Proof of the General Case

- Based on Volker's notes §5.2, §5.3.

The plan.

We will prove the following results.

Theorem 1. (Inverse function theorem) *Let $\varnothing \neq U \subset \mathbb{R}^N$ be open, let $f \in C^1(U, \mathbb{R}^N)$, and let $x_0 \in U$ be such that* $\det J_f(x_0) \neq 0$*. Then there is an open neighborhood $V \subseteq U$ of x_0 such that f is injective on V, $f(V)$ is open, and $f^{-1}\colon f(V) \mapsto \mathbb{R}^N$ is a C^1 function such that $J_{f^{-1}} = J_f^{-1}$.*

Theorem 2. (Implicit function theorem) *Let $\varnothing \neq U \subseteq \mathbb{R}^{M+N}$ be open, let $f \in C^1(U, \mathbb{R}^N)$, and let $(x_0, y_0) \in U$ be such that $f(x_0, y_0) = 0$ and $\det\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$. Then there are neighborhoods $V \subseteq \mathbb{R}^M$ of x_0 and $W \subseteq \mathbb{R}^N$ of y_0 with $V \times W \subseteq U$ and a unique $\varnothing \in C^1(V, \mathbb{R}^N)$ such that*

i. $\phi(x_0) = y_0$, and

ii. $f(x, y) = 0$ if and only if $\phi(x) = y$ for all $(x, y) \in V \times W$.

Moreover, we have

$$J_\phi = -\left(\frac{\partial f}{\partial y}\right)^{-1} \frac{\partial f}{\partial x}. \tag{1}$$

We will prove these results through the following steps.

1. We prove Theorem 1 first, in three steps.
   Let $\varnothing \neq U \subset \mathbb{R}^N$ be open, let $f \in C^1(U, \mathbb{R}^N)$, and let $x_0 \in U$ be such that $\det J_f(x_0) \neq 0$.
   i. There is $V \subseteq U$ such that f is injective on V.
   ii. $W := f(V)$ is open.
   iii. $f^{-1}\colon W \mapsto V$ is a C^1 function.
2. Next we prove Theorem 2 as a corollary of Theorem 1.

Remark 3. The necessity of the conditions $\det J_f(x_0) \neq 0$ and $\det \frac{\partial f}{\partial y}(x_0, y_0) \neq 0$ can be easily understood through the special case where f is a linear function. On the other hand, the assumption that $f \in C^1$ is more of a technical nature.

Proof of the inverse function theorem

In this section we prove Theorem 1. Let $\varnothing \neq U \subset \mathbb{R}^N$ be open, let $f \in C^1(U, \mathbb{R}^N)$, and let $x_0 \in U$ be such that $\det J_f(x_0) \neq 0$.

There is $V \subseteq U$ such that f is injective on V.

As $f \in C^1$, for any $1 \leqslant i, j \leqslant N$, $\frac{\partial f_i}{\partial x_j}$ is a continuous function on U. Thus there is $r > 0$ such that for any N^2 points $x_{11}, x_{12}, \ldots, x_{NN} \in B_r(x_0)$, the matrix $\left(\frac{\partial f_i}{\partial x_j}(x_{ij}) \right)$ is nonsingular.

Take $V := B_r(x_0)$. Let $x, y \in V$ be arbitrary. We apply the mean value theorem for the functions

$$g_i(t) := f_i(x + t\,(y - x)) \tag{2}$$

and obtain

$$f_i(y) - f_i(x) = g_i(1) - g_i(0) = [\mathrm{grad}\, f(\xi_i)] \cdot (y - x) \tag{3}$$

where $\xi_i = x + t_i\,(y - x)$ for some $t_i \in [0, 1]$. Thus we see that

$$f(y) - f(x) = \tilde{J} \cdot (y - x) \tag{4}$$

where $\tilde{J}$ is an $N\times N$ matrix whose (i,j) entry is $\left(\frac{\partial f_i}{\partial x_j}(\xi_i)\right)$.

As $\xi_i\in V$ for $i=1,2,...,N$, the matrix $\tilde{J}$ is non-singular. Consequently $y\neq x\Longrightarrow f(y)\neq f(x)$. We have proved injectivity of f on V.

Remark 4. Note that for $f\colon\mathbb{R}^M\mapsto\mathbb{R}^N$, there does not hold a "mean value theorem" of the form

$$f(y)-f(x)=J_f(\xi)\cdot(y-x) \tag{5}$$

for some ξ "between" x and y, unless $N=1$.

Remark 5. Considering $f(x)=\frac{x}{2}+x^2\sin\frac{1}{x}$ we see that it is not sufficient to assume only differentiability of f.

$W:=f(V)$ is open.

Let $x_0\in V$ and $y_0=f(x_0)$. We try to show that there is $\varepsilon>0$ such that $B_\varepsilon(y_0)\subseteq W$. Let $r>0$ be such that $B_r[x_0]:=\overline{B_r(x_0)}\subset V$. As f is injective, there holds

$$y_0\notin f(\partial B_r(x_0)). \tag{6}$$

Furthermore as f is continuous, $f(\partial B_r(x_0))$ is compact. Consequently

$$\varepsilon:=\frac{1}{3}\inf_{z\in f(\partial B_r(x_0))}\|y_0-z\|>0. \tag{7}$$

We claim that $B_\varepsilon(y_0)\subseteq W$.

To see this, take any $y\in B_\varepsilon(y_0)$ and consider the optimization problem

$$\min_{x\in B_r[x_0]} g(x):=\|f(x)-y\|^2. \tag{8}$$

As g is continuous, the minimum is attained, either inside $B_r(x_0)$, or on $\partial B_r(x_0)$.

We see that $g(x_0)\leqslant\varepsilon$ while on the other hand,

$$\forall x\in\partial B_r(x_0),\qquad g(x)=\|f(x)-y\|\geqslant\|f(x)-y_0\|-\|y-y_0\|\geqslant 2\,\varepsilon. \tag{9}$$

Therefore the minimum is not attained on $\partial B_r(x_0)$.

Let $\tilde{x}$ be a minimizer. We see that $\tilde{x}\in B_r(x_0)$, consequently

$$0=\nabla g(\tilde{x})=2\,J_f(x)\cdot(f(\tilde{x})-y). \tag{10}$$

We recall that V is chosen such that for any N^2 points $x_{11}, x_{12}, ..., x_{NN}\in V$, the matrix $\left(\frac{\partial f_i}{\partial x_j}(x_{ij})\right)$ is non-singular. In particular $J_f(x)$ is non-singular. Consequently $f(\tilde{x})=y$. As y is arbitrary, we see that $B_\varepsilon(y_0)\subseteq W$.