General implicit and inverse function theorems

Theorem 1. (Implicit function theorem) Let \( f: \mathbb{R}^N \rightarrow \mathbb{R}^M \) with \( N > M \). We decompose
\[
\mathbb{R}^N = \mathbb{R}^{N-M} \times \mathbb{R}^M
\]
and denote the first \( N - M \) coordinates by vector \( x \) and the rest \( M \) coordinates by \( y \).

Assume
i. \( f \) is differentiable and has continuous partial derivatives;
ii. \( f(x_0, y_0) = 0 \);
iii. \( \det \left( \frac{\partial f}{\partial y} \right)(x_0, y_0) \neq 0 \) where the Jacobian matrix with respect to \( y \) is defined as
\[
\left( \frac{\partial f}{\partial y} \right) := \begin{pmatrix}
\frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_M} \\
\vdots & \ddots & \vdots \\
\frac{\partial f_M}{\partial y_1} & \cdots & \frac{\partial f_M}{\partial y_M}
\end{pmatrix}.
\]

Then there are open sets \( U \subseteq \mathbb{R}^{N-M}, V \subseteq \mathbb{R}^M \) satisfying \( x_0 \in U, y_0 \in V \) and
i. For every \( x \in U \) the equation \( f(x, y) = 0 \) has one unique solution \( y = Y(x) \in V \);
ii. \( Y(x_0) = y_0 \);
iii. \( Y \) is differentiable with continuous partial derivatives;
iv. For \( x \in U \),
\[
\left( \frac{\partial Y}{\partial x} \right) = -\left( \frac{\partial f}{\partial y} \right)^{-1} \left( \frac{\partial f}{\partial x} \right) = \begin{pmatrix}
\frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_{N-M}} \\
\vdots & \ddots & \vdots \\
\frac{\partial f_M}{\partial x_1} & \cdots & \frac{\partial f_M}{\partial x_{N-M}}
\end{pmatrix}^{-1} \begin{pmatrix}
\frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_M} \\
\vdots & \ddots & \vdots \\
\frac{\partial f_M}{\partial y_1} & \cdots & \frac{\partial f_M}{\partial y_M}
\end{pmatrix}.
\]

Proof. The proof follows exactly the same idea as the \( \mathbb{R}^2 \) case. We only emphasize the difference here. Denote
\[
A := \left( \frac{\partial f}{\partial y} \right)(x_0, y_0).
\]

To make the presentation easier we pre-process as follows. Set \( F(x, y) := A^{-1} f(x, y) \). Then it is easy to verify that it suffices to work with \( F \) and furthermore \( \left( \frac{\partial F}{\partial y} \right)(x_0, y_0) = I \) the identity matrix. We choose \( \delta_1, \delta_2 \) small to satisfy the following:

1. \( \delta_2 \) small enough so that
\[
\frac{\partial f_i}{\partial y_j} < \frac{1}{2M^2}, \quad \left| 1 - \frac{\partial f_i}{\partial y_i} \right| < \frac{1}{2M^2}
\]
for all \( \| (x, y) - (x_0, y_0) \| < 2 \delta_2 \).
2. Fix the above \( \delta_2 \). Now we choose \( \delta_1 \leq \delta_2 \) such that for all \( x \in B(x_0, \delta_1), \| F(x, y_0) \| < \frac{1}{2} \).
Now we first try to show the existence of a unique \( y \) solving

\[
F(x, y) = 0
\]  

(6)

for all \( x \in B(x_0, \delta_1) \). Fix \( x \). Denote

\[
g(y) := F(x, y).
\]  

(7)

We still use the iteration

\[
y_n = y_{n-1} - g(y_{n-1}).
\]  

(8)

Now we have

\[
y_n - y_{n-1} = y_{n-1} - y_{n-2} - [g(y_{n-1}) - g(y_{n-2})].
\]  

(9)

Note that the difference now is that we do not have one single \( \xi \) such that

\[
g(y_{n-1}) - g(y_{n-2}) = \left( \frac{\partial g}{\partial y} \right)(\xi) (y_{n-1} - y_{n-2}).
\]  

(10)

However we still have the following mean value theorem:

\[
g_i(y_{n-1}) - g_i(y_{n-2}) = (\text{grad } g)(\xi) \cdot (y_{n-1} - y_{n-2}).
\]  

(11)

This way we still could prove that \( \{y_n\} \) is Cauchy. □

**Exercise 1.** Complete the proof of the theorem.

**Example 2.** Consider the system

\[
\begin{align*}
x_1 y_2 - 4 x_2 + 2 e^{y_1} + 3 &= 0 \quad (12) \\
2 x_1 - x_3 - 6 y_1 + y_2 \cos y_1 &= 0 \quad (13)
\end{align*}
\]

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

**Solution.** Let

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

(14)

