# Inverse Mapping Theorem

S. Kumaresan
School of Math. and Stat.
University of Hyderabad
Hyderabad 500046
kumaresa@gmail.com

**Theorem 1** (Inverse Mapping Theorem)**.** *Let $f\colon U \subset \mathbf{E} \to \mathbf{E}$ be $C^1$. Assume that $x_0 \in U$ is such that $Df(x_0)$ is invertible. Then there is an open subset $V$ of $U$ containing the point $x_0$ with the following properties:*

*i) on $V$ the map is one-one*

*ii) the image $f(V)$ is open neighborhood of $f(x_0)$*

*iii) $f^{-1}$ is $C^1$ on $f(V)$ with $Df^{-1}(y) = Df(f^{-1}(y))^{-1}$ for all $y \in V$.*

Before proving this, we recall the mean value inequality in the following form:

**Theorem 2** (Mean Value Inequality-2)**.** *Let $f\colon U \subset \mathbf{E} \to \mathbf{F}$ be differentiable on the open set $U$. Let $x, y \in U$ such that $[x, y] \subset U$. Let $T \in BL(\mathbf{E}, \mathbf{F})$. Then we have*

$$\| f(y) - f(x) - T(y-x) \| \leq \| y - x \| \sup_{0<t<1} \| Df(x + t(y-x)) - T \| . \tag{1}$$

*Proof.* Consider $\varphi(x) = f(x) - Tx$ and apply the standard mean value inequality to $\varphi$. □

*Proof.* (of IMT) We may assume that $Df(x_0) = I_{\mathbf{E}}$, the identity map of $\mathbf{E}$. (Justify.)

Since $Df$ is continuous at $x_0$ there exists a $\delta > 0$ such that

$$\| Df(x) - Df(x_0) \| < 1/2, \qquad \text{for all } x \in B(x_0, \delta). \tag{2}$$

For $x_1, x_2 \in B(x_0, \delta)$ we have by the mean value inequality

$$\| f(x_1) - f(x_2) - Df(x_0)(x_1 - x_2) \| \leq \| x_1 - x_2 \| \sup_{0<t<1} \| Df(x_1 + t(x_2 - x_1)) - Df(x_0) \| . \tag{3}$$

Now $x_1 + t(x_2 - x_1) - x_0 = (1-t)(x_1 - x_0) + t(x_2 - x_0)$ so that $\| x_1 + t(x_2 - x_1) - x_0 \| \leq (1-t) \| x_1 - x_0 \| + t \| x_2 - x_0 \| < \delta$. Hence, in view of Eq. 2, Eq. 3 becomes

$$\| f(x_1) - f(x_2) - (x_1 - x_2) \| \leq \frac{1}{2} \| x_1 - x_2 \| . \tag{4}$$

This implies $f$ is 1-1 on $B(x_0, \delta)$.

To show that there exists a ball $B(y_0, \delta')$ contained in the image of $B(x_0, \delta)$, we modify the Newton's method. Newton's method can be briefly described as follows: if $x_0$ is an approximate zero of $f(x) = 0$ and if $f$ has a non vanishing derivative around this point $x_0$ then with $x_{n+1} := x_n - \frac{f(x_n)}{Df(x_n)}$, defined recursively we have $x_n$ tending to a limit $x$ which is a zero of $f$. Draw some pictures to see the geometric idea behind this algorithm. We modify this algorithm below.

Suppose $y \in B(y_0, \delta')$ for some $\delta' > 0$. We want to solve for the equation $f(x) = y$ with $x \in B(x_0, \delta)$. We define recursively $x_k = x_{k-1} + y - f(x_{k-1})$. We need to check whether $x_k \in B(x_0, \delta)$. We have

$$x_k - x_{k-1} = x_{k-1} - x_{k-2} - (f(x_{k-1}) - f(x_{k-2})).$$

Taking norm on both sides, using Eq. 4 and induction we get

$$\|x_k - x_{k-1}\| \leq (1/2)\|x_{k-1} - x_{k-2}\| \leq \cdots \leq \frac{1}{2^{k-1}}\|x_1 - x_0\| \leq 2^{1-k}\delta'.$$

Hence if we choose $\delta' = (\delta/2)$, then we have $\|x_k - x_{k-1}\| \leq 2^{-k}\delta$. In particular, $\|x_0 - x_k\| \leq \sum_{i=0}^{k} \delta 2^{-i} < \delta$. This also shows that $x_k$ is Cauchy. Let $\lim x_k = x$. Clearly we have $f(x) = y$. For, from the recursive definition of $x_k$ by taking the limit as $k \to \infty$ we get $x = \lim x_k = \lim_k (x_{k-1} + y - f(x_{k-1})) = x + y - f(x)$.

