The Contraction Principle

Let (X,d) be a metric space and let $\varphi: X \to X$.

If $\exists c \in \mathbb{R}$, $c<1$ with $d(\varphi(x),\varphi(y)) \leq c\, d(x,y)$ $\forall x,y \in X$, then we say that φ is a contraction (of X into X)

Theorem: Let X be a complete metric space & let $\varphi: X \to X$ be a contraction. Then there exist a unique $x \in X$ with $\varphi(x) = x$.

Proof: (I) Existence: Let $x_0 \in X$ be arbitrary.

and define the sequence $x_j = \varphi(x_{j-1})$

then $d(x_{j+1}, x_j) = d(\varphi(x_j), \varphi(x_{j-1})) \leq c\, d(x_j, x_{j-1})$

$\leq c^j\, d(x_1, x_0)$

$\Rightarrow$ for $n > m$, $d(x_n, x_m) \leq \sum_{j=m+1}^{n} d(x_j, x_{j-1})$

$\leq d(x_1,x_0) \sum_{j=m+1}^{n} c^j = c^{m+1} d(x_1,x_0) \underbrace{\sum_{j=0}^{n-m-1} c^j}_{\frac{1-c^{n-m}}{1-c}}$

$\leq \frac{c^{m+1}}{1-c}\, d_1(x_1,x_0)$

$\Rightarrow \{x_m\}$ is a cauchy sequence & by completeness $\lim_{m\to\infty} x_m = x \in X$

Finally $\varphi(x) = \lim_{m\to\infty} \varphi(x_m) = \lim_{m\to\infty} x_{m+1} = x$ by continuity of φ

(II) Uniqueness: Suppose $\varphi(x) = x$, $\varphi(y) = y$

$\Rightarrow d(x,y) = d(\varphi(x), \varphi(y)) \leq c\, d(x,y)$

So $d(x,y) = 0$ (as $c < 1$) $\Rightarrow x = y$.

The Inverse Function Theorem:

Theorem: Suppose that $f: E \subset \mathbb{R}^n \to \mathbb{R}^n$ (where E is an open set) is continuously differentiable (i.e. $f \in C'$) & suppose that $\vec{f}'(\vec{a})$ is invertible for some $a \in E$ with $b = f(a)$.

Then (a) $\exists$ open sets $U, V \subset \mathbb{R}^n$ where $a \in U$, $b \in V$ f is one-to-one on U & $f(U) = V$.

(b) If $g: V \to U$ is the inverse of f, defined by

$g(f(x)) = x$

then $g \in C'(V)$.

— × —

Interpretation: If the total derivative $\mathcal{D}$ ($f \in C'$) is invertible at $\vec{a}$, then f is invertible in a neighborhood of $\vec{a}$. Also, the inverse, f^{-1}, is continuously diff.

So, given a system of equations $\vec{f}(\vec{x}) = \vec{y}$, where $\vec{f}(\vec{a}) = \vec{b}$ & $\vec{f}'(\vec{a})$ is invertible, the system $\vec{f}(\vec{x}) = \vec{y}$ has a unique soln near $\vec{a}$, $\forall \vec{y}$ in a neighborhood of $\vec{b}$.

Proof of (a) (proof that f is 1-to-1 on U & $f(U)=V$ where $a\in U$, $b\in V$, U, V open)

- Define $A = \vec{f}'(\vec{a})$ (and note this is a linear transf.)

Since $f \in C'$, f' is continuous. So there exists an open ball $U \subset E$ (centered at $\vec{a}$) with

$$\|f'(x) - A\| < \frac{1}{2\|A^{-1}\|}$$

$\forall x \in U$.

- Let us define $\varphi_y(x) = x + A^{-1}(y - f(x))$, $x \in E$

where $\varphi_y(x) = x$ when $y = f(x)$ (i.e. such an x is a fixed pt of φ_y)

We want to make sure $\exists$ a unique fixed pt. in U as this ensures $f(x) = y$ has a unique solution.

