§3.3 The Inverse Function Theorem

Recall: Chain Rule

Let
$$\left.\begin{array}{l} G : U^{\subset \mathbb{R}^n} \longrightarrow \mathbb{R}^m \\ F : V^{\subset \mathbb{R}^m} \longrightarrow \mathbb{R}^\ell, \end{array}\right\} \text{differentiable}$$

U, V open in $\mathbb{R}^n$ & $\mathbb{R}^m$ respectively, and

$$G(U) \subset V.$$

Then $H = F \circ G : U \to \mathbb{R}^\ell$ differentiable and

$$DH(x) = DF(G(x))\, DG(x),$$

where
$$DG(x) = \left(\frac{\partial G_i}{\partial x_j}(x)\right)_{\substack{i=1,\cdots,m \\ j=1,\cdots,n}} = \begin{pmatrix} - \nabla G_1 - \\ \vdots \\ - \nabla G_m - \end{pmatrix}$$

$$= \begin{pmatrix} \frac{\partial G_1}{\partial x_1} & \cdots & \frac{\partial G_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial G_m}{\partial x_1} & \cdots & \frac{\partial G_m}{\partial x_n} \end{pmatrix}$$

and similarly for DF & DH.

We also need

Prop 3.6 Let $F: B \to \mathbb{R}^n$ be $\underline{C^1}$, where B = ball in $\mathbb{R}^n$.

Then $\forall\, \vec{x}_1, \vec{x}_2 \in B$,

matrix acts on column vector.

$$F(\vec{x}_1) - F(\vec{x}_2) = \left(\int_0^1 DF(\vec{x}_2 + t(\vec{x}_1 - \vec{x}_2))\, dt \right) \cdot (\vec{x}_1 - \vec{x}_2)$$

In component form $F = \begin{pmatrix} F_1 \\ \vdots \\ F_n \end{pmatrix}$, this is

j-component of the vector $\vec{X}_1 - \vec{X}_2$

$$F_i(\vec{x}_1) - F_i(\vec{x}_2) = \sum_{j=1}^n \left(\int_0^1 \frac{\partial F_i}{\partial x_j}(\vec{x}_2 + t(\vec{x}_1 - \vec{x}_2))\, dt \right) (\vec{x}_1 - \vec{x}_2)_j .$$

jth-variable of F_i

Pf: For each $i = 1, \cdots, n$,

$$F_i(\vec{x}_1) - F_i(\vec{x}_2) = \int_0^1 \left(\frac{d}{dt} F_i(\vec{x}_2 + t(\vec{x}_1 - \vec{x}_2)) \right) dt$$

$$= \int_0^1 \sum_{j=1}^n \left[\frac{\partial F_i}{\partial x_j}(\vec{x}_2 + t(\vec{x}_1 - \vec{x}_2)) \cdot (\vec{x}_1 - \vec{x}_2)_j \right] dt$$

$$= \int_0^1 \nabla F_i(\vec{x}_2 + t(\vec{x}_1 - \vec{x}_2)) \bullet (\vec{x}_1 - \vec{x}_2)\, dt$$

$$= \left(\int_0^1 \nabla F_i(\vec{x}_2 + t(\vec{x}_1 - \vec{x}_2))\, dt \right) \bullet (\vec{x}_1 - \vec{x}_2)$$

dot product of vectors

$$\therefore \quad F(\vec{x}_1) - F(\vec{x}_2) = \left(\int_0^1 DF(\vec{x}_2 + t(\vec{x}_1 - \vec{x}_2))\, dt \right) \cdot (\vec{x}_1 - \vec{x}_2)$$

#

Recall: If $F: U \subset \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at a point p in an open set U of $\mathbb{R}^n$, then

(Caution: we write "x" for "$\vec{x}$" in order to simplify notations)

$$F(p+x) - F(p) = DF(p)\,x + o(|x|)$$

$\forall\, x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ sufficiently small, (i.e. $|x|$ small)

