The Implicit Function Theorem and the Inverse Function Theorem
Sample solutions

Jean-Baptiste Campesato

January 1st, 2020

Exercise 1. *Let $B = \overline{B}(a,r) \subset \mathbb{R}^n$. Let $f : B \to B$ be a contraction mapping (i.e. a Lipschitz mapping with constant $q \in [0,1)$, or equivalently $\exists q \in [0,1)$, $\forall x, y \in B$, $\|f(x) - f(y)\| \leq q\|x - y\|$).*

1. *Prove that f is continuous.*

   Let $x \in B$. Let $\varepsilon > 0$.
   Set $\delta = \begin{cases} \frac{\varepsilon}{q} & \text{if } q \neq 0 \\ 1 & \text{otherwise} \end{cases}$.
   Let $y \in B$. If $\|x - y\| < \delta$ then $\|f(x) - f(y)\| \leq q\|x - y\| < \varepsilon$.
   Hence f is continuous at x.

We define a sequence inductively by picking $x_0 \in B$ and then setting $x_{n+1} = f(x_n)$.

2. *Prove that $\forall n \in \mathbb{N}_{\geq 0}$, $\|x_{n+1} - x_n\| \leq q^n \|x_1 - x_0\|$.*

   Let's prove it by induction on n.
   Base case at $n = 0$: $\|x_1 - x_0\| \leq q^0 \|x_1 - x_0\| = \|x_1 - x_0\|$

   Induction step: assume that the statement holds for some $n \in \mathbb{N}_{\geq 0}$, then

   $$\begin{aligned} \|x_{n+2} - x_{n+1}\| &= \|f(x_{n+1}) - f(x_n)\| \\ &= q\|x_{n+1} - x_n\| \\ &= q^{n+1}\|x_1 - x_0\| \text{ by the induction hypothesis.} \end{aligned}$$

3. *Prove that $\forall m, n \in \mathbb{N}_{\geq 0}$, $m > n \implies \|x_m - x_n\| \leq \frac{q^n}{1-q}\|x_1 - x_0\|$.*

   Let $m, n \in \mathbb{N}_{\geq 0}$. Assume that $m > n$, then

   $$\begin{aligned} \|x_m - x_n\| &= \left\| \sum_{i=n}^{m-1} (x_{i+1} - x_i) \right\| \\ &\leq \sum_{i=n}^{m-1} \|x_{i+1} - x_i\| \quad \text{by the Triangle Inequality} \\ &\leq \sum_{i=n}^{m-1} q^i \|x_1 - x_0\| \quad \text{by Question 2.} \\ &= q^n \frac{1 - q^{m-n}}{1 - q} \|x_1 - x_0\| \\ &\leq \frac{q^n}{1 - q} \|x_1 - x_0\| \end{aligned}$$

4. *Prove that $(x_n)_{n\in\mathbb{N}_{\geq 0}}$ is a Cauchy sequence.*

   Let $\varepsilon > 0$.
   Case 1: if $x_1 = x_0$ then for any $m, n \in \mathbb{N}_{\geq 0}$ satisfying $m > n$, $\|x_m - x_n\| \leq \frac{q^n}{1-q}\|x_1 - x_0\| = 0 < \varepsilon$.
   Case 2: otherwise, since $\lim_{n\to+\infty} \frac{q^n}{1-q} = 0$ (as $|q| < 1$),

   $$(1) \quad \exists N \in \mathbb{N},\ n > N \implies 0 \leq \frac{q^n}{1-q} \leq \frac{\varepsilon}{\|x_1 - x_0\|}$$

   Let $m, n \in \mathbb{N}_{\geq 0}$, if $m \geq n > N$ then

   $$\begin{aligned}\|x_m - x_n\| &\leq \frac{q^n}{1-q}\|x_1 - x_0\| \text{ by Question 3}\\ &\leq \varepsilon \text{ by (1)}\end{aligned}$$

   We proved that $\forall \varepsilon > 0, \exists N \in \mathbb{N}_{\geq 0}, \forall m, n \in \mathbb{N}_{\geq 0},\ m \geq n > N \implies \|x_m - x_n\| \leq \varepsilon$
   i.e. $(x_n)_{n\in\mathbb{N}_{\geq 0}}$ is a Cauchy sequence.

5. *Prove that $(x_n)_{n\in\mathbb{N}_{\geq 0}}$ is convergent in $\boldsymbol{B}$.*

   Since $(x_n)_{n\in\mathbb{N}_{\geq 0}}$ is a Cauchy sequence in $\mathbb{R}^n$, it admits a limit $x \in \mathbb{R}^n$.
   And since $\forall n \in \mathbb{N},\ x_n \in \boldsymbol{B}$ and $\boldsymbol{B} \subset \mathbb{R}^n$ is a closed subset, its limit x must lie in $\boldsymbol{B}$.

