Math3033 (Fall 2013-2014) **Tutorial Note 3**

Inverse and Implicit Function Theorem

Key Definitions and Results

Definition 1. We say that $F:\mathbb{R}^n \to \mathbb{R}^m$ is $\boldsymbol{C^1}$ **at** $\boldsymbol{a \in \mathbb{R}^n}$ if:

(i) Near a all first order partial derivatives exists.

(ii) All partial derivatives of F are continuous at a.

Definition 2. Let $F:U \to \mathbb{R}^m$ be defined on some $U \subseteq \mathbb{R}^n$, we say that F is $\boldsymbol{C^1}$ **on a subset** $\boldsymbol{E \subseteq U}$ if F is C^1 at each $a \in E$.

Definition 3. Let $x = (x_{i_1},\dots,x_{i_h})$, $F:\mathbb{R}^n \to \mathbb{R}^m$ and $p \in \mathbb{R}^n$. If $F'(p)$ exists, we define

$$F_{x_{i_1},\dots,x_{i_h}}(p) := F_x(p) := \begin{bmatrix} \dfrac{\partial f_1}{\partial x_{i_1}} & \cdots & \dfrac{\partial f_1}{\partial x_{i_h}} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial x_{i_1}} & \cdots & \dfrac{\partial f_m}{\partial x_{i_h}} \end{bmatrix}(p).$$

Theorem 4 (Inverse Function). Let $F:\mathbb{R}^n \to \mathbb{R}^n$ be a function such that:

(i) F is C^1 near $p \in \mathbb{R}^n$.

(ii) $\det(F'(p)) \neq 0$.

Then there is an $r>0$ such that $F|_{B(p,r)}$ has a C^1 inverse $G: F(B(p,r)) \to \mathbb{R}^n$, and by the chain rule we have

$$\boxed{G'(F(x)) = (F'(x))^{-1}} \quad \text{for all } x \in B(p,r).$$

Theorem 5 (Implicit Function). Let $F:\mathbb{R}^n \to \mathbb{R}^m$ $(n>m)$ be a function such that:

(i) F is C^1 near $p = (p_1,\dots,p_n)$.

(ii) $F(p) = 0$.

(iii) The right block matrix indicated below is invertible:

$$\left[\begin{array}{ccc|ccc} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_{n-m}} & \dfrac{\partial f_1}{\partial x_{n-m+1}} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial x_1} & \cdots & \dfrac{\partial f_m}{\partial x_{n-m}} & \dfrac{\partial f_m}{\partial x_{n-m+1}} & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{array}\right](p)$$

- Then near p the variables $(x_{n-m+1},\dots,x_n)$ can be ***solved implicitly*** from the equation $F(x_1,\dots,x_n) = 0$ and expressed as a C^1 function $G(x_1,\dots,x_{n-m})$.
- More precisely, there is an open ball $B(p,r) \subseteq \mathbb{R}^n$ and a C^1 function G defined on the "open set"

$$U := \{(x_1,\dots,x_{n-m}) : (x_1,\dots,x_n) \in B(p,r)\}$$

such that

$$\begin{aligned} &B(p,r) \cap F^{-1}(\{0\}) \\ &= \text{image of } U \text{ under } (\mathrm{id}, G) \\ &= \left\{(x_1,\dots,x_{n-m}, G(x_1,\dots,x_{n-m})) : (x_1,\dots,x_{n-m}) \in U\right\}. \end{aligned} \qquad (*)$$

- Furthermore, by Example 1 of this tutorial note we have

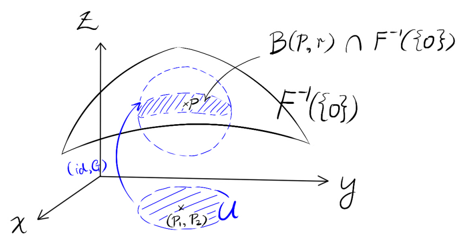
$$\boxed{G'(p_1,\dots,p_{n-m}) = -\left(\left[\begin{array}{c}\textbf{invertible} \\ \textbf{part of } \boldsymbol{F'(p)}\end{array}\right]\right)^{-1} \times \left[\begin{array}{c}\textbf{remaining} \\ \textbf{part of } \boldsymbol{F'(p)}\end{array}\right].} \qquad (\square)$$

Example 1. Consider Implicit Function Theorem 5, explain the geometrical meaning of the set equality $(*)$ when $n=3$ and $m=1$.