So we will use the contraction principle, but to do that, we need $|\varphi_y(x_1) - \varphi_y(x_2)| \leq c|x_1 - x_2|$ for some $c<1$

Observe that $\varphi_y'(x) = I - A^{-1}f'(x) = A^{-1}(A - f'(x))$

$$\Rightarrow \|\varphi_y'(x)\| \leq \|A^{-1}\| \underbrace{\|A - f'(x)\|}_{< \frac{1}{2\|A^{-1}\|} \text{ in } U}$$

$$\Rightarrow \|\varphi_y'(x)\| < \tfrac{1}{2} \quad \forall x \in U.$$

So $|\varphi_y(x_1) - \varphi_y(x_2)| < \frac{1}{2}|x_1 - x_2|$ $\forall x_1, x_2 \in U$ (by a previous theorem)

$\Rightarrow$ φ_y has $\uparrow$ unique fixed pt. in U & $f(x) = y$ has, $\uparrow$ unique soln in U $\Rightarrow$ f is one-to-one in U.

(at most a; at most, one)

We still need to prove that $f(U) =: V$ is open.

- Pick some $y_0 \in V \Rightarrow \exists x_0 \in U$ s.t. $f(x_0) = y_0$ and choose $r > 0$ so small that $rB(x_0)$ (open ball centered at x_0, radius r) satisfies $\overline{rB(x_0)} \subset U$. (we will need this later) To prove that V is open we will show that $y \in V$ whenever $|y - y_0| < \frac{r}{2\|A^{-1}\|}$.

Let y be s.t. $|y - y_0| < \frac{r}{2\|A^{-1}\|}$

Then $|\varphi_y(x_0) - x_0| = |x_0 + A^{-1}(y - f(x_0)) - x_0| = |A^{-1}(y - f(x_0)|$

$\leq \|A^{-1}\| |y - y_0| < \frac{r}{2}$

$\Rightarrow \forall x \in \overline{rB(x_0)} \subset U$:

$|\varphi_y(x) - x_0| \leq \underbrace{|\varphi_y(x_0) - x_0|}_{< r/2} + \underbrace{|\varphi_y(x) - \varphi_y(x_0)|}_{< \frac{1}{2}|x - x_0|}$ (φ_y is a contraction with $c < \frac{1}{2}$)

$$|\varphi_y(x)-x_0| < r/2 + \tfrac{1}{2}\underbrace{|x-x_0|}_{\leq r} \leq r$$

i.e. $|\varphi_y(x)-x_0| < r$ and $\varphi_y(x) \in r\overline{B}(x_0)$

Recall that we showed $|\varphi_y(x_1)-\varphi_y(x_2)| \leq \frac{1}{2}|x_1-x_2|$

$\forall x_1, x_2 \in \mathcal{U}$ but $\overline{r\overline{B}(x_0)} \subset \mathcal{U}$ so φ_y is a contraction on $\overline{r\overline{B}(x_0)}$. Note that since $\overline{r\overline{B}(x_0)} \subset \mathbb{R}^n$ is closed, it is complete

- So by the contraction mapping theorem, φ_y has a unique fixed pt. in $\overline{r\overline{B}(x_0)}$. Call it x.

note that $f(x) = y$.

So we are done with (a).

Proof of (b) (the inverse of f is in $C^1(V)$)

Need to show that the function g: $g(\underbrace{f(x)}_{\in V}) = \underbrace{x}_{\in U}$ is C^1

Strategy: $g(f(x)) = x$ so if everything where nice

$g'(f(x))f'(x) = I$

so $g'(y)f'(g(y)) = I$ so $g'(y) = \underbrace{(f'(g(y)))^{-1}}$

this is our candidate

For the derivative, our task is to show this is true and $g'(y)$ is cont.

Let $y \in V$ & $y+k \in V$

$\Rightarrow \exists x \in U$ & $x+h \in U$ s.t. $f(x) = y$, $f(x+h) = y+k$

Using $\varphi_y(x) = x + A^{-1}(y - f(x))$

we see that $\varphi_y(x+h) - \varphi_y(x) = x+h + A^{-1}(y - f(x+h)) - (x + A^{-1}(y-f(x)))$

$= h + A^{-1}(f(x) - f(x+h))$

$= h + A^{-1}(y - y - k)$

$= h - A^{-1}k$

$\Rightarrow |\varphi_y(x+h) - \varphi_y(x)| = |h - A^{-1}k| \leq \frac{1}{2}|x+h-x|$ (φ_y is a contraction with $c = \frac{1}{2}$)

$\Rightarrow |h - A^{-1}k| \leq \frac{1}{2}|h|$

$\Rightarrow |h| - |A^{-1}k| \leq \frac{1}{2}|h|$ (reverse Δ-ineq)

So $|h| \leq 2|A^{-1}k| \leq 2\|A^{-1}\||k|$ (*)

Recall that on U: $\|f'(x) - A\| \leq \frac{1}{2\|A^{-1}\|}$

So $\|f'(x) - A\|\|A^{-1}\| \leq \frac{1}{2}$

But $\|B - A\|\|A^{-1}\| \leq 1$ implies B is invertible

So $f'(x)$ is invertible

Let T be the inverse of $f'(x)$.

