MATH 2010 Chapter 13

13.1 Implicit Function Theorem

Theorem 13.1 (Implicit Function Theorem)**.** *Let* $\Omega \subseteq \mathbb{R}^{n+k}$ *be open,* $F : \Omega \to \mathbb{R}^k$ *be* C^1

Denote $x = (x_1, \cdots, x_n), y = (y_1, \cdots, y_k)$

$$\begin{aligned} F(\vec{x}, \vec{y}) &= \begin{bmatrix} F_1(\vec{x}, \vec{y}) \\ \vdots \\ F_k(\vec{x}, \vec{y}) \end{bmatrix} \\ &= \begin{bmatrix} F_1(x_1, \cdots, x_n, y_1, \cdots, y_k) \\ \vdots \\ F_k(x_1, \cdots, x_n, y_1, \cdots, y_k) \end{bmatrix} \end{aligned}$$

Suppose $(a, b) \in \Omega$*, where* $a = \vec{a} \in \mathbb{R}^n$*,* $b = \vec{b} \in \mathbb{R}^k$*, such that*

$$F(a, b) = \vec{c} = \begin{bmatrix} c_1 \\ \vdots \\ c_k \end{bmatrix} \in \mathbb{R}^k$$

and the $k \times k$ *matrix*

$$(D_{\vec{y}}F)\,|_{(a,b)} := \left[\frac{\partial F_i}{\partial y_j}(a, b)\right]_{1 \leqslant i,j \leqslant k} = \begin{bmatrix} \frac{\partial F_1}{\partial y_1}(a, b) & \cdots & \frac{\partial F_1}{\partial y_k}(a, b) \\ \vdots & & \vdots \\ \frac{\partial F_k}{\partial y_1}(a, b) & \cdots & \frac{\partial F_k}{\partial y_k}(a, b) \end{bmatrix} \text{ is invertible.}$$

Then, there exist open sets $U \subseteq \mathbb{R}^n$ *containing* a *,* $V \subseteq \mathbb{R}^k$ *containing* b *and a* C^1 *function*

$$\varphi : U \to V$$
$$x = (x_1, \cdots, x_n) \in U \to y = (y_1, \cdots, y_k) \in V$$

such that

1. $\varphi(a) = b$

2. $F(x, \varphi(x)) = c$

3. For any $(x, y) \in U \times V$ *such that* $F(x, y) = c$*, we have* $y = \varphi(x)$*.*

4. For $1 \leq i \leq k$*,* $1 \leq j \leq n$*,*

$$\left[\frac{\partial y_i}{\partial x_j}(a)\right] = \left[\frac{\partial \varphi_i}{\partial x_j}(a)\right] = -\left((D_{\vec{y}}F)^{-1}\, D_{\vec{x}}F\right)_{ij}$$

Remark. Provided that we know $\vec{y} = \vec{y}(\vec{x})$ is a differentiable function of $\vec{x}$, the last item follows from the Chain Rule.

If $\vec{y}$ is a differentiable function of $\vec{x}$, then $(\vec{x}, \vec{y})$ may be viewed as a vector-valued function:

$$\vec{\phi} : \mathbb{R}^n \longrightarrow \mathbb{R}^{n+k}$$

where:

$$\vec{\phi}(\vec{x}) = \begin{bmatrix} \vec{x} \\ \vec{y}(\vec{x}) \end{bmatrix}$$

