Course 224: Geometry - Continuity and Differentiability

Lecture Notes by Prof. David Simms

LaTeXed by Chris Blair

1 Continuity

Consider M, N finite dimensional real or complex vector spaces. What does it mean to say that a map

$$M \supset X \xrightarrow{f} Y \subset N$$

is continuous? Let $a, x \in X \subset M$, and $f(a), f(x) \in Y \subset N$, then we can make "$f(x)$ arbitrarily close to $f(a)$" by taking "x sufficiently close to a." To make this precise, we choose norms on M, N.

Definition: (Norm) A norm or length function on M is a function $||\cdot||$ such that

$$\begin{aligned} M &\to \mathbb{R} \\ x &\mapsto ||x|| \end{aligned}$$

such that:
i) $||x+y|| \leq ||x|| + ||y||$ (triangle inequality),
ii) $||\alpha x|| = |\alpha| ||x||$,
iii) $||x|| \geq 0$ with equality if and only if $x = 0$.

Definition: (Normed Vector Space) A real or complex vector space with a chosen norm is called a normed vector space. From now on, M, N etc. are normed spaces.

Definition: (Ball) Let $X \subset M$, $a \in X, r > 0$, then the ball in X of radius r, centred at a is

$$B_X(a, r) = \{x \in X \mid ||x - a|| < r\}$$

Definition: (Open) A set $V \subset X$ is called open in X if for each $a \in V$ there exists an $r > 0$ such that $B_X(a, r) \subset V$, ie each point of V is an interior point.

Theorem 1.1.
$B_X(a, r)$ is open in X.

Proof. Let $b \in B_X(a, r)$. Then $||b - a|| < r \Rightarrow r - ||b - a|| > 0$. Let $s > 0$ be less than $r - ||b - a||$, then

$$\begin{aligned} x \in B_X(b, s) &\Rightarrow ||x - b|| < s < r - ||b - a|| \\ &\Rightarrow ||x - b|| + ||b - a|| < r \\ &\Rightarrow ||x - a|| < r \Rightarrow x \in B_X(a, r) \end{aligned}$$

using the triangle inequality. Hence $B_X(b, s) \subset B_X(a, r)$ as required. $\square$

Definition: (Continuity) A map $M \supset X \xrightarrow{f} Y \subset N$ is called continuous at $a \in X$ if for each $\varepsilon > 0$ $\exists\, \delta > 0$ such that $f\,[B_X(a,\delta)] \subset B_Y\,(f(a),\varepsilon))$

We call f continuous if it is continuous at a for every $a \in X$.

Theorem 1.2.
Let $M \supset X \xrightarrow{f} Y \subset N$, then f is continuous at $a \Leftrightarrow$ for each V open in Y such that $f(a) \in V$ there exists W open in X such that $fW \subset V$.

Proof. i) Let f be continuous at a, and let V open in Y such that $f(a) \in V$. V is open, so there exists $\varepsilon > 0$ such that $B_Y(f(a),\varepsilon) \subset V$. f is continuous, so therefore there exists $\delta > 0$ such that $fB_X(a,\delta) \subset B_Y(f(a),\varepsilon)$, where $B_X(a,\delta)$ is open in X as required.

ii) Let V be open in Y such that $f(a) \in V \Rightarrow$ there exists W open in X such that $a \in W$ and $fW \subset V$. V is open so there exists $\varepsilon > 0$ such that $B_Y(f(a),\varepsilon)$ open in Y, $f(a) \in B_Y(f(a),\varepsilon)$. Therefore there exists W open in X such that $a \in W$ and $fW \subset B_Y(f(a),\varepsilon)$.

W is open in X, so there exists $\delta > 0$ such that $B_X(a,\delta) \subset fW \subset B_Y(f(a),\varepsilon)$, hence f is continuous at a as required. □

This theorem shows that the continuity at a of $X \xrightarrow{f} Y$ depends only on the open sets of X and Y.

Definition: (Usual Topology) The collection of open sets of X, where X is a subset of a finite dimensional real or complex vector space, is called the usual topology on X, and is independent of the choice of norm.

