor for $f$.

Proof: Given $y, z \in X$, the sequence

$$y, z, f(y), f(z), f^2(y), f^2(z), f^3(y), f^3(z), \ldots$$

is a Cauchy sequence. By completeness of $X$, $\lim_{n \to \infty} f^n(y) = \lim_{n \to \infty} f^n(z)$, for all $y, z \in X$. Let $x$ be that common limit. By continuity of $f$,

$$f(x) = f(\lim_{n} f^n(y)) = \lim_{n} f(f^n(y)) = \lim_{n} f^n(y) = x$$

as desired. ///

2. Inverse function theorem

This is a corollary of the fixed-point lemma.

Recall that the derivative $Df(x_o)$ (evaluated) at $x_o \in \mathbb{R}^m$, of a function $f : U \to \mathbb{R}^n$ with open $U \subset \mathbb{R}^m$, if it exists, is a linear map $Df(x_o) : \mathbb{R}^m \to \mathbb{R}^n$ such that

$$f(x_o + h) = f(x_o) + Df(x_o)(h) + o(h) \quad (\text{as } |h| \to 0 \text{ with } h \in \mathbb{R}^m)$$
where \( |h| \) is the usual norm on \( \mathbb{R}^m \), and where \( o(h) \) is Landau’s little-o notation, meaning that

\[
\frac{1}{|h|} \cdot \left| f(x_o + h) - f(x_o) - Df(x_o)(h) \right| \to 0 \quad (\text{as } |h| \to 0)
\]

Differentiability, especially indefinite/infinite differentiability, can also be discussed in terms of partial derivatives.

Let \( U \subset \mathbb{R}^n \) be open, and \( f : \mathbb{R}^n \to \mathbb{R}^n \) a \( C^k \)-function.

**[2.1] Theorem:** Inverse Function Theorem: For \( x_0 \in \mathbb{R}^n \) 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 \( k \)-times differentiable inverse on \( f(V) \).

**Proof:** For linear \( T : \mathbb{R}^n \to \mathbb{R}^n \) let \( |T| \) be the uniform operator norm

\[
|T| = \sup_{|v|=1} |Tv|
\]

Without loss of generality, \( x = 0 \), \( f(x_0) = 0 \), and \( Df(0) \) is the identity map \( 1 : \mathbb{R}^n \to \mathbb{R}^n \). Letting \( F(x) := x - f(x) \), we have \( F'(0) = 0 \). By continuity, there is \( \delta_0 > 0 \) so that if \( |x| < \delta_0 \) then \( |F'(x)| < \frac{1}{2} \). Then there is \( 0 < \delta < \delta_0 \) so that \( |X| < \delta \) implies \( |F(x)| \leq |x|/2 \). Thus, \( F \) maps the closed ball \( B_\delta \) to itself, and is contractive with constant \( c = \frac{1}{2} \), from

\[
|\Phi_y(x_1) - \Phi_y(x_2)| = |F(x_1) - F(x_2)| \leq |x_1 - x_2|/2 \quad \text{(for } x_1, x_2 \in B_\delta \text{)}
\]

invoking the one-dimensional Mean-Value Theorem and using \( |F'(x)| < \frac{1}{2} \).

By the fixed-point lemma, \( \Phi_y \) has a unique fixed point, the unique solution of the equation \( f(x) = y \), and the fixed point is an attractor. (Emphatically, \( |y| < \delta/2 \)).

For continuity of the inverse map \( \varphi := f^{-1} \), take \( x_1, x_2 \in B_\delta \). From \( |F'(x_1)| < \frac{1}{2} \) we obtain the last inequality in

\[
|x_1 - x_2| \leq |f(x_1) - f(x_2)| + |F(x_1) - F(x_2)| \leq |f(x_1) - f(x_2)| + \frac{1}{2}|x_1 - x_2|
\]

Then

\[
|x_1 - x_2| < 2|f(x_1) - f(x_2)|
\]

expressing the continuity.

