General implicit and inverse function theorems

Theorem 1. (Implicit function theorem) *Let* $\boldsymbol{f}\colon \mathbb{R}^N \mapsto \mathbb{R}^M$ *with* $N > M$*. We decompose*

$$\mathbb{R}^N = \mathbb{R}^{N-M} \times \mathbb{R}^M \tag{1}$$

and denote the first $N - M$ *coordinates by vector* $\boldsymbol{x}$ *and the rest* M *coordinates by* $\boldsymbol{y}$*.*

Assume

i. $\boldsymbol{f}$ *is differentiable and has continuous partial derivatives;*

ii. $\boldsymbol{f}(\boldsymbol{x}_0, \boldsymbol{y}_0) = \boldsymbol{0}$.

iii. $\det\left(\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{y}}\right)(\boldsymbol{x}_0, \boldsymbol{y}_0) \neq 0$ *where the Jacobian matrix with respect to* $\boldsymbol{y}$ *is defined as*

$$\left(\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{y}}\right) := \begin{pmatrix} \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_M} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_M}{\partial y_1} & \cdots & \frac{\partial f_M}{\partial y_M} \end{pmatrix}. \tag{2}$$

Then there are open sets $U \subseteq \mathbb{R}^{N-M}, V \subseteq \mathbb{R}^M$ *satisfying* $\boldsymbol{x}_0 \in U$, $\boldsymbol{y}_0 \in V$ *and*

i. For every $\boldsymbol{x} \in U$ *the equation* $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{0}$ *has one unique solution* $\boldsymbol{y} = \boldsymbol{Y}(\boldsymbol{x}) \in V$*;*

ii. $\boldsymbol{Y}(\boldsymbol{x}_0) = \boldsymbol{y}_0$*;*

iii. $\boldsymbol{Y}$ *is differentiable with continuous partial derivatives;*

iv. For $\boldsymbol{x} \in U$*,*

$$\left(\frac{\partial \boldsymbol{Y}}{\partial \boldsymbol{x}}\right) = -\left(\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{y}}\right)^{-1} \left(\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\right) = \begin{pmatrix} \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_M} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_M}{\partial y_1} & \cdots & \frac{\partial f_M}{\partial y_M} \end{pmatrix}^{-1} \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_{N-M}} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_M}{\partial x_1} & \cdots & \frac{\partial f_M}{\partial x_{N-M}} \end{pmatrix}. \tag{3}$$

Proof. The proof follows exactly the same idea as the $\mathbb{R}^2$ case. We only emphsize the difference here. Denote

$$A := \left(\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{y}}\right)(\boldsymbol{x}_0, \boldsymbol{y}_0). \tag{4}$$

To make the presentation easier we pre-process as follows. Set $\boldsymbol{F}(\boldsymbol{x}, \boldsymbol{y}) := A^{-1}\, \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y})$. Then it is easy to verify that it suffices to work with $\boldsymbol{F}$ and furthermore $\left(\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{y}}\right)(\boldsymbol{x}_0, \boldsymbol{y}_0) = I$ the identity matrix. We choose δ_1, δ_2 small to satisfy the following:

1. δ_2 small enough so that

$$\frac{\partial f_i}{\partial y_j} < \frac{1}{2\, M^2}, \qquad \left|1 - \frac{\partial f_i}{\partial y_i}\right| < \frac{1}{2\, M^2} \tag{5}$$

   for all $\|(\boldsymbol{x}, \boldsymbol{y}) - (\boldsymbol{x}_0, \boldsymbol{y}_0)\| < 2\, \delta_2$.

2. Fix the above δ_2. Now we choose $\delta_1 \leqslant \delta_2$ such that for all $\boldsymbol{x} \in B(\boldsymbol{x}_0, \delta_1)$, $\|\boldsymbol{F}(\boldsymbol{x}, \boldsymbol{y}_0)\| < \frac{1}{2}$.

