MATH2010E LECTURE 12: IMPLICIT FUNCTION THEOREM

1. Introduction

Recall that we have already discussed the following situation: Let $F(x,y)=0$ be a implicit function with F being a C^1 map from $\mathbb{R}^2$ to $\mathbb{R}$. Then we have the formula

$$\frac{dy}{dx}=-\frac{F_x}{F_y} \tag{1.1}$$

when $F_y\neq 0$. This shows us that y can be written as a function of x near a point p if and only if $F_y(p)\neq 0$.

Let us also recall that the way we obtain (1.1). Suppose there exists a variable s with $x=s$ and $y=y(s)$, we have

$$F_x+F_y\frac{dy}{dx}=0 \tag{1.2}$$

by using the chain rule. In the following paragraphs we will generalize this equation.

Let $F:\mathbb{R}^{m+n}\to\mathbb{R}^n$ be a C^1 function. So for any $x\in\mathbb{R}^m$, $y\in\mathbb{R}^n$, we can write $F(x,y)=(f_1(x,y),f_2(x,y),...,f_n(x,y))$ with all $f_i:\mathbb{R}^{m+n}\to\mathbb{R}$ are C^1.

Now, let us suppose that there exist parameters $s_1,...,s_n$ with $x_i=s_i$ and $y=y(s_1,...,s_n)$. Under this setting, we have

$$\begin{aligned}\frac{\partial F}{\partial s_i}&=\begin{pmatrix}\frac{\partial f_1}{\partial x_i}\frac{\partial x_i}{\partial s_i}\\ \frac{\partial f_2}{\partial x_i}\frac{\partial x_i}{\partial s_i}\\ \vdots\\ \frac{\partial f_n}{\partial x_i}\frac{\partial x_i}{\partial s_i}\end{pmatrix}+\begin{pmatrix}\frac{\partial f_1}{\partial y_1}\frac{\partial y_1}{\partial s_i} & \cdots & \frac{\partial f_1}{\partial y_n}\frac{\partial y_n}{\partial s_i}\\ \frac{\partial f_2}{\partial y_1}\frac{\partial y_1}{\partial s_i} & \cdots & \frac{\partial f_2}{\partial y_n}\frac{\partial y_n}{\partial s_i}\\ \vdots & \vdots & \vdots\\ \frac{\partial f_n}{\partial y_1}\frac{\partial y_1}{\partial s_i} & \cdots & \frac{\partial f_n}{\partial y_n}\frac{\partial y_n}{\partial s_i}\end{pmatrix}\\ &=F_x+[F_y]\begin{pmatrix}\frac{\partial y_1}{\partial x_i}\\ \frac{\partial y_2}{\partial x_i}\\ \vdots\\ \frac{\partial y_n}{\partial x_i}\end{pmatrix}=0\end{aligned} \tag{1.3}$$

for any $i=1,2,...,n$. Here $[F_y]$ is a matrix value function defined as the following

$$[F_y]=\begin{pmatrix}\frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_n}\\ \frac{\partial f_2}{\partial y_1} & \cdots & \frac{\partial f_2}{\partial y_n}\\ \vdots & \vdots & \vdots\\ \frac{\partial f_n}{\partial y_1} & \cdots & \frac{\partial f_n}{\partial y_n}\end{pmatrix}. \tag{1.4}$$

To obtain (1.1) from (1.2), divide the both sides of (1.1) by F_y when it is non zero. In this general case, the situation is similar. We have to "divide" (1.3) by $[F_y]$.

This idea leads the following theorem.

Theorem 1.1. *Let $F \in C^1$. Suppose $[F_y]$ is invertible at a point p, then we can write y as a function of x near p and*

$$\frac{\partial y}{\partial x_i} = \begin{pmatrix} \frac{\partial y_1}{\partial x_i} \\ \frac{\partial y_2}{\partial x_i} \\ \vdots \\ \frac{\partial y_n}{\partial x_i} \end{pmatrix} = -[F_y]^{-1}F_x \tag{1.5}$$

One simple example we can think of is the implicit function of $n-$dimensional unit sphere S^n.

$$S^n = \{x_1^2 + x_2^2 + \cdots + x_n^2 - 1 = 0\}. \tag{1.6}$$

We can see that by the implicit function theorem: If $F_{x_i}(p)$ is non zero for some $i = 1, 2, ..., n$, then x_i can be written as the function of other variables. So at $p = (0, 0, ..., 1, 0, ..., 0)$ which has 0 for all but $i-$th component, $F_{x_k}(p) = 0$ for all $k \neq i$ and $F_{x_i}(p) = 2$. In this case, we can conclude that x_i can be written as a function of other variables.

2. Proof

Proof. Here we only prove the case that F is a C^∞ (This assumption can actually be changed to $F \in C^1$). Also, in our proof, we always write F as a column vector.

Let $p = (p_1, p_2) \in \mathbb{R}^{m+n}$. for any $x \in \mathbb{R}^m$ near p_1, we can write down following formula by the error bound estimate.

$$0 = F(x, y) = F(x, p_2) + [F_y](x, p_2)(y - p_2)^t + E_2(y) \tag{2.1}$$

where $|E_2(y)| \leq C|y - p_2|^2$.

Recall by the assumption in the statement of Theorem 1.1, the matrix $[F_y](p)$ is invertible, so for any x which is sufficiently close to p_1, $[F_y](x, p_2)$ is invertible. Therefore, we have

$$(y - p_2)^t = [F_y]^{-1}(x, p_2)F(x, p_2) + [F_y]^{-1}(x, p_2)E_2(y). \tag{2.2}$$

Not that the last term $[F_y]^{-1}(x, p_2)E_2(y)$ is still bounded by some $C|y - p_2|^2$. We denote $[F_y]^{-1}(x, p_2)E_2(y)$ by $G(y - p_2)$. So by taking $y \in \mathbb{R}^n$ also sufficiently close to p_2, we have $|G(y - p_2)| \leq \frac{1}{100}|y - p_2|$.

Now, (1.8) implies that

$$(I - G)(y - p_2) = [F_y]^{-1}(x, p_2)F(x, p_2). \tag{2.3}$$

Here we claim that $(I - G)$ gives us the 1-1 map near p_2. So we can define the $(I - G)(y - p_2) = y' - p_2$ and obtain the map from x to y' near the point p. Therefore, we prove this theorem. $\square$

To show $(I-G)$ is 1-1, it is actually a implication of a fix point theorem. I will prove this in a moment. However, we can easily be convinced by this claim, because I is 1-1 and G is a such small error term.

To prove that $I-G$ is one to one, we can see this from the inequality

$$|(I-G)(y_1-p_2)-(I-G)(y_2-p_2)| \geq |y_1-y)2|-|G||y_1-y_2| \geq |y_1-y_2|. \tag{2.4}$$

To prove $I-G$ is onto, we take any q near p_2 fixed. Then we want to find y such that

$$G(y)+q=y. \tag{2.5}$$

Let us consider $B_\delta(p_2)$ with $\delta > 2dist(q,p_2)$. Then $G(y)+q \in B_\delta(p_2)$ for any $y \in B_\delta(p_2)$. Therefore, if we start with any $y_0 \in B_\delta(p_2)$ and consider the sequence

$$y_i = G(y_{i-1})+q. \tag{2.6}$$

we can see that $|y_i - y_{i+1}| \leq \frac{1}{100}|y_{i-1}-y_i|$. So $\{y_i\}$ converges to some y with $y = G(y)+q$. This proves (1.11)