For differentiability, let \( y_i = f(x_i) \) with \( |y_i| < \delta/2 \). We wish to show that the derivative \( \varphi'(y_2) \) of the inverse \( \varphi \) of \( f \) at \( y_1 \) is simply

\[
\varphi'(y_2) = f'(\varphi(y_2))^{-1}
\]

To this end, estimate

\[
|\varphi(y_1) - \varphi(y_2) - f'(y_2)^{-1}(y_1 - y_2)| = |x_1 - x_2 - f'(y_2)^{-1}(f(x_1) - f(x_2))|
\]

By \( |F'(0)| = 0 \),

\[
f'(x_2) = 1_m + T
\]

where \( T \) depends upon \( x_2 \) and \( |T| = o(x_2) \) as \( x_2 \) goes to 0. By hypothesis, \( |T| < \frac{1}{2} \) for all such \( x_2 \), giving a uniform estimate

\[
|f'(x_2)^{-1}| < |1_m - T + T^2 - T^3 + \ldots| \leq \sum_{i \geq 0} 2^{-i} = 2
\]
Then
\[ x_1 - x_2 = f'(x_2)(x_1 - x_2) - T(x_1 - x_2) \]
and so
\[ |x_1 - x_2 - f'(x_2)^{-1}(f(x_1) - f(x_2))| \leq |f'(x_2)^{-1}(x_1 - x_2) - f'(x_2)^{-1}(f(x_1) + f(x_2))| + |T| \cdot |x_1 - x_2| \]
\[ \leq |f'(x_2)^{-1}| \cdot |x_1 - x_2 - f(x_1) + f(x_2)| + |T| \cdot |x_1 - x_2| \]
Since \( f \) is differentiable,
\[ |x_1 - x_2 - f(x_1) + f(x_2) = o(x_1 - x_2) \]
And \( |T| = o(x_2) \), so this is
\[ |f'(x_2)^{-1}| \cdot |F(x_1) - F(x_2)| + \frac{1}{2}|x_1 - x_2| < \frac{5}{2}|x_1 - x_2| \leq |x_1 - x_2 - f(x_1) + f(x_2)| \]
From the differentiability of \( F \), the latter expression is \( o(x_1 - x_2) \) as \( |x_1 - x_2| \) goes to zero, so is also \( o(y_1 - y_2) \) by continuity of \( \varphi \). Thus, \( \varphi \) is differentiable and its derivative is indeed \( \varphi'(y) = f'(\varphi(y))^{-1} \), for \( |y| < \delta/2 \).
Since inversion (of invertible linear maps) is an infinitely differentiable map, \( f' \) is \( C^{k-1} \), so \( \varphi' \) is also \( C^{k-1} \).

///

### 3. Implicit function theorem

The implicit function theorem can be made a corollary of the inverse function theorem. Let \( U \subset \mathbb{R}^m \) and \( V \subset \mathbb{R}^n \) be open. Let \( F : U \times V \to \mathbb{R}^n \) be a \( C^k \) mapping. Let \( F_2 \) denote the derivative of \( f \) with respect to its second argument.

**[3.1] Theorem:** Suppose that \( F_2(x_0, y_0) : \mathbb{R}^n \to \mathbb{R}^n \) is a linear isomorphism. For a sufficiently small neighborhood \( U_0 \subset U \) of \( x_0 \), there is a unique continuous map \( g : U_0 \to V \) so that \( g(x_0) = y_0 \) and \( F(x, g(x)) = 0 \) for all \( x \in U_0 \).

**Proof:** Without loss of generality, \( F_2(x_0 + 0, y_0) = 1 \). The map \( \varphi : U \times V \to \mathbb{R}^m \oplus \mathbb{R}^n \) defined by \( \varphi(x, y) = (x, F(x, y)) \) has derivative (in \( \mathbb{R}^m \oplus \mathbb{R}^n \)-coordinates)
\[
\begin{pmatrix}
1_m & 0 \\
F_1 & F_2
\end{pmatrix} = \begin{pmatrix}
1_m & 0 \\
F_1 & 1_n
\end{pmatrix}
\]
Thus, \( \varphi \) has an inverse \( \psi \) near \((x_0, y_0)\), with derivative
\[
\begin{pmatrix}
1_m & 0 \\
-F_1(x_0, y_0) & 1_n
\end{pmatrix}
\]
Then \( \psi(x, z) = (x, G(x, z)) \) where \( G \) is some \( C^k \) map. Define \( g(x) = G(x, 0) \). Then \( g \) is \( C^k \) and
\[
(x, F(x, g(x))) = \varphi(x, g(x)) = \varphi(x, G(x, 0)) = \varphi(\psi(x, 0)) = (x, 0)
\]
This proves existence.