Applying the chain rule for differentiation with respect to $\vec{x}$ to both sides of:

$$\underbrace{F(\vec{x}, \vec{y})}_{F\left(\vec{\phi}(\vec{x})\right)} = \vec{c},$$

we have:

$$\underbrace{\left[\begin{array}{ccc|ccc} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_n} & \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_k} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_k}{\partial x_n} & \frac{\partial F_k}{\partial y_1} & \cdots & \frac{\partial F_k}{\partial y_k} \end{array}\right]}_{DF = \left[D_{\vec{x}}F \,\mid\, D_{\vec{y}}F\right]} \underbrace{\left[\begin{array}{ccc} \frac{\partial \phi_1}{\partial x_1} & \cdots & \frac{\partial \phi_1}{\partial x_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial \phi_n}{\partial x_1} & \cdots & \frac{\partial \phi_n}{\partial x_n} \\ \hline \frac{\partial \phi_{n+1}}{\partial x_1} & \cdots & \frac{\partial \phi_{n+1}}{\partial x_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial \phi_{n+k}}{\partial x_1} & \cdots & \frac{\partial \phi_{n+k}}{\partial x_n} \end{array}\right]}_{D\vec{\phi}} = \vec{0}$$

$$\begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_k}{\partial x_n} \end{bmatrix} I_n + \begin{bmatrix} \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_k} \\ \vdots & \vdots & \vdots \\ \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_k}{\partial y_k} \end{bmatrix} \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial y_k}{\partial x_1} & \cdots & \frac{\partial y_k}{\partial x_n} \end{bmatrix} = \vec{0}$$

Hence,

$$\begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial y_k}{\partial x_1} & \cdots & \frac{\partial y_k}{\partial x_n} \end{bmatrix} = -\begin{bmatrix} \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_k} \\ \vdots & \vdots & \vdots \\ \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_k}{\partial y_k} \end{bmatrix}^{-1} \begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \vdots & \vdots & \vdots \\ \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_k}{\partial x_n} \end{bmatrix}$$

Remark. Applying the theorem to the special case where we have a real-valued differentiable function $F : \mathbb{R}^2 \longrightarrow \mathbb{R}$ satisfying:

$$F(x, y) = c,$$

we have:

$$\frac{\partial y}{\partial x} = -\frac{F_x}{F_y}$$

Example 13.2. $n = 2, k = 1$.

Consider the surface (a sphere to be precise) desribed by the following equation:

$$x^2 + y^2 + z^2 = 2.$$

Is z implicitly a function of (x, y) near the point $(0, 1, 1)$ on the surface?

If so, what is $\dfrac{\partial z}{\partial x}, \dfrac{\partial z}{\partial y}$ at $(x, y, z) = (0, 1, 1)$?

Solution. The equation describing the surface is equivalent to $F(x, y, z) = 2$, where:

$$F(x, y, z) = x^2 + y^2 + z^2 : \mathbb{R}^3 \longrightarrow \mathbb{R}^1$$

Observe that F is C^1 on $\mathbb{R}^3$. In the context of the IFT, we have $k = 1, n = 3-1 = 2$.

So, it is possible that any one ($k = 1$) of (x, y, z) is implicitly a function of the other two ($n = 2$).

We have:

$$D_{(z)}F = [F_z] = [2z].$$

At $(0, 1, 1)$, we have $[F_z]|_{(0,1,1)} = [2 \cdot 1] = [2]$, which is an invertible 1×1 matrix, with inverse:

$$\left(D_{(z)}F|_{(0,1,1)}\right)^{-1} = [1/2]$$

Hence, the conditions of IFT are satisfied. The variable z is implicitly a C^1 function of (x, y) near $(0, 1, 1)$ on the surface.

Moreover, we have:

$$\begin{aligned}
\begin{bmatrix} \frac{\partial z}{\partial x} & \frac{\partial z}{\partial y} \end{bmatrix}\Big|_{(0,1,1)} &= -\left(D_{(z)}F|_{(0,1,1)}\right)^{-1} \begin{bmatrix} F_x & F_y \end{bmatrix}|_{(0,1,1)} \\
&= -[1/2] \begin{bmatrix} F_x & F_y \end{bmatrix}|_{(0,1,1)} \\
&= -[1/2] \begin{bmatrix} 2x & 2y \end{bmatrix}|_{(0,1,1)} \\
&= -[1/2] \begin{bmatrix} 0 & 2 \end{bmatrix}
\end{aligned}$$