So:

$F(x+h) = y+k \Rightarrow \quad x+h = g(y+k)$

$$|g(y+k)-g(y)-Tk| = |x+h-x-Tk|$$
$$= |h-Tk|$$
$$= |T(T^{-1}h - k)|$$
($T^{-1}h \to f'(x)$, $k \to f(x+h)-f(x)$)
$$\leq \|T\| \, |f'(x)h - f(x+h) + f(x)|$$

$$\Rightarrow \frac{|g(y+k)-g(y)-Tk|}{|k|} \leq \|T\| \frac{|f(x+h)-f(x)-f'(x)h|}{|k|}$$
$$\geq |h| / (\|k\| \|A^{-1}\|) \text{ by } \circledast$$
$$\leq 2\|A^{-1}\| \, \|T\| \, \frac{|f(x+h)-f(x)-f'(x)h|}{|h|}$$

by $\circledast$ $h \to 0$ as $k \to 0$

Moreover the RHS $\to 0$ as $h \to 0$

$\Rightarrow$ LHS $\to 0$ $\Rightarrow$ T is the derivative of g!

i.e., $g'(y) = T$ where T was the inverse of $f'(x)$

$f'(g(y))$

$$\Rightarrow g'(y) = (f'(g(y)))^{-1} \quad -- (**)$$

So we have shown that g is differentiable, hence a continuous map of V to U

but f is in C^1 so f' is also a cont. map from U to the set Ω of invertible elements of $L(\mathbb{R}^n)$

and inversion is cont. map of $\Omega \to \Omega$

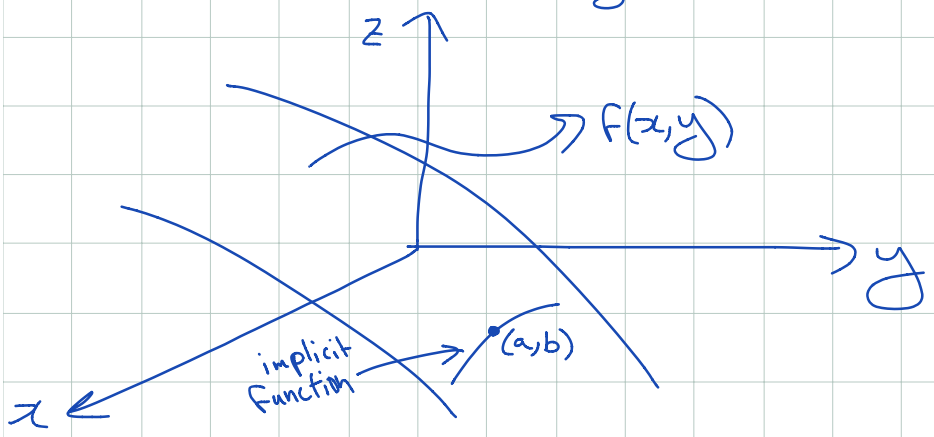
So their composition $(**)$ is cont. $\Rightarrow g \in C_1(V)$ ∎

Theorem: If f is a C' mapping of an open set $E \subset \mathbb{R}^n$ into $\mathbb{R}^n$ & if $f'(x)$ is invertible $\forall x \in E$ then $f(W)$ is an open subset of $\mathbb{R}^n$ for every open set $W \subset E$.

Proof: Follows from part (a) of Inv. Ft. theorem.