For uniqueness: Let \( h \) be a continuous function on a neighborhood of \( x_0 \) so that \( h(x_0) = y_0 \) and \( f(x, h(x)) = 0 \) for all \( x \) near \( x_0 \). By continuity, \( h(x) \) is near \( y_0 \) for such \( x \), so \( \varphi(x, h(x)) = (x, 0) \). Since \( \varphi \) is invertible near \((x_0, y_0)\), there is a unique \((x, y)\) so that \( \varphi(x, y) = (x, 0) \). Then \( g \) and \( h \) must be equal on a neighborhood of \( x_0 \) (possibly smaller than \( U_0 \)).

The closed set of \( 0 \leq t \leq 1 \) so that \( g(x_0 + t(x - x_0)) = h(x_0 + t(x - x_0)) \) is thus non-empty. If its least upper bound is \( \tau < 1 \), then by continuity \( g(x_0 + \tau(x - x_0)) = h(x_0 + \tau(x - x_0)) \) implies the same equality for some \( t > \tau \). Thus, \( \tau = 1 \). That is, \( h = g \) on all of \( U_0 \).

///
4. Real-differentiability versus complex-differentiability

Again, for $U \subset \mathbb{R}^2$ non-empty open, for $f : U \to \mathbb{R}^2$, the derivative $Df(x_o)$ at $x_o \in U$, if it exists, is a linear map $Df(x_o) : \mathbb{R}^2 \to \mathbb{R}^2$ such that

$$f(x_o + h) = f(x_o) + Df(x_o)(h) + o(h) \quad \text{as } |h| \to 0, \ h \in \mathbb{R}^2$$

and $f$ is real-differentiable (on $U$) when $Df(x)$ exists (for all $x \in U$).

[4.1] Claim: Let $U$ be a non-empty open $U \subset \mathbb{C}$. Identifying $\mathbb{C} \approx \mathbb{R}^2$, let $f : U \to \mathbb{C}$ be real-differentiable. Then $f$ is complex-differentiable on $U$ if and only if $Df(z) : \mathbb{C} \to \mathbb{C}$ is complex-linear for every $z \in U$.

Proof: The trivial direction is that $f$ is complex-differentiable. Then $Df$ is just $f'$, with the real-linear maps $Df : \mathbb{C} \to \mathbb{C}$ being by complex multiplication.

In the slightly less trivial direction, suppose that $Df(z)$ is complex-linear for all $z \in U$, namely, $Df(z)(\alpha \cdot h) = \alpha \cdot Df(z)(h)$ for $h, \alpha \in \mathbb{C}$. For any field, the collection of $k$-linear maps of $k$ to itself is exactly the collection of multiplications by elements of $k$. Explicitly,

$$Df(z)(h) = Df(z)(h \cdot 1) = h \cdot Df(z)(1) \quad \text{(with complex multiplication)}$$

The value $Df(z)(1)$ is a complex number depending only on $z$, and fills the role of $f'(z)$, because

$$f(z + h) = f(z) + Df(z)(1) \cdot h + o(h)$$

Thus, $f$ is complex-differentiable. ///

5. Holomorphic inverse function theorem

[5.1] Theorem: For holomorphic $f$ on a neighborhood of $z_o$, if $f'(z_o) \neq 0$, then there is holomorphic $g$ on a neighborhood of $w_o = (z_o)$ giving a two-sided inverse of $f$:

$$g(f(z)) = z \quad \text{and} \quad f(g(w)) = w \quad \text{(for } z \text{ near } z_o \text{ and } w \text{ near } w_o)$$

Further, as expected, $g'(w_o) = 1/f'(z_o)$.

Proof: With $f$ considered as a real-differentiable map $f : \mathbb{R}^2 \to \mathbb{R}^2$, obtain a real-differentiable inverse $g$, and observe that complex differentiability of $f$ implies that of $g$.