6. *Prove that f admits a fixed point, i.e. $\exists x \in \boldsymbol{B},\ f(x) = x$.*

   By the previous question, there exists $x \in \boldsymbol{B}$ such that $\lim_{n\to+\infty} x_n = x$.
   Hence, by continuity of f (cf Question 1.), $\lim_{n\to+\infty} f(x_n) = f(x)$.
   Since $\forall n \in \mathbb{N}_{\geq 0},\ f(x_n) = x_{n+1}$, by uniqueness of the limit, we get that $f(x) = x$.

7. *Prove that f admits only one fixed point (i.e. if $y \in \boldsymbol{B}$ is another fixed of f point then $x = y$).*

   Assume by contradiction that there exists $y \in \boldsymbol{B}$ another fixed point, i.e. $x \neq y$ and $f(y) = y$.
   Then $\|x - y\| = \|f(x) - f(y)\| \leq q\|x - y\|$.
   Since $\|x - y\| > 0$, we get that $q \geq 1$, which contradicts the assumption $q \in [0, 1)$.

You just proved the

Theorem 2 (Banach fixed point theorem)**.** *A contraction mapping $f : \overline{B}(a, r) \to \overline{B}(a, r)$ admits a unique fixed point.*

Exercise 3. *Let $U \subset \mathbb{R}^n$ be an open subset containing $\mathbf{0}$ and $f : U \to \mathbb{R}^n$ be a C^1 function satisfying $f(\mathbf{0}) = \mathbf{0}$ and $Df(\mathbf{0}) = I_{n,n}$.*

1. *Prove that there exists $t > 0$ such that $B(\mathbf{0}, t) \subset U$ and $\forall x \in B(\mathbf{0}, t)$, $\det(Df(x)) \neq 0$.*

   Since f is C^1, $x \mapsto Df(x)$ is continuous (the entries of $Df(x)$ are the partial derivatives of the components of f which are continuous by definition of C^1).
   Then $\varphi : U \to \mathbb{R}$ defined by $\varphi(x) = \det(Df(x))$ is continuous too by composition of continuous functions.
   Since φ is continuous at $\mathbf{0}$, there exists $t_1 > 0$ such that
   $$\|x - \mathbf{0}\| < t_1 \implies |\varphi(x) - \varphi(\mathbf{0})| < \frac{1}{2}$$
   i.e.
   $$\|x\| < t_1 \implies |\varphi(x) - 1| < \frac{1}{2}$$
   Since U is open, there exists $t_2 > 0$ such that $B(\mathbf{0}, t_2) \subset U$.
   Set $t = \min(t_1, t_2)$.

   Then $B(\mathbf{0}, t) \subset B(\mathbf{0}, t_2) \subset U$.

   Now let $x \in B(\mathbf{0}, t)$, then, since $\|x\| < t \leq t_1$,
   $$1 = |1 - \varphi(x) + \varphi(x)| \leq |1 - \varphi(x)| + |\varphi(x)| < \frac{1}{2} + |\varphi(x)|$$
   Hence $|\varphi(x)| > \frac{1}{2}$ and therefore $\varphi(x) \neq 0$.
   We just proved that $\forall x \in B(\mathbf{0}, t)$, $\det(Df(x)) \neq 0$.

2. *Define $F : U \to \mathbb{R}^n$ by $F(x) = f(x) - x$.*

   (a) *Compute $DF(\mathbf{0})$.*

   $DF(\mathbf{0}) = Df(\mathbf{0}) - D\mathrm{id}(\mathbf{0}) = I_{n,n} - I_{n,n} = 0 \in M_{n,n}(\mathbb{R})$

   (b) *Prove that there exists $r \in (0, t)$ such that $\forall x \in B(\mathbf{0}, r)$, $\|DF(x)\| \leq \frac{1}{2}$.*

   Since F is C^1, DF is continuous, hence there exists $r_1 > 0$ such that
   $$\|x - \mathbf{0}\| < r_1 \implies \|DF(x) - DF(\mathbf{0})\| < \frac{1}{2}$$
   i.e.
   $$\|x\| < r_1 \implies \|DF(x)\| < \frac{1}{2}$$
   Therefore we can take $r = \min\left(r_1, \frac{t}{2}\right)$.