Solution. $(*)$ says that the surface

$$F^{-1}(\{0\}) = \{(x,y,z) \in \mathbb{R}^3 : F(x,y,z) = 0\}$$

near p can be parametrized as the graph of a C^1 function G defined near (p_1,p_2).



Thus generally Implicit Function Theorem is nothing but parametrization of "abstract surfaces" near some point, in such general cases G will be vector-valued and cannot be easily visualized, unless we are god :).

Example 2. Let

$$u = x + y,$$
$$v = \sin x + \cos y.$$

Show that when (x, y) is near $(0,1)$, x and y can be expressed as a differentiable function in (u,v). Compute x_u, x_v, y_u, y_v at $(1,\cos 1)$.

Solution. Define

$$(u,v) = F(x, y).$$

To express x, y as a differentiable function of (u,v) is to find differentiable F^{-1} such that $(x, y) = F^{-1}(u,v)$. For this, we use Inverse Function Theorem.

Step 1: We need to show that F is C^1 near $(0,1)$.

Procedure: The standard way to do this is to find an $r > 0$ such that all partial derivatives of F **exist** and are **continuous on $B((0,1),r)$**. In many cases it is easy to choose suitable r.

Now

$$u_x = 1, \quad u_y = 1$$

and

$$v_x = \cos x, \quad v_y = -\sin y.$$

As they exist and are continuous on $\mathbb{R}^2$, they exist and are continuous on $B((0,1),1)$, so F is C^1 near $(0,1)$.

Step 2: We need to show that $F'(0,1)$ is invertible. This is a routine calculation:

$$\det F'(0,1) = \begin{bmatrix} 1 & 1 \\ 1 & -\sin 1 \end{bmatrix} = -\sin 1 - 1 \neq 0.$$

Step 3: Make conclusion (=.=). By Inverse Function Theorem near $(0,1)$ F has a C^1 inverse F^{-1}, thus

$$(x, y) = (x(u,v), y(u,v)) = F^{-1}(u,v) : \text{near } F(0,1) \to \text{near } (0,1)$$

is now a differentiable function of (u,v). Moreover,

$$\begin{aligned} \begin{bmatrix} x_u & x_v \\ y_u & y_v \end{bmatrix}(1,\cos 1) &= (F^{-1})'(F(0,1)) \\ &= (F'(0,1))^{-1} \\ &= \begin{bmatrix} 1 & 1 \\ 1 & -\sin 1 \end{bmatrix}^{-1} \\ &= \frac{1}{1+\sin 1}\begin{bmatrix} \sin 1 & 1 \\ 1 & -1 \end{bmatrix}. \end{aligned}$$

Example 3. Consider the following equations

$$x^2 + 2y^2 + u^2 + v = 6, \quad (1)$$
$$2x^3 + 4y^2 + u + v^2 = 9. \quad (2)$$

(a) Show that near $p = (1,-1,-1,2)$, (u,v) can be expressed as a differentiable function of (x, y).

(b) Compute u_x and v_x at $(1,-1)$.

Solution. (a) We define

$$F(x, y,u,v) = (f, g) = \begin{bmatrix} x^2 + 2y + 2 + u^2 + v - 6 \\ 2x^3 + 4y^2 + u + v^2 - 9 \end{bmatrix}.$$

To "solve" (u,v) out from the equations $F = 0$ for (x, y,u,v) near p, we try to prove $F_{u,v}(p)$ is an invertible matrix plus some extra conditions.