First, we give a straightforward proof using the idea of the previous section, that complex-linearity of the real derivative implies complex differentiability. From $f(g(z)) = z$ with $f$ complex differentiable and $g$ real-differentiable,

$$z + h = f(g(z + h)) = f(g(z) + Dg(z)(h) + o(h)) = f(g(z)) + f'(g(z)) \cdot Dg(z)(h) + o(h)$$

so

$$h = (z + h) - z = f'(g(z)) \cdot Dg(z)(h) + o(h)$$

Replace $h$ by $t \cdot \alpha \cdot h$ with $\alpha \in \mathbb{C}$ and $t$ real:

$$t \cdot \alpha \cdot h = f'(g(z)) \cdot Dg(z)(t \cdot \alpha \cdot h) + o(t \cdot \alpha \cdot h)$$
The real-linearity of $Dg(z)$ lets us divide through by $t$:

$$\alpha \cdot h = f'(g(z)) \cdot Dg(z)(\alpha \cdot h) + o(t \cdot \alpha \cdot h) \over t$$

Taking the limit as $t \to 0$ gives

$$\alpha \cdot h = f'(g(z)) \cdot Dg(z)(\alpha \cdot h)$$

At $z$ such that $f'(g(z)) \neq 0$, this shows that

$$Dg(z)(\alpha \cdot h) = \alpha \cdot \frac{h}{f'(g(z))} = \alpha \cdot Dg(z)(h)$$

where the latter equality follows from the previous by taking $\alpha = 1$. Thus, $Dg(z)$ is complex-linear, and $g$ is complex-differentiable.

\[\Box\]

6. Holomorphic implicit function theorem

Since we do not have adequate several-complex-variables set-up, we only prove a limited version of an implicit function theorem:

Let $F(z, w)$ be a real-smooth function of two complex variables, holomorphic in each one. Let $F_z, F_w$ be the complex derivatives with respect to the first and second arguments. Assume that these are still real-smooth.

[6.1] Theorem: At $z_0, w_0$ satisfying $F(z_0, w_0) = 0$ and $F_w(z_0, w_0) \neq 0$, there is a unique holomorphic function $f(z)$ on a neighborhood $U$ of $z_0$ such that $f(z_0) = w_0$ and $F(z, f(z))$ for $z \in U$.

Proof: The implicit function theorem for smooth functions gives a real-differentiable $g : \mathbb{C} \to \mathbb{C}$ such that $F(z, g(z)) = 0$ for $z$ near $z_0$, and with $g(z_0) = w_0$. We show that $g$ is complex-differentiable.

The real-differentiability of $g$ gives

$$g(z + h) = g(z) + Dg(z)(h) + o(h) \quad \text{(for small } h \in \mathbb{C})$$

It suffices to show that $Dg$ is complex-linear. Using the assumptions on $F$,

$$0 = F(z + h, g(z + h)) = F(z + h, g(z) + h) = F(z, g(z)) + F_z(z, g(z)) \cdot h + F_w(z, g(z)) \cdot Dg(z)(h) + o(h)$$

with multiplication and addition in $\mathbb{C}$. Using $F(z, g(z)) = 0$,

$$0 = F_z(z, g(z)) \cdot h + F_w(z, g(z)) \cdot Dg(h) + o(h)$$

For fixed $\alpha \in \mathbb{C}$, and for $h$ small, replacing $h$ by $\alpha \cdot h$ gives

$$0 = F_z(z, g(z)) \cdot (\alpha \cdot h) + F_w(z, g(z)) \cdot Dg(\alpha \cdot h) + o(\alpha \cdot h)$$

Dividing through by $\alpha$ gives

$$0 = F_z(z, g(z)) \cdot h + F_w(z, g(z)) \cdot \frac{Dg(\alpha \cdot h)}{\alpha} + o(h)$$

Comparing this with the analogous earlier identity,

$$F_w(z, g(z)) \cdot \left( \frac{Dg(\alpha \cdot h)}{\alpha} - Dg(h) \right) = o(h)$$
Replacing \( h \) by \( t \cdot h \) with small real \( t \),

\[
F_w(z, g(z)) \cdot \left( \frac{Dg(\alpha \cdot t \cdot h)}{\alpha} - Dg(t \cdot h) \right) = o(t \cdot h)
\]

Using the \( \mathbb{R} \)-linearity of \( Dg \), and dividing through by \( t \), gives

\[
F_w(z, g(z)) \cdot \left( \frac{Dg(\alpha \cdot h)}{\alpha} - Dg(h) \right) = \frac{o(t \cdot h)}{t}
\]

Letting \( t \to 0 \) gives

\[
\frac{Dg(\alpha \cdot h)}{\alpha} - Dg(h) = 0
\]

This is the \( \mathbb{C} \)-linearity of \( Dg \), which is the \textit{complex}-differentiability of \( g \). //