The proof of the Inverse Function Theorem

Perturbation Lemma: Assume that X is a complete normed space and $r>0$. Assume also that $\vec{H}: \overline{B}(0,r) \to X$ has the following properties

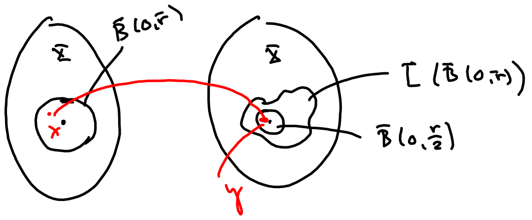
(i) $\vec{H}(\vec{0}) = \vec{0}$

(ii) $\|\vec{H}(\vec{x}) - \vec{H}(\vec{y})\| \leq \frac{1}{2}\|\vec{x}-\vec{y}\|$ for all $x, y \in \overline{B}(0,r)$.

Define $L: \overline{B}(0,r) \to X$ by

$$\vec{L}(\vec{x}) = \underbrace{\vec{x}}_{\text{identity}} + \underbrace{\vec{H}(\vec{x})}_{\text{perturbation}}$$

Then L is injective and $\overline{B}(\vec{0}, \frac{r}{2}) \subseteq L(\overline{B}(\vec{0},r))$, i.e. for any $\vec{y} \in \overline{B}(\vec{0}, \frac{r}{2})$ there is an $\vec{x} \in \overline{B}(\vec{0}, r)$ such that $\vec{y} = L(\vec{x})$.



Proof: L is injective: Assume $L(\vec{x}_1) = L(\vec{x}_2)$. Then

$$\vec{L}(\vec{x}_1) = \vec{x}_1 + \vec{H}(\vec{x}_1)$$
$$\|$$
$$\vec{L}(\vec{x}_2) = \vec{x}_2 + \vec{H}(\vec{x}_2)$$

Hence $\vec{x}_1 + \vec{H}(\vec{x}_1) = \vec{x}_2 + \vec{H}(\vec{x}_2)$

$$\|\vec{x}_1 - \vec{x}_2\| = \|H(\vec{x}_2) - H(\vec{x}_1)\| \leq \frac{1}{2}\|\vec{x}_1 - \vec{x}_2\|$$

Thus $\|\vec{x}_1 - \vec{x}_2\| = 0 \Rightarrow \vec{x}_1 = \vec{x}_2$.

We next need to prove that for any $\vec{y} \in \overline{B}(0;\frac{r}{2})$, there is an $\vec{x} \in \overline{B}(0;r)$ such that $\vec{y} = L\vec{x} = \vec{x} + \vec{H}(\vec{x})$, i.e. $\vec{x} = \underbrace{\vec{y} - \vec{H}(\vec{x})}_{\vec{F}_{\vec{y}}(\vec{x})}$. Need to find a fixed point for $\vec{F}_{\vec{y}}$.

We want to apply Banach's Fixed Point Theorem to

$$\vec{F}_y: \overline{B}(\vec{0}, r) \to \overline{B}(\vec{0}, r)$$

Need to check that $\vec{F}_y(\vec{x}) \in \overline{B}(0, r)$: $\frac{1}{2}\|\vec{x}-0\| = \frac{1}{2}\|\vec{x}\|$

$$\|\vec{F}_y(\vec{x})\| = \|\vec{y} - H(\vec{x})\| \leq \underbrace{\|\vec{y}\|}_{\leq \frac{r}{2}} + \|H(\vec{x})\| = \|\vec{y}\| + \|\vec{H}(\vec{x}) - \vec{H}(\vec{0})\|$$
$$\leq \|y\| + \frac{1}{2}\|x\| \leq \frac{r}{2} + \frac{r}{2} = r.$$

Since $\overline{B}(\vec{0}, r)$ is complete, we can apply Banach to get that $\vec{F}_y$ has a fixed point since it is a contraction.

$$\|\vec{F}_y(\vec{x}_1) - \vec{F}_y(\vec{x}_2)\| = \|(\cancel{\vec{y}} - \vec{H}(\vec{x}_1)) - (\cancel{\vec{y}} - \vec{H}(\vec{x}_2))\|$$
$$= \|\vec{H}(\vec{x}_2) - \vec{H}(\vec{x}_1)\| \leq \frac{1}{2}\|\vec{x}_1 - \vec{x}_2\|$$

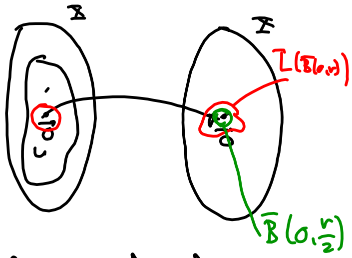
Proposition (A ridiculous simple case of IFT). Assume that $\mathfrak{X}$ is a complete normed space and that $\mathcal{U}$ is an open subset of $\mathfrak{X}$ containing $\vec 0$.