Then we have

\[
\begin{pmatrix}
\frac{\partial F}{\partial x} \\
\frac{\partial F}{\partial y}
\end{pmatrix} = \begin{pmatrix}
y_2 & -4 & 0 \\
2 & 0 & -1
\end{pmatrix}, \quad \begin{pmatrix}
\frac{\partial F}{\partial x} \\
\frac{\partial F}{\partial y}
\end{pmatrix} = \begin{pmatrix}
2 e^{y_1} & x_1 \\
-6 - y_2 \sin y_1 & \cos y_1
\end{pmatrix}.
\]  

(15)

At the specified point we have

\[
\begin{pmatrix}
\frac{\partial F}{\partial x} \\
\frac{\partial F}{\partial y}
\end{pmatrix} = \begin{pmatrix}
1 & -4 & 0 \\
2 & 0 & -1
\end{pmatrix}, \quad \begin{pmatrix}
\frac{\partial F}{\partial x} \\
\frac{\partial F}{\partial y}
\end{pmatrix} = \begin{pmatrix}
2 & -1 \\
-6 & 1
\end{pmatrix}.
\]  

(16)

We see that

\[
\begin{pmatrix}
\frac{\partial Y}{\partial x} \\
\frac{\partial Y}{\partial y}
\end{pmatrix} = \frac{1}{4} \begin{pmatrix}
3 & -4 & -1 \\
10 & -24 & -2
\end{pmatrix}.
\]  

(17)
Example 3. Let \( z = f(x, y), g(x, y) = 0 \). Calculate \( \frac{dz}{dx} \).

Solution. Let

\[ F(x, y, z) = \left( \begin{array}{c} f(x, y) - z \\ g(x, y) \end{array} \right). \]

Then we have

\[ \left( \frac{\partial F}{\partial x} \right) = \left( \begin{array}{c} \frac{\partial f}{\partial x} \\ \frac{\partial g}{\partial x} \end{array} \right), \quad \left( \frac{\partial F}{\partial (y, z)} \right) = \left( \begin{array}{c} \frac{\partial f}{\partial y} - 1 \\ \frac{\partial g}{\partial y} 0 \end{array} \right) \]

which gives

\[ \frac{\partial (Y, Z)}{\partial x} = - \left( \begin{array}{c} \frac{\partial f}{\partial y} - 1 \\ \frac{\partial g}{\partial y} 0 \end{array} \right)^{-1} \left( \begin{array}{c} \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial y} \end{array} \right) \cdot \left( \begin{array}{c} 0 1 \\ - \frac{\partial g}{\partial y} \frac{\partial f}{\partial y} \end{array} \right). \]

Finally we have

\[ \frac{dZ}{dx} = \frac{1}{\frac{\partial g}{\partial y}} \det \left( \frac{\partial (f, g)}{\partial (x, y)} \right). \]

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

i. \( f \) is differentiable with continuous partial derivatives;

ii. \( f(y_0) = x_0 \);

iii. \( \det \left( \frac{\partial f}{\partial y} \right)(y_0) \neq 0 \).

Then there are two open sets \( U, V \) such that \( x_0 \in U \), \( y_0 \in V \) and there is a function \( g: U \mapsto V \) which is the inverse of \( f \). Furthermore we have

\[ \left( \frac{\partial g}{\partial x} \right)(x_0, y_0) = \left[ \left( \frac{\partial f}{\partial y} \right)(x_0, y_0) \right]^{-1}. \]

Proof. Most of the theorem follow directly from implicit function theorem, from which we obtain the existence of \( I, J, g \) such that

\[ f(g(x)) = x \]

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

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

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

Now we check one-to-one. Assume that \( f(y_1) = f(y_2) \). Then we know there are \( \xi_1, \ldots, \xi_N \) such that

\[ (\text{grad } f_i)(\xi_i) \cdot (y_1 - y_2) = 0. \]

If we set

\[ A := \begin{pmatrix} (\text{grad } f_1)(\xi_1)^T \\ \vdots \\ (\text{grad } f_N)(\xi_N)^T \end{pmatrix}, \]

then
we would have
\[ A \left( y_1 - y_2 \right) = 0. \]  
(26)

But by our choice of \( U \), \( \det A \neq 0 \). Consequently \( y_1 - y_2 = 0 \implies y_1 = y_2 \).

**Exercise 2.** Let \( A, B \) be sets. \( f: A \to B \) a function. If there is \( g: B \to A \) such that
\[ f(g(y)) = y \]
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 \( \frac{\partial (r, \theta)}{\partial (x, y)} \).

**Exercise 4.** Let \( x = r \cos \theta, y = r \sin \theta \).

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

b) Does the inverse function exist globally?

**Problem 1.** Let \( f: \mathbb{R}^N \to \mathbb{R}^N \) be differentiable with continuous partial derivatives. Assume that \( \det J_f(x_0) \neq 0 \). Then there is \( r > 0 \) such that for any open set \( U \subseteq B(x_0, r) \), \( f(U) \) is open.