Hence:
$$\frac{\partial z}{\partial x}|_{(0,1,1)} = -\frac{1}{2}\cdot 0 = 0, \quad \frac{\partial z}{\partial y}|_{(0,1,1)} = -\frac{1}{2}\cdot 2 = -1$$

Example 13.3. $n = 1, k = 2$ Consider the curve (a circle) which is the intersection the following two surfaces:

$$\begin{cases} x^2 + y^2 + z^2 & = 2 \\ x + z & = 1 \end{cases}$$

IFRAME

What does IFT say in this case?

Define $F : \mathbb{R}^{2+1} \longrightarrow \mathbb{R}^2$ as follows:

$$F(x, y, z) = \begin{bmatrix} F_1 \\ F_2 \end{bmatrix} = \begin{bmatrix} x^2 + y^2 + z^2 \\ x + z \end{bmatrix}$$

Then, the curve in question corresponds to the constraint:

$$F(x, y, z) = \vec{c} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Since the codomain of F is $\mathbb{R}^2$ ($k = 2$), and F has 3 variables, by IFT, any two ($k = 2$) of the variables of F could implicitly be a function of the remaining one variable ($n = 3 - k = 1$).

We have:
$$D_{(y,z)}F = \begin{bmatrix} F_{1,y} & F_{1,z} \\ F_{2,y} & F_{2,z} \end{bmatrix} = \begin{bmatrix} 2y & 2z \\ 0 & 1 \end{bmatrix}$$

Hence, at, for example, the point $(0, 1, 1)$, we have:

$$D_{(y,z)}F\big|_{(0,1,1)} = \begin{bmatrix} 2 & 2 \\ 0 & 1 \end{bmatrix},$$

which is invertible.

Hence, by IFT, on the curve the variables y, z are implicitly differentiable functions of x near the point $(0, 1, 1)$, with:

$$\begin{bmatrix} \frac{\partial y}{\partial x} \\ \frac{\partial z}{\partial x} \end{bmatrix} = -\left. D_{(y,z)}F\right|^{-1}_{(0,1,1)} \left.\begin{bmatrix} F_{1,x} \\ F_{2,x} \end{bmatrix}\right|_{(0,1,1)} = -\frac{1}{2}\begin{bmatrix} 1 & -2 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 2\cdot 0 \\ 1 \end{bmatrix}$$
$$= \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Remark. There is no intrinsic ordering to the variables x, y, z, so we might equally ask whether x, y are implicitly functions of z on the curve.

Since

$$D_{(x,y)}F\big|_{(0,1,1)} = \begin{bmatrix} F_{1,x} & F_{1,y} \\ F_{2,x} & F_{2,y} \end{bmatrix}\Bigg|_{(0,1,1)} = \begin{bmatrix} 2x & 2y \\ 1 & 0 \end{bmatrix}\Bigg|_{(0,1,1)} = \begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix}$$

is invertible, the IFT says that on the curve the variable x, y are implicity functions of z near the point $(0, 1, 1)$, with:

$$\begin{aligned} \begin{bmatrix} \frac{\partial x}{\partial z} \\ \frac{\partial y}{\partial z} \end{bmatrix} &= -\, D_{(x,y)}F\big|_{(0,1,1)}^{-1} \begin{bmatrix} F_{1,z} \\ F_{2,z} \end{bmatrix}\Bigg|_{(0,1,1)} = \frac{1}{2}\begin{bmatrix} 0 & -2 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} 2\cdot 1 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} -1 \\ -1 \end{bmatrix} \end{aligned}$$

Example 13.4. Consider the constraints

$$\begin{cases} xz + \sin(yz - x^2) &= 8 \\ x + 4y + 3z &= 18 \end{cases}$$

Near $(2, 1, 4)$, can we express 2 of the variables as functions of the remaining variable?