Assume also that $\vec L : \mathcal{U} \to \mathfrak{X}$ is a differentiable function s.t.

(i) $\vec L(\vec 0) = \vec 0$

(ii) $\vec L'(\vec 0) = I$

(iii) $\vec L'$ is continuous at $\vec 0$.



Then there is an $r>0$ such that ~~\~~ $\vec L$ restricted to $\bar B(\vec 0, r)$ has an inverse function M defined on a set containing $\bar B(\vec 0, \frac{r}{2})$. Moreover, M is differentiable at $\vec 0$ and

$$\vec M'(\vec 0) = I.$$

Proof: Define $\vec H(\vec x) = \vec L(\vec x) - \vec x$. Then

$$\vec L(\vec x) = \vec x + \vec H(\vec x) \quad (\text{as in the perturbation lemma})$$

First observe that $\vec H(\vec 0) = \vec L(\vec 0) - \vec 0 = \vec 0$.

Also $\vec H' = L' - I$, hence $\vec H'(\vec 0) = \vec L'(\vec 0) - I = I - I = 0$.

Since $\vec H'$ is continuous, there is an $r>0$ such that if $\vec x \in \bar B(\vec 0, r)$, then $\|\vec H'(\vec x)\| \le \frac{1}{2}$. But then if $\vec x, \vec y \in \bar B(\vec 0, r)$, then

$$\|\vec H(\vec x) - \vec H(\vec y)\| \le \frac{1}{2}\|\vec x - \vec y\| \quad (\text{Cor to MVT}).$$

Hence $\vec H$ satisfies the condition of the Perturbation Lemma. This means that $\vec L$ restricted to $\bar B(\vec 0, r)$ is injective and $\vec L(\bar B(\vec 0, r)) \supseteq \bar B(0, \frac{r}{2})$. Hence L has an inverse function $\vec M$ defined on a set containing $\bar B(\vec 0, \frac{r}{2})$.

We still need to prove that $\vec M'(\vec 0) = I$.

Observe: $\|\vec x\| = \|\vec L(\vec x) - \vec H(\vec x)\| \le \|L\vec x\| + \|\vec H\vec x\| \le \|L\vec x\| + \frac{1}{2}\|\vec x\|$

$$\boxed{\begin{aligned}\vec H(\vec x) &= \vec L(\vec x) - \vec x\\ \vec x &= \vec L(\vec x) - \vec H(\vec x).\end{aligned}}$$

Hence

$$\frac{1}{2}\|\vec x\| \le \|\underbrace{L\vec x}_{\vec y}\|$$

which means that

$$\frac{\|\vec x\|}{\|\vec y\|} \le 2 \qquad \|\vec x\| \le 2\|\vec y\|$$

Let us check that $\vec M$ is differentiable at $\vec 0$ with $\vec M'(\vec 0) = I$.

$\sigma_M(\vec y) = \vec M(\vec 0 + \vec y) - \vec M(\vec 0) - I(\vec y)$ (need to prove this goes to 0 faster than $\|\vec y\|$)

$$\therefore = \vec M(\vec y) - \vec y = \vec x - L(\vec x)$$

$$= -\left(\vec L(\vec x) - \vec L(\vec 0) - I(\vec x)\right) = -\sigma_L(\vec x) \qquad \vec y = L(\vec x), \quad \vec x = M(\vec y)$$

Hence

$$\frac{\|\sigma_M(\vec y)\|}{\|\vec y\|} = \frac{\|\sigma_L(\vec x)\|}{\|\vec y\|} = \frac{\|\sigma_L(\vec x)\|}{\|\vec x\|}\underbrace{\frac{\|\vec x\|}{\|\vec y\|}}_{\le 2} \to 0 \text{ as } \vec y \to 0$$

Inverse Function Theorem: Assume that $\bar{X},\bar{Y}$ are two complete normed spaces and that U is an open subset of $\bar{X}$ containing the point $\vec{a}$. Assume further that $\vec{F}: U \to \bar{Y}$ is a differentiable function such that

(i) $\vec{F}'(\vec{a})$ is invertible

(ii) $\vec{F}'$ is continuous at $\vec{a}$.

Then $\vec{F}$ has a local inverse $\vec{G}$ at $\vec{a}$ and $\vec{G}$ is differentiable at $\vec{b} = \vec{F}(\vec{a})$ and

$$\underbrace{\vec{G}'(\vec{b}) = (\vec{F}'(\vec{a}))^{-1}}_{}$$

Proved already!