Step 1: Show that F is C^1 near p. It is similar to Example 2, indeed,

$$F'(x, y,u,v) = \begin{bmatrix} f_x & f_y & f_u & f_v \\ g_x & g_y & g_u & g_v \end{bmatrix} = \begin{bmatrix} 2x & 4y & 2u & 1 \\ 6x^2 & 8y & 1 & 2v \end{bmatrix},$$

clearly all partial derivatives exist and are continuous on $B(p,1)$, so F is C^1 near p.

Step 2: Show that $F(p) = 0$. This can be done by direct calculation, make sure that you finish this step in your midterm/final exams.

Step 3: Show that $F_{u,v}(p)$ is invertible. We have

$$F'(p) = F'(1,-1,-1,2) = \begin{bmatrix} 2 & -4 & -2 & 1 \\ 6 & -8 & 1 & 4 \end{bmatrix}.$$

Therefore $F_{u,v}(p) = \begin{bmatrix} -2 & 1 \\ 1 & 4 \end{bmatrix}$ with $\det F_{u,v}(p) = -9$, thus $F_{u,v}(p)$ is invertible. By Implicit Fucntion Theorem u,v can be expressed as a function $G(x, y)$ defined near $(1,-1)$ (such that $F(x, y,G(x, y)) = 0$ for (x, y) near $(1,-1)$).

(b) **Method 1.** We use the equation on the first page of tutorial note 3 (or the exact version in Example 4) directly to get

$$G'(1,-1) = -\begin{bmatrix} -2 & 1 \\ 1 & 4 \end{bmatrix}^{-1}\begin{bmatrix} 2 & -4 \\ 6 & -8 \end{bmatrix} = \begin{bmatrix} 2/9 & -8/9 \\ -14/9 & 12/9 \end{bmatrix}.$$

Method 2. We can try to create a system of linear equations and then solve for u_x and v_x. Indeed, we differentiate (1) and (2) w.r.t. x to obtain

$$2x + 2uu_x + v_x = 0 \quad \text{and} \quad 6x^2 + u_x + 2vv_x = 0.$$

Now we put $(x, y) = (1,-1)$ and use the fact that $u(1,-1) = -1$ and $v(1,-1) = 2$, then

$$2u_x - v_2 = 2 \quad \text{and} \quad u_x + 4v_x = -6,$$

we solve them to get $u_x(1,-1) = 2/9$ and $v_x(1,-1) = -14/9$.

***Example 1.** In this example we consider $n = 4$ and $m = 2$ of Implicit Function Theorem, *for simplicity*. Let $F : \mathbb{R}^4 \to \mathbb{R}^2$ and $p_0 = (x_0, y_0, u_0, v_0) \in \mathbb{R}^4$. Suppose:

(i) F is C^1 near p_0.

(ii) $F(p_0) = 0$.

(iii) $F_{u,v}(p_0)$ is invertible.

Show that near p_0, u, v can be solved from the equation $F(x,y,u,v) = 0$ and expressed as a C^1 function $G(x,y)$ near (x_0,y_0). Not only that, we have for (x,y) near (x_0,y_0),

$$\begin{bmatrix} u_x & u_y \\ v_x & v_y \end{bmatrix}(x,y) = G'(x,y) = -[F_{u,v}\big(x,y,G(x,y)\big)]^{-1} F_{x,y}\big(x,y,G(x,y)\big). \quad (\square')$$

Solution. (i), (ii) and (iii) are just the conditions in Implicit Function Theorem, thus near p_0, (u,v) can be expressed as a C^1 function $G(x,y)$ such that $F(x,y,G(x,y)) = 0$ for (x,y) near (x_0,y_0). Now we differentiate both sides (taking Jacobian matrix) to get

$$0 = F'(x,y,G(x,y)) \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \multicolumn{2}{c}{G'(x,y)} \end{bmatrix} = F_{x,y}(x,y,G(x,y)) + F_{u,v}(x,y,G(x,y))G'(x,y),$$

here last equality follows from block matrix multiplication, note that $G'(x,y)$ is a 2×2 matrix, now we solve from the above equation to get

$$G'(x,y) = -F_{u,v}(x,y,G(x,y))^{-1} F_{x,y}(x,y,G(x,y)).$$

When $(x,y) = (x_0,y_0)$, then

$$\boxed{G'(x_0,y_0) = -F_{u,v}(p_0)^{-1} F_{x,y}(p_0),}$$

this is the version we use in Example 3.

