MATH 3472

§13.2 Functions with $J_f \neq 0$.

We already saw that $J_f(\vec{x}) \neq 0$ for functions $f: \mathbb{R}^n \longrightarrow \mathbb{R}^n$ means that (subject to a few conditions) every ball B has an image $f(B)$ with nonempty interior. Moreover, if B is centred at a point $\vec{a} \in \mathbb{R}^n$, then $f(a)$ is the point that serves as a witness to the fact that $\operatorname{int}(f(B))$ is nonempty.

Our goal is to show that $J_f \neq 0$ implies f is one-to-one and <u>open</u>. Recall a map $f: X \longrightarrow Y$ is open if, for every open set $U \subset X$, the set $f(U)$ is open. The next theorem shows that if we can show f is one-to-one, then that is enough.

<u>Theorem 13.3</u>: Suppose $U \subseteq \mathbb{R}^n$ is open, and that $f: U \longrightarrow \mathbb{R}^n$ is continuous with finite partials $D_j f_i$ on U. If

(i) f is one-to-one

(ii) $J_f(\bar{x}) \neq 0 \quad \forall\, \bar{x} \in U$

Then $f(U)$ is open.

<u>Proof</u>: Let $b \in f(U)$ be given, and suppose $b = f(a)$ for some $a \in U$.

Because U is open, $\exists$ a ball $B \subset U$, with a at its centre, and by our assumptions this B will satisfy the hypotheses of the previous theorem. I.e:

- f is cts on $\overline{B}$
- $D_j f_i(\vec{x})$ exist for all $\vec{x} \in B$
- $J_f(\vec{x}) \neq 0 \;\; \forall \vec{x} \in B$
- $f(\vec{a}) \neq f(\vec{x}) \;\; \forall \vec{x} \in \partial B$, since f is one-to-one.

Therefore there exists an open ball B' containing $f(a)$ with $f(a) \in B' \subset f(B) \subset f(U)$. So
$b = f(a) \in \mathrm{int}(f(U))$, so $f(U)$ is open.

Thus our goal becomes: Show that $J_f(\vec{x}) \neq 0$ implies that f is one-to-one.

Theorem 13.4: Assume that $f = (f_1, \ldots, f_n)$ has continuous partials $D_j f_i$ on some open set $U \subseteq \mathbb{R}^n$. Suppose further that $J_f(\vec{a}) \neq 0$ for some point $\vec{a} \in U$. Then there exists a ball B with centre $\vec{a}$ such that f is one-to-one on B.

Proof: Choose n points $z_1, \ldots, z_n$ in U.
Use the notation

$$Z = (z_1, z_2, z_3, \ldots, z_n)$$

to denote the obvious element of $\mathbb{R}^{n^2}$.

Define a function $h: \mathbb{R}^{n^2} \longrightarrow \mathbb{R}$ as follows:

$$h(z) = \det [D_j f_i(z_i)]$$

This function is only defined at certain points in $\mathbb{R}^{n^2}$, namely those points of the form

$$z = (z_1, z_2, z_3, \ldots, z_n)$$

where $z_i \in U$ for all i. At points of this form, it is continuous, however, for the following reason:

By assumption, each $D_j f_i$ is continuous on all of U. The determinant function is a polynomial in the n^2 entries of $[D_j f_i(z_i)]$, so overall, h is a polynomial function composed with n^2 continuous partials - and is thus continuous.

Now let z denote the special point in the domain of h given by:

$$z = \underbrace{(\vec{a}, \vec{a}, \ldots, \vec{a})}_{n \text{ times}} \quad \left(\begin{array}{l} \vec{a} \in U \text{ is where} \\ J_f(\vec{a}) \neq 0 \end{array} \right).$$

Then $h(z) = \det [D_j f_i(\vec{a})] = J_f(\vec{a}) \neq 0$, and so by continuity of h there is some ball $B' \subseteq \mathbb{R}^{n^2}$ containing z such that h is nonzero on B'.

Correspondingly, there is a ball $B \subseteq \mathbb{R}^n$ such that $\det [D_j f_i(z_i)] \neq 0$ if each $z_i \in B$ (with B centred at $\vec{a}$).