Now we first try to show the existence of a unique $\boldsymbol{y}$ solving

$$\boldsymbol{F}(\boldsymbol{x},\boldsymbol{y})=\boldsymbol{0} \tag{6}$$

for all $\boldsymbol{x}\in B(\boldsymbol{x}_0,\delta_1)$. Fix $\boldsymbol{x}$. Denote

$$\boldsymbol{g}(\boldsymbol{y}):=\boldsymbol{F}(\boldsymbol{x},\boldsymbol{y}). \tag{7}$$

We still use the iteration

$$\boldsymbol{y}_n=\boldsymbol{y}_{n-1}-\boldsymbol{g}(\boldsymbol{y}_{n-1}). \tag{8}$$

Now we have

$$\boldsymbol{y}_n-\boldsymbol{y}_{n-1}=\boldsymbol{y}_{n-1}-\boldsymbol{y}_{n-2}-[\boldsymbol{g}(\boldsymbol{y}_{n-1})-\boldsymbol{g}(\boldsymbol{y}_{n-2})]. \tag{9}$$

Note that the difference now is that we do not have one single $\boldsymbol{\xi}$ such that

$$\boldsymbol{g}(\boldsymbol{y}_{n-1})-\boldsymbol{g}(\boldsymbol{y}_{n-2})=\left(\frac{\partial\boldsymbol{g}}{\partial\boldsymbol{y}}\right)(\boldsymbol{\xi})\,(\boldsymbol{y}_{n-1}-\boldsymbol{y}_{n-2}). \tag{10}$$

However we still have the following mean value theorem:

$$g_i(\boldsymbol{y}_{n-1})-g_i(\boldsymbol{y}_{n-2})=(\operatorname{grad} g)(\boldsymbol{\xi})\cdot(\boldsymbol{y}_{n-1}-\boldsymbol{y}_{n-2}). \tag{11}$$

This way we still could prove that $\{\boldsymbol{y}_n\}$ is Cauchy. □

Exercise 1. Complete the proof of the theorem.

Example 2. Consider the system

$$\begin{aligned} x_1\,y_2-4\,x_2+2\,e^{y_1}+3 &= 0 \quad (12)\\ 2\,x_1-x_3-6\,y_1+y_2\cos y_1 &= 0 \quad (13)\end{aligned}$$

Calculate the Jacobian of the implicit function $\boldsymbol{Y}(\boldsymbol{x})$ at $x_1=-1, x_2=1, x_3=-1, y_1=0, y_2=1$.

Solution. Let

$$\boldsymbol{F}(\boldsymbol{x},\boldsymbol{y}):=\begin{pmatrix} x_1\,y_2-4\,x_2+2\,e^{y_1}+3\\ 2\,x_1-x_3-6\,y_1+y_2\cos y_1\end{pmatrix}. \tag{14}$$

Then we have

$$\left(\frac{\partial\boldsymbol{F}}{\partial\boldsymbol{x}}\right)=\begin{pmatrix} y_2 & -4 & 0\\ 2 & 0 & -1\end{pmatrix},\qquad \left(\frac{\partial\boldsymbol{F}}{\partial\boldsymbol{y}}\right)=\begin{pmatrix} 2\,e^{y_1} & x_1\\ -6-y_2\sin y_1 & \cos y_1\end{pmatrix}. \tag{15}$$

At the specified point we have

$$\left(\frac{\partial\boldsymbol{F}}{\partial\boldsymbol{x}}\right)=\begin{pmatrix} 1 & -4 & 0\\ 2 & 0 & -1\end{pmatrix},\qquad \left(\frac{\partial\boldsymbol{F}}{\partial\boldsymbol{y}}\right)=\begin{pmatrix} 2 & -1\\ -6 & 1\end{pmatrix}. \tag{16}$$

We see that

$$\left(\frac{\partial\boldsymbol{Y}}{\partial\boldsymbol{x}}\right)=\frac{1}{4}\begin{pmatrix} 3 & -4 & -1\\ 10 & -24 & -2\end{pmatrix}. \tag{17}$$

Example 3. Let $z = f(x, y),\ g(x, y) = 0$. Calculate $\frac{\mathrm{d}z}{\mathrm{d}x}$.

Solution. Let

$$\boldsymbol{F}(x, y, z) = \begin{pmatrix} f(x, y) - z \\ g(x, y) \end{pmatrix}. \tag{18}$$

Then we have

$$\left(\frac{\partial \boldsymbol{F}}{\partial x}\right) = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial g}{\partial x} \end{pmatrix}, \qquad \left(\frac{\partial \boldsymbol{F}}{\partial (y, z)}\right) = \begin{pmatrix} \frac{\partial f}{\partial y} & -1 \\ \frac{\partial g}{\partial y} & 0 \end{pmatrix} \tag{19}$$

which gives

$$\frac{\partial (Y, Z)}{\partial x} = -\begin{pmatrix} \frac{\partial f}{\partial y} & -1 \\ \frac{\partial g}{\partial y} & 0 \end{pmatrix}^{-1} \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial g}{\partial x} \end{pmatrix} = -\frac{1}{\frac{\partial g}{\partial y}} \begin{pmatrix} 0 & 1 \\ -\frac{\partial g}{\partial y} & \frac{\partial f}{\partial y} \end{pmatrix} \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial g}{\partial x} \end{pmatrix}. \tag{20}$$