The Implicit Function theorem

Let's start with $f: \mathbb{R}^2 \to \mathbb{R}$ with $f(a,b) = 0$ for some (a,b) satisfying $\frac{\partial f}{\partial y}(a,b) \neq 0$



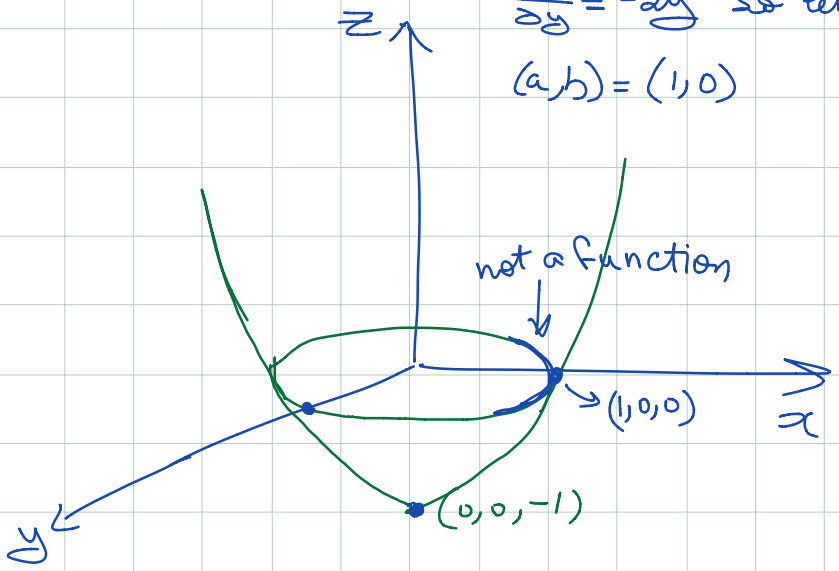
around (a,b) we can write $y = g(x)$ for some implicit $g(x)$ & still have $f(x, g(x)) = 0$

now to see why $\frac{\partial f}{\partial y}(a,b) \neq 0$ is important, consider

$f(x,y) = 1 - x^2 - y^2$

$\frac{\partial f}{\partial y} = -2y$ so let's look at

$(a,b) = (1,0)$



$f(1,0) = 0$ & $\frac{\partial f}{\partial y}\Big|_{(1,0)} = 0$ but the points around $(1,0)$ where the surface cuts the xy-plane (i.e. $f(x,y)=0$) don't form a function of x.

On the other hand, $f(1,0) = 0$ & $\frac{\partial f}{\partial x}\Big|_{(1,0)} = 2 \neq 0$

So we can write $x = g(y)$ and have $f(g(y), y) = 0$ around $(1,0)$.

Linear version of the implicit function theorem

Theorem: If $A \in L(\mathbb{R}^{n+m}, \mathbb{R}^n)$ & if A_x is invertible then there corresponds to every $k \in \mathbb{R}^m$ a unique $h \in \mathbb{R}^n$ such that $A\begin{pmatrix} h \\ k \end{pmatrix} = 0$. Moreover $\vec{h} = -A_x^{-1} A_y \vec{k}$.

(Here $\begin{pmatrix} \vec{h} \\ \vec{k} \end{pmatrix} := (h_1, \dots, h_n, k_1, \dots, k_m) =: (\vec{h}, \vec{k})$ & A_x: $A_x \vec{h} = A\begin{pmatrix} \vec{h} \\ \vec{0} \end{pmatrix}$.
A_y: $A_y \vec{k} = A\begin{pmatrix} \vec{0} \\ \vec{k} \end{pmatrix}$)

Proof: Note that $A_x h + A_y k = A\begin{pmatrix} \vec{h} \\ \vec{k} \end{pmatrix}$

$A\begin{pmatrix} \vec{h} \\ \vec{k} \end{pmatrix} = 0 \iff A_x h + A_y k = 0$

So $h = -A_x^{-1} A_y k$ ∎

Theorem (Implicit fct theorem)

Let $E \subset \mathbb{R}^{n+m}$ be open & let $f: E \subset \mathbb{R}^{n+m} \to \mathbb{R}^n$ satisfy $f(\vec{a}, \vec{b}) = 0$ for some $\vec{a} \in \mathbb{R}^n$, $\vec{b} \in \mathbb{R}^m$, $(\vec{a}, \vec{b}) \in E$

Set $A = f'(\vec{a}, \vec{b})$ and assume A_x is invertible.

Then $\exists$ open sets $U \subset \mathbb{R}^{n+m}$ & $W \subset \mathbb{R}^m$ with $(\vec{a}, \vec{b}) \in U$ & $\vec{b} \in W$ s.t.

(i) $\forall \vec{y} \in W$ $\exists$ a unique $\vec{x}$ with $(\vec{x}, \vec{y}) \in U$ & $f(\vec{x}, \vec{y}) = 0$

(ii) Writing $\vec{x} = g(\vec{y})$, g is a C' map of W into $\mathbb{R}^n$ with $g(\vec{b}) = \vec{a}$, $f(g(\vec{y}), \vec{y}) = 0$ & $g'(b) = -(A_x)^{-1} A_y$.