Solution. Let $F(x, y, z) = \begin{bmatrix} F_1(x, y, z) \\ F_2(x, y, z) \end{bmatrix} = \begin{bmatrix} xz + \sin(yz - x^2) \\ x + 4y + 3z \end{bmatrix}$.

Then:

$$DF = \begin{bmatrix} z - 2x\cos(yz - x^2) & z\cos(yz - x^2) & x + y\cos(yz - x^2) \\ 1 & 4 & 3 \end{bmatrix}$$

$$DF(2, 1, 4) = \begin{bmatrix} 0 & 4 & 3 \\ 1 & 4 & 3 \end{bmatrix}$$

$D_{(x,y)}F|_{(2,1,4)}$:

$$\begin{vmatrix} 0 & 4 \\ 1 & 4 \end{vmatrix} = -4 \neq 0 \overset{IFT}{\Rightarrow} x, y \text{ can be expressed as functions of } z \text{ near } (2, 1, 4)$$

$D_{(x,z)}F|_{(2,1,4)}$:

$$\begin{vmatrix} 0 & 3 \\ 1 & 3 \end{vmatrix} = -3 \neq 0 \overset{IFT}{\Rightarrow} x, z \text{ can be expressed as functions of } y \text{ near } (2, 1, 4)$$

$D_{(y,z)}F|_{(2,1,4)}$:

$$\begin{vmatrix} 4 & 3 \\ 4 & 3 \end{vmatrix} = 0$$

Hence, $D_{(y,z)}F|_{(2,1,4)}$ is not invertible. We may not conclude from the IFT whether y, z are locally functions of x near $(2, 1, 4)$.

Remark. The variables y, z are in fact not differentiable functions of x near $(2, 1, 4)$.

Otherwise, by implicit differentiation, we have:

$$\begin{aligned} D_{(y,z)}F|_{(2,1,4)} \left.\begin{bmatrix} \frac{dy}{dx} \\ \frac{dz}{dx} \end{bmatrix}\right|_{(2,1,4)} &= -D_xF|_{(2,1,4)} = -\begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ \begin{bmatrix} 4 & 3 \\ 4 & 3 \end{bmatrix} \left.\begin{bmatrix} \frac{dy}{dx} \\ \frac{dz}{dx} \end{bmatrix}\right|_{(2,1,4)} &= -\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \end{aligned}$$

which has no solutions, i.e. not satisfied for any values of $\frac{\partial y}{\partial x}|_{(2,1,4)}, \frac{\partial z}{\partial x}|_{(2,1,4)}$.

Hence, y, z cannot be differentiable functions of x near $(2, 1, 4)$.

Remark. Implicit Function Theorem has many important applications, such as rigorous proofs of

1. Implicit differentiation
2. Tangent plane of surface $F(x, y, z) = c$
3. Lagrange Multipliers

13.2 Inverse Function Theorem

Theorem 13.5 (Inverse Function Theorem)**.** *Let $\Omega \subseteq \mathbb{R}^n$ be open.*

$f : \Omega \to \mathbb{R}^n$ be $C^1, f(a) = b$.

Suppose $Df(a)$ is an invertible $n \times n$ matrix

Then, there exist open sets $U_1, U_2 \subseteq \mathbb{R}^n, a \in U_1, b \in U_2$ and a C^1 function $g : U_2 \to U_1$ such that:

1. *$g(b) = a$*
2. *$g(f(x)) = x \quad$ for all $x \in U_1$*

   $f(g(y)) = y \quad$ for all $y \in U_2$

   (g is a local inverse of f : $g = (f|_{U_1})^{-1}$)

3. $Dg(b) = Df(a)^{-1}$

Remark. The Inverse Function Theorem is equivalent to Implicit Function Theorem.

Idea:

$\Rightarrow$: Given $F(x, y) = c$, where $x \in \mathbb{R}^n$ and $y, c \in \mathbb{R}^k$, apply the Inverse Function Theorem to: $H(x, y) = (x, F(x, y)) : \mathbb{R}^{n+k} \longrightarrow \mathbb{R}^{n+k}$.