   (c) *Prove that there exists $s \in (0, r)$ such that*
   $$\forall x, y \in \overline{B}(\mathbf{0}, s),\ \|F(x) - F(y)\| \leq \left(\sup_{z \in \overline{B}(\mathbf{0}, s)} \|DF(z)\|\right) \|x - y\|$$
   Comment: we use the Frobenius norm for matrices as in PS4 (recall that $\|AB\| \leq \|A\|\|B\|$).

   Take $s = \frac{r}{2}$ and then follow one of the proofs from the document "an MVT-like inequality".

   (d) *Prove that $\forall x, y \in \overline{B}(\mathbf{0}, s)$, $\|F(x) - F(y)\| \leq \frac{1}{2}\|x - y\|$.*

   Let $x, y \in \overline{B}(\mathbf{0}, s)$, then
   $$\begin{aligned}\|F(x) - F(y)\| &\leq \left(\sup_{z \in \overline{B}(\mathbf{0}, s)} \|DF(z)\|\right) \|x - y\| \text{ by Question 2.(c)} \\ &\leq \frac{1}{2}\|x - y\| \text{ by Question 2.(b) since } \overline{B}(\mathbf{0}, s) \subset B(\mathbf{0}, r)\end{aligned}$$

3. *Let* $y \in B\left(\mathbf{0}, \frac{s}{2}\right)$. *Define* $\theta_y : U \to \mathbb{R}^n$ *by* $\theta_y(x) = y - F(x)$.

   (*a*) *Prove that* $\theta_y\left(\overline{B}(\mathbf{0}, s)\right) \subset B(\mathbf{0}, s)$.

   Let $x \in \overline{B}(\mathbf{0}, s)$ *then*

$$\begin{aligned}\|\theta_y(x)\| &= \|y - F(x)\| \\ &\leq \|y\| + \|F(x)\| \\ &< \frac{s}{2} + \|F(x) - F(\mathbf{0})\| \quad \text{since } \|y\| < \frac{s}{2} \text{ and } F(\mathbf{0}) = \mathbf{0} \\ &\leq \frac{s}{2} + \frac{1}{2}\|x - \mathbf{0}\| \quad \text{by Question 2.}(d)\text{, since } x \in \overline{B}(\mathbf{0}, s) \\ &\leq s \text{ since } x \in \overline{B}(\mathbf{0}, s)\end{aligned}$$

   Hence $\forall x \in \overline{B}(\mathbf{0}, s)$, $\|\theta_y(x)\| < s$.

   (*b*) *Prove that* $\theta_y : \overline{B}(\mathbf{0}, s) \to \overline{B}(\mathbf{0}, s)$ *is a contraction mapping with constant* $\frac{1}{2}$.

   We first notice that since $\theta_y\left(\overline{B}(\mathbf{0}, s)\right) \subset B(\mathbf{0}, s) \subset \overline{B}(\mathbf{0}, s)$, *the function* $\theta_y : \overline{B}(\mathbf{0}, s) \to \overline{B}(\mathbf{0}, s)$ *is well-defined.*

   Let $x, x' \in \overline{B}(\mathbf{0}, s)$, *then*

$$\begin{aligned}\|\theta_y(x) - \theta_y(x')\| &= \|y - F(x) - y + F(x')\| \\ &= \|F(x) - F(x')\| \\ &\leq \frac{1}{2}\|x - x'\| \text{ by Question 2.}(d)\end{aligned}$$

   Therefore θ_y *is a contraction mapping with constant* $\frac{1}{2}$.

   (*c*) *We set* $V = B(\mathbf{0}, s) \cap f^{-1}\left(B\left(\mathbf{0}, \frac{s}{2}\right)\right)$ *and* $W = B\left(\mathbf{0}, \frac{s}{2}\right)$.
   Prove that $f : V \to W$ *is a well-defined bijection between two open subsets of* $\mathbb{R}^n$ *containing* $\mathbf{0}$.

   We first notice that V *and* W *are obviously open subsets.*

   Then we check that $f : V \to W$ *is well-defined:*

   - *Since* $B(\mathbf{0}, s) \subset B(\mathbf{0}, t) \subset U$, *we have well that* $V \subset U$. *Hence* f *is well-defined on* V.
   - *Moreover* $f(V) \subset W$ *by definition of* V.