Remark: $A_{\vec{x}}$ is invertible means $[A_{\vec{x}}] := \begin{bmatrix} D_1 f_1 & \cdots & D_n f_1 \\ \vdots & & \\ D_1 f_n & \cdots & D_n f_n \end{bmatrix}$ is invertible; that is $\det([A_{\vec{x}}]) \neq 0$.

Proof of Theorem (Part I)

Strategy: • We will define $F(x,y) = (f(x,y), y)$ and show that $F'(a,b)$ is invertible, so we may use the inverse function theorem to conclude that $\exists$ open sets U & V in $\mathbb{R}^{n+m}$ with $(a,b) \in U$, $(0,b) \in V$

and F is 1-1 and onto from $U \to V$.

Let $W = \{ y \in \mathbb{R}^m : (0,y) \in V \}$ (W is open since V is open)

$\forall y \in W$, $(0,y) \in V \Rightarrow (0,y) = F(x,y)$ for some $(x,y) \in U$ (since $V = F(U)$). For this (x,y), $f(x,y) = 0$

Remains to show x is unique: suppose that with the same y: $f(x',y) = 0$ for some $(x',y) \in U$

$$\Rightarrow (f(x',y), y) = (f(x,y), y)$$

$$\Rightarrow F(x',y) = F(x,y)$$

but F is 1-1 on $U \Rightarrow x = x'$

So, we're done once we show that $F'(a,b)$ is invertible.

$$F(x,y) = (f(x,y), y) \quad \forall (x,y) \in E$$

Since $f \in C'$, F is also C', hence differentiable

$$f(a+h, b+k) = \underbrace{f(a,b)}_{0} + A(h,k) + \underbrace{r(h,k)}_{\text{remainder}} \quad (*)$$

$$\Rightarrow F(a+h, b+k) - F(a,b)$$

$$= \left(f(a+h,b+k), b+k\right) - \left(\underbrace{f(a,b)}_{0}, b\right)$$

$$= \left(f(a+h,b+k) - 0, b+k-b\right)$$

$$= \left(f(a+h,b+k), k\right)$$

$$= \left(A(h,k) + r(h,k), k\right) \quad (\text{by } *)$$

$$= \left(A(h,k), k\right) + \underbrace{\left(r(h,k), 0\right)}_{\frac{\|r(h,k)\|}{\|(h,k)\|} \xrightarrow[(h,k)\to 0]{} 0}$$

$\Rightarrow F'(a,b)$ is the linear operator

$$(h,k) \longmapsto (A(h,k),k)$$

To see that it is invertible take $(A(h,k),k)=0 \Rightarrow k=0$ & $A(h,0)=0 \Rightarrow h=0$ by the linear version of the IFT.

$\Rightarrow F'(a,b)$ is 1-1 $\Rightarrow$ invertible

Proof of part 2: (If $x=:g(y)$, then $g: W \to \mathbb{R}^n$ is C^1 $g(b)=a$, $f(g(y),y)=0$ $\forall y \in W$)

For $y \in W$, Pick $g(y)$ s.t. $f(g(y),y)=0$ & $(g(y),y) \in U$

$$\Rightarrow F(g(y),y) = (\underbrace{f(g(y),y)}_{0},y) = (0,y) \quad \forall y \in W$$

Let $G: V \to U$ be the inverse of $F \Rightarrow G$ is C^1 (IFT) & $(g(y),y) = G(0,y)$ $\forall y \in W$

$\Rightarrow g$ is C^1 (since G is C^1)

Remains to show $g'(b) = -(A_x)^{-1} A_y$

Let $\varphi(y) = (g(y),y) \Rightarrow \varphi'(y)k = (g'(y)k, k)$ (**)

$$\Rightarrow f(\varphi(y)) = f(g(y),y) = 0 \quad \forall y \in W$$

$\overset{\text{chain rule}}{\Rightarrow}$ $f'(\varphi(y))\varphi'(y) = 0$

Take $y=b$, then $F'(\varphi(b))\,\varphi'(b)=0$

$(g(b),b)=(a,b)$

$\Rightarrow F'(a,b)\,\varphi'(b)=0$

A $\quad (g'(b),I)$ by (**)

$\Rightarrow A_x g'(b) + A_y = 0$ ∎