For $y \in B(y_0, \delta/2)$ we define $g(y) := x$ where $x \in B(x_0, \delta)$. We want to prove that $g$ is differentiable on $B(y_0, \delta/2)$ and that $Dg(y) = Df(x)^{-1}$. To see this, we let $g(y) = x$, $g(y + k) = x + h$ so that $f(x + h) = y + k$. Then we have

$$k = f(x + h) - f(x) = Df(x)h + \psi. \tag{5}$$

where $\lim_{h\to 0} \frac{\psi}{\|h\|} = 0$. From Eq. 4 it follows that

$$\|k - h\| = \|f(x + h) - f(x) - h\| \leq (1/2)\|h\|,$$

so that $(1/2)\|h\| \leq \|k\| \leq (3/2)\|h\|$. In view of Eq. 5 we have $h - Df(x)^{-1}\psi$. But $\lim_{k\to 0} \frac{Df(x)^{-1}\psi}{\|k\|} = 0$ since $2\|h\| \geq \|k\| \geq 2/3\|h\|$ and $\left\|Df(x)^{-1}\right\| \leq 2$. Then

$$\frac{2}{3}\frac{Df(x)^{-1}\psi}{\|h\|} \leq \frac{Df(x)^{-1}\psi}{\|k\|} \leq 2\frac{Df(x)^{-1}\psi}{\|h\|}$$

This proves the result. □

We shall explain the geometric meaning of IMT in the finite dimensional case.

Let $f\colon \mathbb{R}^{n+1} \to \mathbb{R}$ be $C^1$. Assume that 0 lies in the image of $f$, without loss of generality. In general, the set $S := f^{-1}(0)$ does not have any nice geometric property. However, if we assume that

$$f'(p) := \text{grad}\, f(p) \neq 0 \text{for all } p \in S$$

then $S$ "looks locally like a hyperplane". Of course this needs explanation! Given $p \in S$, since $f'(p) \neq 0$, we assume without loss of generality that $\frac{\partial f}{\partial x_{n+1}} \neq 0$. Then consider the map $\Phi\colon \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ given by $(x_1, \ldots, x_n, x_{n+1}) \mapsto (x_1, \ldots, x_n, f(x))$. Then $\Phi'(p)$ has the Jacobian

$$\begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \ldots & \frac{\partial f}{\partial x_{n+1}} \end{pmatrix}.$$

The determinant of this matrix is $\pm\frac{\partial f}{\partial x_{n+1}} \neq 0$ and hence $\Phi'(p)$ is invertible. Thus $\Phi$ is a "$C^1$-diffeomorphism" of an open neighborhood of $V$ of $p$ in $\mathbb{R}^{n+1}$ onto an open set $\Phi(V) \subset \mathbb{R}^{n+1}$, by IMT. (By a $C^1$-diffeomorphism we mean a map $F$ which is $C^1$, one-one on its domain $U$ and $F(U)$ is open and $F^{-1}$ is also $C^1$ on $F(U)$.) Now we introduce a new set of coordinates on $V$ by setting $y_i(r) := u_i \circ \Phi(r)$, where $u_i$ are the 'usual' coordinates on $\mathbb{R}^{n+1}$: $u_i(x) := x_i$. In plain language this is:

$$y_i(r) = \begin{cases} r_i & \text{if } 1 \leq i \leq n \\ f(r) & \text{if } i = n+1. \end{cases}$$

With respect to this new set of coordinates $y_i$, $S$ has a local description around $p$ on $V \cap S$: it is the "hyperplane" $\{y_{n+1} = 0\}$. Thus by taking a suitable system of coordinates we "straighten" the hypersurface to a hyperplane locally.

To see how this change of coordinates can help us we pose the following question: Suppose the hypersurface $S$ is also described around $p$ as $g^{-1}(0)$. That is, there exists an open set $U \ni p$ such that $S \cap U = g^{-1}(0)$, with $g\colon U \to \mathbb{R}$ being a $C^1$-function. Is $g$ divisible by $f$ at least locally around $p$? That is, does there exist another function $h$ defined in an open set containing $p$ on which we can write $g = fh$?

Let $F := f \circ \Phi^{-1}$ and $G := g \circ \Phi^{-1}$. Then it follows that $\Phi(V \cap S) = \{y_{n+1} = 0\} \cap \Phi(V)$. The above question then reduces to an equivalent one: Is $G$ divisible by $F = y_{n+1}$ locally around 0? This is certainly easy to answer (in the affirmative, by Taylor expansion).

**The moral therefore is that the IMT allows us to use a coordinate system that is most convenient or that simplifies the geometric problem on hand.**

Before closing this section we shall state the implicit function theorem and leave its proof to the reader. First recall that if $f\colon U \subset \mathbf{E} \times \mathbf{F} \to \mathbf{G}$ is differentiable at $(x, y)$ then we have

$$f'(x, y)(u, v) = D_x f(x, y)(u, 0) + D_y f(x, y)(0, v)$$

where $D_x f(x, y)$ is the *partial derivative* of $f$ in the first variable etc.

**Theorem 3.** *Assume that $f\colon U \subset \mathbf{E} \times \mathbf{F} \to \mathbf{G}$ is such that*
*i) $f(x_0, y_o) = 0$ for some $(x_0, y_0) \in U$*
*ii) $D_y f$ exists on $U$ and $Df(x_0, y_0) : \mathbf{F} \to \mathbf{G}$ is continuous and bijective*
*iii) $f$ and $D_y f$ are continuous at $(x_0, y_0)$.*

*Then there exist $\rho$ and $r$ such that for all $x \in B(x_0, \rho)$ there is a unique $y(x) \in \mathbf{F}$ for which $\|y(x) - y_o\| \leq r$ and $f(x, y(x)) = 0$. Furthermore, if $f$ is $C^1$ for $0 \leq k \leq \infty$ around $(x_0, y_0)$, then so is $x \mapsto y(x)$ around $x_0$.* □