Definition: (Topology on X) Let X be a set. A topology on X is a collection of subsets of X, called the open sets of the topology, such that

i) $\emptyset, X$ open

ii) If $\{V_i\}_{i \in I}$ a collection of sets V_i open in X then their union $\bigcup V_i$ is open in X

iii) If $V_1 \ldots V_k$ a finite collection of sets V_i open in X, then their intersection $V_1 \cap V_2 \cap \cdots \cap V_k$ is open in X

Definition: (Topological Space) A set X together with a topology on X is called a topological space.

Definition: (Continuous at a) Let $X \xrightarrow{f} Y$, X, Y topological spaces, and $a \in X$. Then f is continuous at a if for each V open in Y such that $f(a) \in V$ there exists W open in X such that $a \in W$ and $fW \subset V$.
f is continuous if it is continuous at a for all $a \in X$.

Note: If X a topological space and if $V \subset X$, and if for each $a \in V$ there exists open W such that $a \in W_a \subset V$, then V is open.

Proof. $V = \bigcup W_a$ is a union of open sets and is therefore open. □

Theorem 1.3.
Let $X \xrightarrow{f} Y$, X, Y topological spaces, then f is continuous $\Leftrightarrow V$ open in $Y \Rightarrow f^{-1}V$ open in X.

Proof. i) Let f be continuous, and V be open in Y. Let $a \in f^{-1}V$, then there exists W_a open in X such that $a \in W_a \subset f^{-1}V$ (as f continuous). $f^{-1}V = \bigcup W_a$ a union of open sets, hence $f^{-1}V$ is open.

ii) Let V be open in $Y \Rightarrow f^{-1}V$ open in X. Let $a \in X$, and let V open in Y be such that $f(a) \in V$. Therefore $a \in f^{-1}V$ open in X, and $f\left[f^{-1}V\right] \subset V$. $f^{-1}V$ is open, so then f is continuous at a for all $a \in X$, and so f is continuous. $\square$

Theorem 1.4.
Let

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[dr, "gf"'] & Y \arrow[d, "g"] \\ & Z \end{tikzcd}$$

then f, g continuous $\Rightarrow gf$ continuous.

Proof. Let V be open in Z, then $(gf)^{-1}V = f^{-1}\left[g^{-1}V\right]$, and $g^{-1}V$ is open as g is continuous, and $f^{-1}g^{-1}V$ is open as f is continuous, hence gf is continuous. $\square$

Thus we have a category *top*, whose objects are topological spaces $X, Y, Z \dots$ and whose morphisms

$$X \xrightarrow{f} Y$$

are continuous maps.

We consider, in calculus, the category whose objects are open subsets $V, U, W \dots$ of finite dimensional real or complex vector spaces, and whose morphisms are continuous maps

$$V \xrightarrow{f} W$$

The isomorphisms in these categories are called homeomorphisms, hence we could ask

"Wherefore art thou homeomorphism?"

2 Differentiability

Definition: (Differentiable) Let $M \supset V \xrightarrow{f} W \subset N$, where V, W are open subsets of real or complex vector spaces M, N. Let $a \in V$, then f is differentiable at a if there exists a linear operator

$$M \xrightarrow{f'(a)} N$$

called the derivative of f at a, such that

$$f(a+h) = f(a) + f'(a)h + \phi(h)$$

where $f'(a)h$ is a linear approximation to the change in f when a changes by h, and $\phi(h)$ is a remainder term such that

$$\frac{||\phi(h)||}{||h||} \to 0 \text{ as } ||h|| \to 0$$

i.e. for each open V, $0 \in V$, in $\mathbb{R}$ there exists open W, $0 \in W$, in M such that $\frac{||\phi(h)||}{||h||} \in V$ for all $h \in W, h \neq 0$. Or equivalently, for any chosen norm, for each $\varepsilon > 0$ there exists $\delta > 0$ such that $\frac{||\phi(h)||}{||h||} < \varepsilon \ \forall ||h|| < \delta, h \neq 0$.