To see this, observe that if $\vec{a} = (a_1, \dots, a_n)$ then the n^2-ball B' contains a set of the form:

$$(a_1 - \varepsilon_{1,1}, a_1 + \varepsilon_{1,1}) \times \dots \times (a_n - \varepsilon_{n,1}, a_n + \varepsilon_{n,1})$$

$$\times (a_1 - \varepsilon_{1,2}, a_1 + \varepsilon_{1,2}) \times \dots \times (a_n - \varepsilon_{n,2}, a_n + \varepsilon_{n,2}) \times \dots$$

$$\dots \times (a_1 - \varepsilon_{1,n}, a_1 + \varepsilon_{1,n}) \times \dots \times (a_n - \varepsilon_{n,n}, a_n + \varepsilon_{n,n}),$$

and then choosing B to be a ball of radius $r = \min_{i,j} \{\varepsilon_{i,j}\}$ suffices.

<u>Claim</u>: f is one-to-one on B.

Assume not, and choose $\overset{x \neq y}{\vec{x}, \vec{y}}$ in B with $f(\vec{x}) = f(\vec{y})$. Since $L(\vec{x}, \vec{y}) \subseteq B$, we can apply the MVT to each component function f_i of f to write:

$$0 = f_i(y) - f_i(x) = \nabla f_i(z_i) \cdot (\vec{y} - \vec{x}) \quad i = 1, 2, \dots, n.$$

for some $z_i \in L(\vec{x}, \vec{y}) \subseteq B$. (Here the book remarks: the MVT applies because f is differentiable on S).

Now interpreting our ~~vector~~ equations above as a system of equations, we have

$$\sum_{k=1}^{n} (y_k - x_k) a_{ik} = 0, \quad \text{where } a_{ik} = D_k f_i(z_i).$$

But the determinant of this coefficient matrix is nonzero, since it's equal to $h(z_1, z_2, .., z_n)$ where $z_i \in B$ for all i (so $(z_1, .., z_n) \in B'$).

Thus $y_k = x_k \; \forall k$, so that f is one-to-one on B.

Caution: This theorem gives the result locally.

That is, for such an $\vec{a}$ there exists a ball B (depending on $\vec{a}$) where f is one-to-one. A nearby point $\vec{a}'$ could have associated with it a different ball B', and on the union $B \cup B'$ f may not be one-to-one.

E.g. if $f(r, \theta) = (r\cos\theta, r\sin\theta)$ then

$$J_f(r,\theta) \overset{\det}{=} \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix} = r(\cos^2\theta + \sin^2\theta) = r,$$

so the Jacobian is nonzero on the open set $\mathbb{R}^2 \setminus \{(0,0)\}$. However f is not one-to-one there, since $(r\cos\theta, r\sin\theta) = (r\cos\theta', r\sin\theta')$ whenever $\theta = \theta' + 2k\pi$.

Globally, what we have is:

We just finished:

Theorem 13.5: Let $U \subset \mathbb{R}^n$ be open, and suppose that $f : U \to \mathbb{R}^n$ is such that $D_j f_i$ are continuous on U $\forall i, j$. If $J_f(\vec{x}) \neq 0$ for all $\vec{x} \in U$, then f is an open map.

Proof: Let $S \subset U$ be open, and choose $x \in S$. Then there is a ball B_x centred at $\vec{x}$ on which f is one-to-one, by Theorem 13.4. By Theorem 13.3, $f(B_x)$ is open in $\mathbb{R}^n$. Now write $U = \bigcup_{x \in U} B_x$, then $f(U) = f\left(\bigcup_{x \in U} B_x\right)$

$$= \bigcup_{x \in U} f(B_x),$$

which is open since it is a union of open sets.

We've already seen that if $J_f(\vec{a}) \neq 0$, then there's a ball around $\vec{a}$ where f is one-to-one. Thus, f has a "local inverse", ie. because f is one-to-one on B there's a function $f^{-1} : f(B) \to B$ that is defined by $f^{-1}(y) = x \iff f(x) = y$.

A priori the function f^{-1} may be discontinuous and badly behaved, however we will see that is not the case.