Proof: Define a new function $\vec{L}$ by

$$\vec{L}(\vec{z}) = A(\vec{F}(\vec{z}+\vec{a}) - \vec{b})$$ defined for z in a neighborhood of $\vec{0}$.

where $A = \vec{F}'(\vec{a})^{-1}$

Note that: $\vec{L}(\vec{0}) = A(\underbrace{\vec{F}(\vec{0}+\vec{a}) - \vec{b}}_{\vec{0}}) = \vec{0}$

Also:

$$\vec{L}'(\vec{z}) = A(\vec{F}'(\vec{z}+\vec{a}))$$

$$\vec{L}'(\vec{0}) = A(\vec{F}'(\vec{a})) = I \quad \text{because } A = (\vec{F}'(\vec{a}))^{-1}$$

This means that $\vec{L}$ has a local inverse function $\vec{M}$.

Assume now that $\vec{y} = \vec{F}(\vec{x})$ and consider

$$\vec{L}(\vec{z}) = A(\vec{F}(\vec{z}+\vec{a}) - \vec{b})$$

Put $\vec{x} = \vec{z} + \vec{a}$

Then $\vec{z} = \vec{x} - \vec{a}$

$\vec{y} = \vec{F}(\vec{x}) = \vec{F}(\vec{z}+\vec{a})$

hence

$$\vec{L}(\vec{x}-\vec{a}) = A(\vec{y} - \vec{b})$$

Since M is the inverse of $\vec{L}$, we get

$$\vec{x} - \vec{a} = \vec{M}(A(\vec{y}-\vec{b}))$$

$$\vec{x} = \vec{M}(\underbrace{A(\vec{y}-\vec{b})}) + \vec{a} = \underbrace{\vec{G}(\vec{y})}_{\text{differentiable.}} \quad \text{Inverse function.}$$

For which y's is $\vec{G}$ defined?

$\vec{M}(\vec{u})$ is defined $\vec{u} \in \bar{B}(\vec{0}, \frac{r}{2})$. Hence we need

$$\|A(\vec{y} - \vec{b})\| \leq \frac{r}{2}$$

$$\|A(\vec{y}-\vec{b})\| \leq \|A\| \, \|\vec{y}-\vec{b}\| \leq \frac{r}{2} \Rightarrow \|\vec{y} - \vec{b}\| \leq \frac{r}{2\|A\|}$$

Implicit Function Theorem

Problem: $f(x_1,\dots,x_n,y)=0$

$$\Big\Downarrow \text{ solve for } y$$

$$y(x_1,\dots,x_n)=g(x_1,\dots,x_n)$$

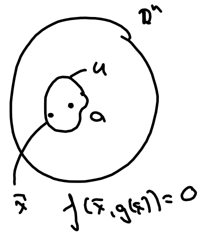
At least prove the existence of g such that

$$f(x_1,\dots,x_n,g(x_1,\dots,x_n))=0$$

Implicit Function Theorem: Assume that $f:\mathbb{R}^n\to\mathbb{R}$ is a differentiable function, $f(x_1,\dots,x_n,y)$, and that there is a point $(\vec a,b)$ such that $f(\vec a,b)=0$. If $\frac{\partial f}{\partial y}(\vec a,b)\neq 0$, and $\frac{\partial f}{\partial y}$ is continuous at $(\vec a,b)$, then there is an open neighborhood U of $\vec a$ and a function $g:U\to\mathbb{R}$ such that $f(\underbrace{x_1,\dots,x_n}_{\vec x},g(x_1,\dots,x_n))=0$ for $\vec x\in U$.

Moreover, g is differentiable at $\vec a$ and

$$\frac{\partial g}{\partial x_i}(\vec a)=-\frac{\frac{\partial f}{\partial x_i}(\vec a,b)}{\frac{\partial f}{\partial y}(\vec a,b)}$$



Proof of formula:

$$f(x_1,\dots,x_n,g(x_1,\dots,x_n))=0$$

Partial derivative on both sides

$$\frac{\partial}{\partial x_i}\big(f(x_1,\dots,x_n,\underline{g(x_1,\dots,x_n)})\big)=0$$

$$\Vert$$

$$\frac{\partial f}{\partial x_i}+\frac{\partial f}{\partial y}\cdot\frac{\partial g}{\partial x_i}=0\Rightarrow\frac{\partial g}{\partial x_i}=-\frac{\frac{\partial f}{\partial x_i}}{\frac{\partial f}{\partial y}}.$$

How do we prove it? (Sketch). $f(\vec x,y)=0$

Look: $\underbrace{(\vec x,y)}_{\mathbb{R}^{n+1}}\xrightarrow{k}\underbrace{(\vec x,f(\vec x,y))}_{\mathbb{R}^{n+1}}$ satisfies the inverse function theorem.

$$(\vec x,z)\xrightarrow{L}L(\vec x,z)\to(\vec x,h(\vec x,z))$$

$$(\vec x,z)=k(L(\vec x,z))=k(\vec x,h(\vec x,z))=(\vec x,f(\vec x,h(\vec x,z)))$$

Hence $z=f(\vec x,h(\vec x,z))$ for all $\vec x$ and z

$z=0$: $0=f(\vec x,\underbrace{h(\vec x,0)}_{g(\vec x)})$