Finally we have

$$\frac{\mathrm{d}Z}{\mathrm{d}x} = \frac{1}{\frac{\partial g}{\partial y}} \det \frac{\partial (f, g)}{\partial (x, y)}. \tag{21}$$

Theorem 4. (Inverse function theorem) *Let* $\boldsymbol{f}: \mathbb{R}^N \mapsto \mathbb{R}^N$ *satisfy*

i. $\boldsymbol{f}$ *is differentiable with continuous partial derivatives;*

ii. $\boldsymbol{f}(\boldsymbol{y}_0) = \boldsymbol{x}_0$*;*

iii. $\det\left(\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{y}}\right)(\boldsymbol{y}_0) \neq 0$.

Then there are two open sets U, V *such that* $\boldsymbol{x}_0 \in U$, $\boldsymbol{y}_0 \in V$ *and there is a function* $\boldsymbol{g}: U \mapsto V$ *which is the inverse of* $\boldsymbol{f}$*. Furthermore we have*

$$\left(\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{x}}\right)(\boldsymbol{x}_0, \boldsymbol{y}_0) = \left[\left(\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{y}}\right)(\boldsymbol{x}_0, \boldsymbol{y}_0)\right]^{-1}. \tag{22}$$

Proof. Most of the theorem follow directly from implicit function theorem, from which we obtain the existence of $I, J, \boldsymbol{g}$ such that

$$\boldsymbol{f}(\boldsymbol{g}(\boldsymbol{x})) = \boldsymbol{x} \tag{23}$$

for all $\boldsymbol{x} \in I$, with $\boldsymbol{g}(\boldsymbol{x})$ unique and belong to J.

Notice that to show $\boldsymbol{g}$ is the inverse, we need to further check the following: There is $V \subseteq J$ open such that $\boldsymbol{f}$ is one-to-one on V. (Think: Why do we need this?)

We take $V = \boldsymbol{f}^{-1}(U) \cap J$. Since $\boldsymbol{f}$ is continuous, $\boldsymbol{f}^{-1}(U)$ is open and thus V is open.

Now we check one-to-one. Assume that $\boldsymbol{f}(\boldsymbol{y}_1) = \boldsymbol{f}(\boldsymbol{y}_2)$. Then we know there are $\boldsymbol{\xi_1}, \ldots, \boldsymbol{\xi_N}$ such that

$$(\text{grad } f_i)(\boldsymbol{\xi}_i) \cdot (\boldsymbol{y}_1 - \boldsymbol{y}_2) = 0. \tag{24}$$

If we set

$$A := \begin{pmatrix} (\text{grad } f_1)(\boldsymbol{\xi}_1)^T \\ \vdots \\ (\text{grad } f_N)(\boldsymbol{\xi}_N)^T \end{pmatrix}, \tag{25}$$

we would have

$$A\,(\boldsymbol{y}_1 - \boldsymbol{y}_2) = 0. \tag{26}$$

But by our choice of U, $\det A \neq 0$. Consequently $\boldsymbol{y}_1 - \boldsymbol{y}_2 = 0 \Longrightarrow \boldsymbol{y}_1 = \boldsymbol{y}_2$. □

Exercise 2. Let A, B be sets. $f: A \mapsto B$ a function. If there is $g: B \mapsto A$ such that

$$f(g(y)) = y \tag{27}$$

for all $y \in B$, can we say g is an inverse of f? What if we further assume f is one-to-one?

Exercise 3. (Polar coordinates) Let $x = r\cos\theta$, $y = r\sin\theta$. Calculate $\left(\frac{\partial(r,\theta)}{\partial(x,y)}\right)$.

Exercise 4. Let $x = r\cos\theta$, $y = r\sin\theta$.

a) Show that $\det\left(\frac{\partial(x,y)}{\partial(r,\theta)}\right) \neq 0$ for all $r > 0$.

b) Does the inverse function exist globally?

Problem 1. Let $\boldsymbol{f}: \mathbb{R}^N \mapsto \mathbb{R}^N$ be differentiable with continuous partial derivatives. Assume that $\det J_f(\boldsymbol{x}_0) \neq 0$. Then there is $r > 0$ such that for any open set $U \subseteq B(\boldsymbol{x}_0, r)$, $\boldsymbol{f}(U)$ is open.