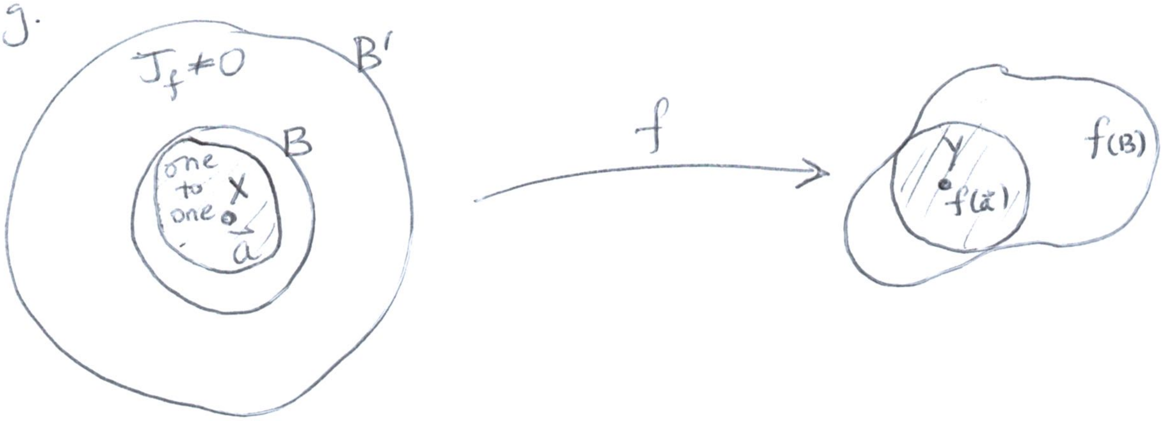
Theorem (Inverse function theorem)

Let $U \subseteq \mathbb{R}^n$ be open and suppose $f: U \to \mathbb{R}^n$, $f = (f_1, \dots, f_n)$, is such that $D_r f_k(\vec{x})$ exists and is cts for all r, k and $\forall \vec{x} \in U$. Suppose $\exists\, a \in U$ s.t. $J_f(\vec{a}) \neq 0$. Then there exist open sets X and $Y = f(X)$ with $a \in X$, such that f is one-to-one on X and its inverse $g: Y \to X$ satisfies $D_r g_k(\vec{y})$ exists and is cts for all $\vec{y} \in Y$ and $\forall r, k$.

Proof: Here is how we arrive at the sets X, Y and the function g.

The function $J_f: U \to \mathbb{R}$ is continuous (by continuity of $D_r f_k$) and thus there is an n-ball B' with $\vec{a}$ as its centre such that $J_f(\vec{x}) \neq 0$ on B'.
By Theorem B.4, we may choose a smaller ball $B \subset B'$ centred at $\vec{a}$ where f is also one-to-one. Last, by Theorem B.2 $f(B)$ contains an n-ball Y with centre $f(\vec{a})$, set $X = f^{-1}(Y) \cap B$, note X is open.

E.g.



Note: $f^{-1}(Y)$ could contain things outside of B', since f is not globally one-to-one!

Now the set $\overline{B}$ is compact and f is one-to-one there, and continuous. By the theorem below, there exists an inverse $g: f(\overline{B}) \longrightarrow \overline{B}$ that is continuous on $f(\overline{B})$. Since ~~X~~ $X \subset \overline{B}$ and $Y \subseteq f(\overline{B})$, all that remains to show is that $D_i g_k$ exist and are continuous on Y.

First we show the theorem used above

Theorem: A continuous, one-to-one function with compact domain has continuous inverse. (defined on its image).

Proof: Let $f: X \longrightarrow Y$ be continuous, one-to-one and onto, suppose that X is compact (here, X, Y are subsets of Euclidean space). Let $g: Y \longrightarrow X$ denote the inverse function, and let $U \subset X$ be open. We show $g^{-1}(U)$ is open.

To see this, observe that U^c is closed, thus $U^c \cap X$ being a closed subset of X, is compact. Since f is continuous, $f(U^c \cap X)$ is compact and so closed in Y. Since f is one-to-one and onto, $f(U^c \cap X)^c = f(U)$, and thus $f(U) = g^{-1}(U)$ is open in Y.