Exercise 1 (Application of Im.F.T.). Let $f(x_1,\dots,x_k)$ be a homogeneous polynomial of degree $d \ge 1^{(*)}$, i.e.,

$$f(tx_1,\dots tx_k) = t^d f(x_1,\dots,x_k) \quad \text{for every } t \in \mathbb{R}.$$

Show that if $\boldsymbol{c \neq 0}$, then the surface $f^{-1}(\{c\}) = \{x \in \mathbb{R}^k : f(x_1,\dots,x_k) = c\}$ can be locally parametrized as the graph of a ***smooth*** (i.e., has partial derivatives of any order) function. *You may try to plot out* $x^2+y^2-z^2 = 1$ *and* $x^2+y^2-z^2 = 0$ *on* WolframAlpha *to see their difference.*

$^{(*)}$ For example, $x^3 + x^2y + y^3$ and $x^4 - xy^3$ are homogeneous polynomial of degree 3 and 4 respectively.

***Example 2 (HKU, Analysis II, Spring 2009).** Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable. Suppose that there is a number $c > 0$ such that

$$\|F(x) - F(y)\| \ge c\|x-y\|, \quad \text{for all } x,y \in \mathbb{R}^n.$$

Show that F is one-one and $DF(x) \neq 0$ for all $x \in \mathbb{R}^n$.

Solution. We note that the map from $\mathbb{R}^n$ to $\mathbb{R}$ defined by

$$x \mapsto \|F'(x)\|$$

is continuous since F is continuously differentiable (i.e., C^1). Therefore if $F'(x_0) = 0$, for some x_0, then for every $\epsilon > 0$ there is a $\delta > 0$ such that

$$\|x - x_0\| < \delta \implies \|F'(x)\| < \epsilon.$$

Therefore by Exercise 4 of tutorial note 2 we have for every $x,y \in B(x_0,\delta)$,

$$\|F(x) - F(y)\| \le \epsilon\|x-y\|.$$

By hypothesis of this example we have

$$c\|x-x_0\| \le \epsilon\|x-x_0\|,$$

therefore if we take $x \in B(x_0,\delta)\setminus\{x_0\}$, then

$$c \le \epsilon.$$

But $\epsilon > 0$ is arbitrary, we have $0 < c \le 0$ by taking $\epsilon \to 0^+$, a contradiction.

The following exercise is for those who knows:

1) The definition of openness, closedness, connectedness of subsets in $\mathbb{R}^n$.

2) The "local version" of inverse function theorem instead of the weaker one stated in this course. Namely, we need the following version:

Theorem. *If $F = (f_1,\dots,f_n)$ is <u>defined near $a \in \mathbb{R}^n$</u>, <u>C^1 near a</u> and <u>$\det f'(a) \neq 0$</u>, then f is a local C^1-diffeomorphism at a.*

Exercise 2 (Application of In.F.T.). Let $F : \mathbb{R}^2 \to \mathbb{R}^2$ be C^1 on $\mathbb{R}^2$. Suppose that:

- $\det F'(x) = 0$ for at most finitely many $x \in \mathbb{R}^2$.
- For every $M > 0$ the set $\{x \in \mathbb{R}^2 : |F(x)| \le M\}$ is bounded$^{(\dagger)}$;

Prove that F maps $\mathbb{R}^2$ **onto** $\mathbb{R}^2$.

Hint: Prove by contradiction.

$^{(\dagger)}$ Such a continuous map is said to be **proper**.