   Finally, we check that $f : V \to W$ *is a bijection:*
   Let $y \in W$.
   We know from Question 3.(b) that $\theta_y : \overline{B}(\mathbf{0}, s) \to \overline{B}(\mathbf{0}, s)$ *is a contraction mapping, hence, by the Banach fixed point theorem, there exists a unique* $x \in \overline{B}(\mathbf{0}, s)$ *such that* $x = \theta_y(x)$.
   Notice that $x = \theta_y(x) \Leftrightarrow x = y - f(x) + x \Leftrightarrow y = f(x)$.
   Hence $x \in f^{-1}\left(B\left(\mathbf{0}, \frac{s}{2}\right)\right)$ *since* $y \in B\left(\mathbf{0}, \frac{s}{2}\right)$.
   Moreover, by Question 3.(a), $x = \theta_y(x) \in B(\mathbf{0}, s)$. *So* $x \in V$.
   To summarize, we just proved that

$$\forall y \in W, \exists! x \in V, y = f(x)$$

   i.e. $f : V \to W$ *is a bijection.*

4. (a) *Prove that $f^{-1} : W \to V$ is Lipschitz with constant 2 and then that it is continuous.*

Let $y_1, y_2 \in W$.
Set $x_1 = f^{-1}(y_1)$ and $x_2 = f^{-1}(y_2)$ which are well-defined by Question 3.(c).
Notice that

$$\begin{aligned}\|x_1 - x_2\| &= \left\|\left(f(x_1) - f(x_2)\right) - \left(F(x_1) - F(x_2)\right)\right\| \\ &\leq \left\|f(x_1) - f(x_2)\right\| + \left\|F(x_1) - F(x_2)\right\| \quad \text{by the Triangle Inequality}\end{aligned}$$

So

$$\begin{aligned}\|x_1 - x_2\| - \|f(x_1) - f(x_2)\| &\leq \|F(x_1) - F(x_2)\| \\ &\leq \frac{1}{2}\|x_1 - x_2\| \quad \text{by Question 2.(d)}\end{aligned}$$

Hence

$$\|x_1 - x_2\| \leq 2\|f(x_1) - f(x_2)\|$$

i.e.

$$\left\|f^{-1}(y_1) - f^{-1}(y_2)\right\| \leq 2\|y_1 - y_2\|$$

We just proved that f^{-1} is Lipschitz with constant 2.
Then f^{-1} is continuous as a Lipschitz function (you can adapt Question 1. of Exercise 1.).

(b) *Prove that $f^{-1} : W \to V$ is differentiable.*
(Hint: use the very definition of differentiability together with Question 1.)

Let $y_0 \in W$ and set $x_0 = f^{-1}(y_0)$.
Since f is differentiable at x_0, we have

$$f(x) = f(x_0 + x - x_0) = f(x_0) + d_{x_0}f(x - x_0) + E(x) \tag{2}$$

where

$$\lim_{x \to x_0} \frac{E(x)}{\|x - x_0\|} = \mathbf{0} \tag{3}$$

Since $V \subset B(\mathbf{0}, t)$, $d_{x_0}f$ is invertible by Question 1.
Hence we may compose (2) with $\left(d_{x_0}f\right)^{-1}$ in order to obtain that

$$\left(d_{x_0}f\right)^{-1}(f(x) - f(x_0)) = x - x_0 + \left(d_{x_0}f\right)^{-1}(E(x))$$

which we may rewrite in terms of $y = f(x)$:

$$\left(d_{x_0}f\right)^{-1}(y - y_0) = f^{-1}(y) - f^{-1}(y_0) + \left(d_{x_0}f\right)^{-1}\left(E\left(f^{-1}(y)\right)\right)$$

Hence

$$f^{-1}(y) = f^{-1}(y_0) + \left(d_{x_0}f\right)^{-1}(y - y_0) - \left(d_{x_0}f\right)^{-1}\left(E\left(f^{-1}(y)\right)\right)$$

We already know that $\left(d_{x_0}f\right)^{-1}$ is linear, so, to prove that f^{-1} is differentiable at y_0, it is enough to check that

$$\lim_{y \to y_0} \frac{\left(d_{x_0}f\right)^{-1}\left(E\left(f^{-1}(y)\right)\right)}{\|y - y_0\|} = \mathbf{0}$$

We first notice that $\|x - x_0\| = \|f^{-1}(y) - f^{-1}(y_0)\| \leq 2\|y - y_0\|$ by Question 4.(a) from which we derive that

$$\frac{1}{\|y - y_0\|} \leq \frac{2}{\|x - x_0\|}$$

and next that

$$\left\| \frac{\left(d_{x_0} f\right)^{-1} \left(E\left(f^{-1}(y)\right)\right)}{\|y - y_0\|} \right\| \leq 2 \left\| \left(d_{x_0} f\right)^{-1} \left(\frac{E(x)}{\|x - x_0\|}\right) \right\|$$