where $o(|x|)$ is a remaining term such that

$$\frac{o(|x|)}{|x|} \to 0 \quad \text{as} \quad |x| \to 0.$$

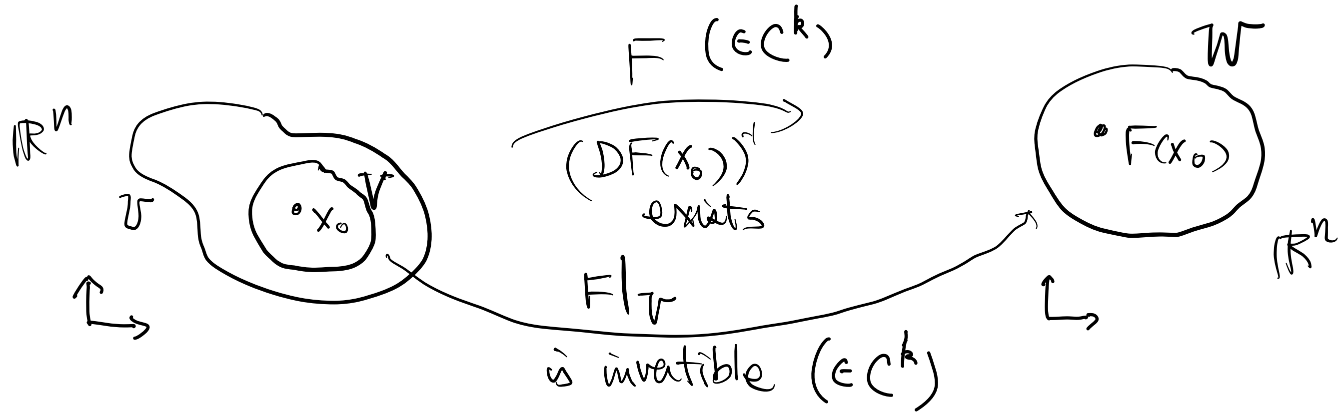
Thm 3.7 (<u>Inverse Function Theorem</u>)

Let $F: U \to \mathbb{R}^n$ be a <u>C^1-map</u> from an open set $U \subset \mathbb{R}^n$.

Suppose $x_0 \in U$ and $DF(x_0)$ is <u>invertible</u> (as a matrix or linear transformation).

(a) Then $\exists$ open sets V & W containing x_0 and $F(x_0)$ respectively such that the restriction of F on V is a <u>bijection onto</u> W with a <u>C^1-inverse</u>.

(b) The inverse is <u>C^k</u> when <u>F is C^k</u>, $1 \le k \le \infty$, in V.



<u>Note</u>: We only have <u>local</u> invertibility by the IFT.

Let see some examples before proving the IFT.

<u>eg 3.8</u>: Let

$$F : (0,\infty) \times (-\infty,\infty) \to \mathbb{R}^2$$
$$(r, \theta) \mapsto (r\cos\theta, r\sin\theta)$$

Then

$$DF = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} \quad \text{invertible} \quad \forall (r,\theta)$$

Then IFT $\Rightarrow$ F is locally invertible at every point $(r,\theta) \in (0,\infty) \times (-\infty,\infty)$.

But F is clearly <u>not globally invertible</u> as it is not one-to-one:

$$F(r, \theta + 2\pi) = F(r, \theta).$$

eg 3.9 (1-dim. is special)

U = open interval (a,b) in $\mathbb{R}$ $(n=1)$

C^1 function $f:(a,b)\to\mathbb{R}$ with $f'\neq 0$

$\Rightarrow$ f strictly increasing or decreasing

$\Rightarrow$ global inverse exists.

eg 3.10 : (i) $F:\mathbb{R}^2\to\mathbb{R}^2 : (x,y)\mapsto(x^2,y)$. (Caution: as 2-d is simply, we use notation as in Advanced Calculus)

Then $DF=\begin{pmatrix}2x & 0\\ 0 & 1\end{pmatrix}$ singular at $(x,y)=(0,0)$.

F doesn't satisfy the condition DF invertible in the IFT (at $(0,0)$)

And clearly F is not invertible near $(x,y)=(0,0)$ as

$F(\pm a,b)=(a^2,b)$ (2-to-1 near $(0,0)$).

$\therefore$ "DF invertible" condition can't be removed from IFT.

(ii) $H:\mathbb{R}^n\to\mathbb{R}^n : (x,y)\mapsto(x^3,y)$ is bijective

& $H^{-1}(x,y)=(x^{1/3},y)$ exists.

But $DH=\begin{pmatrix}3x^2 & 0\\ 0 & 1\end{pmatrix}$ singular at $(x,y)=(0,0)$.

The point is: H^{-1} is not C^1 near $(x,y)=(0,0)$.

$\therefore$ "DF invertible" is only a "sufficient" condition, but not "necessary", for "local invertibility".

Terminology: The condition in IFT that $DF(x_0)$ is invertible is called the nondegeneracy condition (at x_0)

By eg 3.10, without nondegeneracy condition, the map may or may not be local invertible.

But Nondegeneracy condition is necessary for the differentiability of the local inverse:

Prop 3.8: Let $F: U \subset \mathbb{R}^n \text{ open} \longrightarrow \mathbb{R}^n$ be C^1, and $x_0 \in U$.

Suppose $\exists$ open V s.t. $x_0 \in V \subset U$, and $F|_V$ has a differentiable inverse.

Then $DF(x_0)$ is non-singular. (ie. invertible).

Pf : Suppose the local inverse $(F|_V)^{-1}$ exists and is differentiable at the point $y_0 = F(x_0)$.

Then Chain rule $\Rightarrow$

$$D(F^{-1})(y_0)\, DF(x_0) = \text{Identity}$$

$\Rightarrow$ $DF(x_0)$ is invertible. #