$\Leftarrow$: Given $H : \mathbb{R}^n \longrightarrow \mathbb{R}^n$, apply the Implicit Function Theorem to $F(x, y) = 0$, where $F(x, y) = y - H(x) : \mathbb{R}^{2n} \longrightarrow \mathbb{R}^n$.

Example 13.6. $f : \mathbb{R}^2 \to \mathbb{R}^2$

$f(x, y) = (x^2 - y^2, 2xy)$

Clearly $f(-x, -y) = f(x, y)$

$\Rightarrow f$ is not injective and has no global inverse.

How about local inverse?

Solution.

$$f(x, y) = \begin{bmatrix} x^2 - y^2 \\ 2xy \end{bmatrix}$$

$$Df(x, y) = \begin{bmatrix} 2x & -2y \\ 2y & 2x \end{bmatrix}$$

$$\det Df(x, y) = 4(x^2 + y^2) > 0 \Leftrightarrow (x, y) \neq (0, 0)$$

By the Inverse Function Theorem, f is locally invertible with differentiable local inverse.

For instance, let $g(u, v)$ be a local inverse of $f(x, y)$ near $(x, y) = (1, -1)$.

Then $f(1, -1) = (0, -2) \Rightarrow g(0, -2) = (1, -1)$, and:

$$Dg(0, -2) = Df(1, -1)^{-1} = \begin{bmatrix} 2 & 2 \\ -2 & 2 \end{bmatrix}^{-1} = \frac{1}{4} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

In fact, we can find $g(u, v)$ explicitly.

Let:

$$\begin{cases} u &= x^2 - y^2 \\ v &= 2xy \end{cases}$$

Near $(x, y) = (1, -1), x \neq 0 \Rightarrow y = \frac{v}{2x}$

$\therefore u = x^2 - (\frac{v}{2x})^2$

$$4x^4 - 4ux^2 - v^2 = 0$$

$$\begin{aligned} x^2 &= \frac{4u \pm \sqrt{(-4u)^2 - 4(4)(-v^2)}}{8} \\ &= \frac{u \pm \sqrt{u^2 + v^2}}{2} \end{aligned}$$

Let $(x, y) = (1, -1)$, then $(u, v) = (0, -2)$.

$$\Rightarrow 1^2 = \frac{0 \pm \sqrt{4}}{2}$$

So, we may reject the negative sign, and it follows that:

$$x^2 = \frac{u + \sqrt{u^2 + v^2}}{2}$$

$$\Rightarrow x = \sqrt{\frac{u + \sqrt{u^2 + v^2}}{2}}, y = \frac{2v}{x} = \frac{2\sqrt{2}v}{\sqrt{u + \sqrt{u^2 + v^2}}}$$

Hence,

$$g(u, v) = \left(\sqrt{\frac{u + \sqrt{u^2 + v^2}}{2}}, \frac{2\sqrt{2}v}{\sqrt{u + \sqrt{u^2 + v^2}}} \right)$$

In both Implicit and Inverse Function Theorems,
we assume a Jacobi matrix to be invertible.

Without this assumption , the theorems are **inconclusive** on the existence of local implicit or inverse function. See the examples below:

Implicit Function Theorem

- $F(x, y) = x^2 - y^2 = 0$
  $\frac{\partial F}{\partial y}|_{(0,0)} = 0$
- $F(x, y) = x^3 - y^3 = 0$
  $\frac{\partial F}{\partial y}|_{(0,0)} = 0$
  $y = x$ locally (and in fact globally).

Inverse Function Theorem

- $f(x) = x^2$

  $f'(0) = 0$

  Not injective near $x = 0$.

  Hence, no local inverse near $x = 0$.

- $f(x) = x^3$

  $f'(0) = 0$

  The function f has a global inverse: $g(y) = \sqrt[3]{y}$