Theorem 2.1.
Let $M \supset V \xrightarrow{f} W \subset N$ be differentiable at $a \in V$. Then the derivative $f'(a)$ is uniquely determined by the formula

$$\begin{aligned} f'(a)h &= \lim_{t\to 0} \frac{f(a+th)-f(a)}{t} \\ &= \frac{d}{dt} f(a+th)\Big|_{t=0} \\ &= \textit{the directional derivative of } f \textit{ at } a \textit{ along } h \end{aligned}$$

Proof.

$$f(a+th) = f(a) + f'(a)th + \phi(th)$$

$$\Rightarrow \left|\left|\frac{f(a+th)-f(a)}{t} - f'(a)h\right|\right| = \frac{||\phi(th)||}{||th||}||h||$$

which $\to 0$ as $t \to 0$, as required. □

Definition: (Function of n independent variables) Let $\mathbb{R}^n \supset V \xrightarrow{f} \mathbb{R}$, V open in $\mathbb{R}^n$, then f is called a real-valued function of n independent variables.

Definition: (Partial Derivative) If $a = (a_1, \dots, a_n) \in V$, and $x = (x_1, \dots, x_n)$, $f(x) = f(x_1, \dots, x_n)$, then we define

$$\begin{aligned} \frac{\partial f}{\partial x^j}(a) &= \frac{\partial f}{\partial x^j}(a_1, \dots, a_n) \\ &= \lim_{t\to 0} \frac{f(a_1, \dots, a_j + t, \dots, a_n) - f(a_1, \dots, a_j, \dots a_n)}{t} \\ &= \lim_{t\to 0} \frac{f(a+te_j) - f(a)}{t} \\ &= \frac{d}{dt} f(a+te_j)\Big|_{t=0} \\ &= \text{directional derivative of } f \text{ at } a \text{ along } e_j \end{aligned}$$

called the partial derivative of f at a with respect to the j^{th} usual coordinate function x^j.

Theorem 2.2.
Let $\mathbb{R}^n \supset V \xrightarrow{f} \mathbb{R}^m$ be differentiable, V open, where $f(x) = (f^1(x), \dots, f^m(x))$, $f^i(x) = f^i(x_1, \dots, x_n)$. Then the derivative

$$\mathbb{R}^n \xrightarrow{f'(a)} \mathbb{R}^m$$

is the $m \times n$ matrix

$$f'(a) = \left(\frac{\partial f^i(a)}{\partial x^j}\right) \quad i = 1 \dots m \ , \ j = 1 \dots n$$

Proof. The j^{th} column of $f'(a) = f'(a)e_j = \frac{d}{dt}\Big(f(a+te_j)\Big)\Big|_{t=0} = \frac{\partial f(a)}{\partial x^j} = \left(\frac{\partial f^1(a)}{\partial x^j}, \dots, \frac{\partial f^m(a)}{\partial x^j}\right)$ as required. □

Definition: (Operator Norm) Let $M \xrightarrow{T} N$ be a linear operator on M, N finite dimensional normed vector spaces. We write

$$||T|| = \sup_{||u||=1} ||Tu||$$

called the operator norm of T. This satisfies:

i) $||S+T|| \leq ||S|| + ||T||$
ii) $||\alpha T|| = |\alpha| ||T||$
iii) $||Tx|| \leq ||T|| \, ||x||$
iv) $||ST|| \leq ||S|| \, ||T||$
v) $||T|| \geq 0$ with equality if and only if $T = 0$

Proof. eg. iii) $||Tx|| = ||x|| \left|\left| T \frac{x}{||x||} \right|\right| \leq ||x|| \, ||T||$ since $\frac{x}{||x||}$ a unit vector.

iv)

$$\begin{aligned} ||ST|| &= \sup_{||u||=1} ||STu|| \\ &\leq \sup_{||u||=1} ||S|| \, ||Tu|| \\ &\leq \sup_{||u||=1} ||S|| \, ||T|| \, ||u|| \\ &= ||S|| \, ||T|| \end{aligned}$$

using iii). □

Theorem 2.3.
(Chain Rule For Functions on Finite Dimensional Real or Complex Vector Space)