Note: We only needed to know that a closed subset of a compact set X is compact, and that a compact subset of a closed set Y is closed. These abstraction notions form the foundation of topology. I suggest studying it further.

MATH 3472

continuing the proof of the inverse function theorem:

We've arrive at (by using $J_f \neq 0$ and continuity) sets X, Y with $f: X \longrightarrow Y$ and $g: Y \longrightarrow X$, f and g both continuous inverses of one another. Moreover $D_j f_i$ are continuous on X, and $\det[D_j f_i(z_i)] \neq 0$ $\forall z_1, \ldots, z_n$ satisfying $z_i \in X$ $\forall i$.

Next we show: g has continuous partials on Y. Suppose $g = (g_1, \ldots, g_n)$. Let $\vec{y} \in Y$ and let $\vec{u}_r$ denote the r^{th} coordinate vector, and consider the quotient

$$\frac{g_k(\vec{y} + t\vec{u}_r) - g_k(\vec{y})}{t}.$$

For sufficiently small t, $g_k(\vec{y} + t\vec{u}_r)$ is defined since $\vec{y} + t\vec{u}_r \in Y$ as ~~it~~ is open.

Set $\vec{x} = g(\vec{y})$ and $\vec{x}' = g(\vec{y} + t\vec{u}_r)$, with t chosen sufficiently small that $L(\vec{x}, \vec{x}') \subseteq X$.

Apply the MVT to f, and we arrive at (for each coordinate function $f_i(\vec{x})$):

$$\frac{f_i(\vec{x}') - f_i(\vec{x})}{t} = \nabla f_i(z_i) \cdot \frac{(x' - x)}{t} \quad \text{for } i = 1, \ldots, n.$$

Note that since $f(\vec{x}') - f(\vec{x}) = \vec{y} + t\vec{u}_r - \vec{y} = t\vec{u}_r$, the quantity $f_i(\vec{x}') - f_i(\vec{x})$ is zero unless $i = r$, in which case it is equal to t.

Thus $\nabla f_i(\vec{z}_i) \cdot \left(\dfrac{\vec{x}' - \vec{x}}{t} \right) = \begin{cases} 0 & \text{if } i \neq r \\ 1 & \text{if } i = r \end{cases}$.

This is a system of n linear equations in n unknowns:

$$\begin{bmatrix} D_1 f_1(\vec{z}_1) & D_2 f_1(\vec{z}_1) & \cdots \\ D_1 f_2(\vec{z}_2) & D_2 f_2(\vec{z}_2) & \cdots \\ \vdots & \vdots & \\ & & \end{bmatrix} \begin{bmatrix} \frac{x_1' - x_1}{t} \\ \vdots \\ \vdots \\ \frac{x_n' - x_n}{t} \end{bmatrix} = \begin{array}{c} \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \end{array} \begin{array}{l} \\ \\ \leftarrow r^{th} \text{ position.} \\ \\ \\ \end{array}$$

Since $\det[D_j f_i(\vec{z}_i)] \neq 0$, there's a unique solution. We can solve for the k^{th} unknown, $\dfrac{x_k' - x_k}{t} = \dfrac{g_k(\vec{y} + t\vec{u}_r) - g_k(\vec{y})}{t}$

by using Cramer's rule and get

$$\frac{g_k(\vec{y} + t\vec{u}_r) - g_k(\vec{y})}{t} = \frac{\det J_k}{\det J},$$ where J is the Jacobian determinant $\det[D_j f_i(z_i)]$ and J_k is the determinant of the matrix obtained by replacing the k^{th} column with $\begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \begin{array}{l} \\ \\ \leftarrow r^{th} \text{ position.} \\ \\ \\ \end{array}$

Now considering the limit of this expression as $t \to 0$:

Since g is continuous, $\lim_{t\to 0} g(\vec{y}+t\vec{u}_r) = g(\vec{y})$, and therefore $\lim_{t\to 0}\vec{x}' = \vec{x}$. But as $\vec{z}_i$ is on the segment $L(\vec{x},\vec{x}')$ for all i, this means $\lim_{t\to 0}\vec{z}_i = \vec{x}$. Therefore

$$\lim_{t\to 0} \det[D_j f_i(z_i)] = \cancel{\lim} \det[D_j f_i(\vec{x})] = J_f(\vec{x}),$$

which is nonzero since $\vec{x}\in X$. Thus

$$\lim_{t\to 0} \frac{g_k(\vec{y}+t\vec{u}_r)-g_k(\vec{y})}{t} = \lim_{t\to 0}\frac{\det J_k}{\det J} = \frac{\lim \det J_k}{\lim \det J}$$

this is valid since the bottom is not going to zero

and the limit $\lim_{t\to 0}\det J_k$ clearly exists since it is the limit of a continuous function (det composed with partials $D_j f_i$, and one column of 0's and 1's).

Therefore $\lim_{t\to 0} \frac{g_k(\vec{y}+t\vec{u}_r)-g_k(\vec{y})}{t}$ exists, ie $D_r g_k(\vec{y})$ exists for all $\vec{y}\in Y$.

Moreover, continuity of $D_r g_k$ follows from continuity of $D_j f_i$ appearing in the determinants of J_k and J:

Since $\lim_{t\to 0}\det[D_j f_i(\vec{x})] = J_f(\vec{x})$ and $\lim_{t\to 0}\det J_k$

$= \det A_k(\vec{x})$, where $A_k(\vec{x})$ is the matrix $[D_j f_i(\vec{x})]$ with $\begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$ ← r^{th} position appearing as its k^{th} column, we can write (recall $\vec{x}$ is defined as $g(\vec{y})$)

$$D_r g_k(y) = \frac{\det A_k(g(\vec{y}))}{J_f(g(\vec{y}))}$$

which is evidently a quotient of continuous functions whose denominator is never zero. Thus $D_r g_k(y)$ is continuous.

Note: The proof above also gives a method for computing $D_r g_k(\vec{y})$, but it is not the best such method. An easier approach is as follows:

If f and g are inverses, then $f \circ g = id$ and therefore

$$Df(g(\vec{y})) \cdot Dg(\vec{y}) = I \quad \text{(identity matrix)}$$

and

$$Df(\vec{x}) \cdot Dg(f(\vec{x})) = I$$

(depending on how you apply the chain rule).

Taking the first equation and writing $\vec{x}$ for $g(\vec{y})$, we get n^2 equations

$$\sum_{k=1}^{n} D_k g_i(\vec{y}) D_j f_k(\vec{x}) = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i \neq j. \end{cases}$$

For a fixed i, there are only n equations, as $j=1,\ldots,n$. Solving these n linear equations yields unknowns $D_1 g_i(\vec{y}), \dots, D_n g_i(\vec{y})$. The recommended method for solving for an individual partial derivative would be (again) Cramer's rule.

<u>Example</u>: Consider $f: \mathbb{R}^2 \longrightarrow \mathbb{R}$ given by

$$f(r, \theta) = (r\cos\theta, r\sin\theta). \quad \text{(Think polar)}.$$

Then
$$Df = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} \Rightarrow J_f(r,\theta) = r.$$

Fix $\vec{x} = (\sqrt{2}, \pi/4)$. Then $Df(\vec{x}) = \begin{pmatrix} \frac{1}{\sqrt{2}} & -1 \\ \frac{1}{\sqrt{2}} & 1 \end{pmatrix}$,

and at the point $f(\sqrt{2}, \pi/4) = (1,1)$ the total derivative of the inverse g must be

$$Dg(\vec{y}) = (Df(\vec{x}))^{-1} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/2 & 1/2 \end{pmatrix},$$

so we can read off the values of the derivatives $D_r g_k(1,1)$ for various r, k. E.g. $D_1 g_1(1,1) = \frac{1}{\sqrt{2}}$.

But we know a formula for g in this case!

$$g(x,y) = \left(\sqrt{x^2+y^2}, \arctan\left(\tfrac{y}{x}\right)\right) \text{ and so}$$

$$D_1 g_1(x,y) = \frac{1}{2} \cdot \frac{1}{\sqrt{x^2+y^2}} \cdot 2x \Rightarrow D_1 g_1(1,1) = \frac{1}{\sqrt{2}},$$

it works. (Try other entries of $Dg(1,1)$ to convince yourself).