When $y \to y_0$ we have $x \to x_0$ by continuity of f (Question 4.(a)) and hence that $\frac{E(x)}{\|x-x_0\|} \to \mathbf{0}$ by (3).
We know that $\left(d_{x_0} f\right)^{-1}$ is continuous (since linear) and that $\|\cdot\|$ is continuous too, hence

$$\left\| \frac{\left(d_{x_0} f\right)^{-1} \left(E\left(f^{-1}(y)\right)\right)}{\|y - y_0\|} \right\| \leq 2 \left\| \left(d_{x_0} f\right)^{-1} \left(\frac{E(x)}{\|x - x_0\|}\right) \right\| \xrightarrow[y \to y_0]{} 0$$

Therefore

$$\lim_{y \to y_0} \frac{\left(d_{x_0} f\right)^{-1} \left(E\left(f^{-1}(y)\right)\right)}{\|y - y_0\|} = \mathbf{0}$$

and f^{-1} is differentiable at y_0.

(c) *Prove that $f^{-1} : W \to V$ is C^1.*
(Hint: study $D\left(f^{-1}\right)$ using the Chain Rule.)

Since f^{-1} is differentiable, we may apply the chain rule to the LHS of the identity $f \circ f^{-1} = \mathrm{id}$ in order to obtain

$$Df\left(f^{-1}(y)\right) D\left(f^{-1}\right)(y) = I_{n,n}$$

Since $Df\left(f^{-1}(y)\right)$ is invertible by Question 1, we deduce that

$$D\left(f^{-1}\right)(y) = \left(Df\left(f^{-1}(y)\right)\right)^{-1}$$

The entries of the RHS matrix are continuous by the formula giving the matrix inverse in terms of the cofactor matrix, hence the entries of the LHS matrix are continuous too.
But the entries of the LHS are the partial derivatives of the components of f^{-1}, so they are continuous.
Therefore f^{-1} is C^1.

You just proved that

Claim 4. *Let $U \subset \mathbb{R}^n$ be an open subset containing $\mathbf{0}$ and $f : U \to \mathbb{R}^n$ be of class C^1 such that $f(\mathbf{0}) = \mathbf{0}$ and $Df(\mathbf{0}) = I_{n,n}$. Then there exist $V, W \subset \mathbb{R}^n$ two open subsets such that $\mathbf{0} \in V \subset U$, $\mathbf{0} \in W$ and $f : V \to W$ is a C^1-diffeomorphism (i.e. f is C^1, bijective and f^{-1} is C^1).*

Exercise 5. *Prove the Inverse Function Theorem (the statement is below).*
(Hint: reduce to the case considered in the above claim.)

Theorem 6 (The Inverse Function Theorem)**.** *Let $U \subset \mathbb{R}^n$ open, $f : U \to \mathbb{R}^n$ of class C^1 and $a \in U$. Assume that $Df(a)$ is invertible.*
Then there exist $V, W \subset \mathbb{R}^n$ two open subsets such that $a \in V \subset U$, $f(a) \in W$ and $f : V \to W$ is a C^1-diffeomorphism (i.e. f is C^1, bijective and f^{-1} is C^1).

Notice that $\tilde{f}(x) = (d_a f)^{-1}(f(a+x) - f(a))$ is well-defined in a neighborhood of the origin.
Moreover $\tilde{f}(\mathbf{0}) = \mathbf{0}$ and $D\tilde{f}(\mathbf{0}) = I_{n,n}$ by the chain rule.
So we may apply the previous claim to $\tilde{f}$ in order to deduce the statement for f.

Exercise 7. *Derive the Implicit Function Theorem from the Inverse Function Theorem (the statement is below).*

Theorem 8 (The Implicit Function Theorem).
Let $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^p$ be two open subsets. Let $(x_0, y_0) \in U \times V$.
Let $F : \begin{array}{ccc} U \times V & \to & \mathbb{R}^p \\ (x, y) & \mapsto & F(x, y) \end{array}$ *be of class C^1.*
If $D_y F(x_0, y_0)$ is invertible then there exist $r, s > 0$ such that $B(x_0, r) \subset U$, $B(y_0, s) \subset V$ and there exists $\varphi : B(x_0, r) \to B(y_0, s)$ of class C^1 such that

$$\forall (x, y) \in B(x_0, r) \times B(y_0, s),\ F(x, y) = F(x_0, y_0) \Leftrightarrow y = \varphi(x)$$

Hint: apply the Inverse Function Theorem to $\psi(x, y) = (x, F(y))$.