Let $U \xrightarrow{g} V \xrightarrow{f} W$ *and* $U \xrightarrow{f \cdot g} W$ *where* U, V, W *are open subsets of finite dimensional real or complex vector spaces. Let* g *be differentiable at* a*,* f *differentiable at* $g(a)$*, then* $f \cdot g$ *is differentiable at* a *and*

$$\Big(f \cdot g\Big)' = f'\Big(g(a)\Big) g'(a)$$

Proof.

$$\begin{aligned} f\Big(g(a+h)\Big) &= f\Big[g(a) + g'(a)h + \phi(h)\Big] && \text{where } \frac{||\phi(h)||}{||h||} \to 0 \text{ as } ||h|| \to 0 \\ &= f\Big[g(a) + y\Big] && \text{where } y = g'(a)h + \phi(h) \\ & && \Rightarrow ||y|| \leq ||g'(a)|| \, ||h|| + ||\phi(h)|| \\ & && \Rightarrow \frac{||y||}{||h||} \leq ||g'(a)|| + \frac{||\phi(h)||}{||h||} \\ &= f\Big(g(a)\Big) + f'\Big(g(a)\Big) y + \psi(y) && \text{where } \frac{||\psi(y)||}{||y||} \to 0 \text{ as } ||y|| \to 0 \\ &= f\Big(g(a)\Big) + f'\Big(g(a)\Big) g'(a)h \\ &\quad + f'\Big(g(a)\Big)\phi(h) + ||y|| \, \theta(y) && \text{where } \theta(y) = \begin{cases} \frac{\psi(y)}{||y||} & y \neq 0 \\ 0 & y = 0 \end{cases} \end{aligned}$$

Now,

$$\frac{\left|\left| f'\Big(g(a)\Big)\phi(h) + ||y||\theta(y) \right|\right|}{||h||} \leq \left|\left| f'\Big(g(a)\Big) \right|\right| \frac{||\phi(h)||}{||h||} + \frac{||y||}{||h||} ||\theta(y)||$$

which $\to 0$ as $||h|| \to 0$ (since $\theta(y) \to 0$ as $||y|| \to 0$ as $||h|| \to 0$), hence result. □

Example. Consider $\frac{d}{dt} f\Big[g^1(t), \dots, g^n(t)\Big]$, where

$$\begin{aligned} \mathbb{R} \supset U &\xrightarrow{g} V \subset \mathbb{R}^n \\ t &\mapsto g(t) = \Big(g^1(t), \dots, g^n(t)\Big) \end{aligned}$$

with g differentiable on U, and

$$\begin{aligned} \mathbb{R}^n \supset V &\xrightarrow{f} \mathbb{R} \\ (x_1, \dots, x_n) &\mapsto f(x_1, \dots, x_n) \end{aligned}$$

so that $\mathbb{R} \supset U \xrightarrow{f \cdot g} \mathbb{R}$. Then,

$$\begin{aligned} \frac{d}{dt} f\Big[g^1(t), \dots, g^n(t)\Big] &= \frac{d}{dt}\Big(f \cdot g\Big)(t) \\ &= \Big(f \cdot g\Big)'(t) && (1 \times 1 \text{ matrix}) \\ &= f'\Big(g(t)\Big) g'(t) && (n \times 1 \text{ and } 1 \times n \text{ matrices}) \\ &= \left(\frac{\partial f(g(t))}{\partial x^i}\right)\left(\frac{d}{dt} g^i(t)\right) \\ &= \sum_{i=1} n \frac{\partial f\Big(g(t)\Big)}{\partial x^i} \frac{d}{dt} g^i(t) \\ &= \frac{\partial f(g(t))}{\partial x^1} \frac{dg^1(t)}{dt} + \\ &\quad \cdots + \frac{\partial f(g(t))}{\partial x^n} \frac{dg^n(t)}{dt} \end{aligned}$$

illustrating the "chain rule."

Example. Consider $f(x, y)$, then

$$\frac{d}{dt} f\Big[g(t), g(t)\Big] = \frac{\partial f}{\partial x^1}\Big[g(t), g(t)\Big] \frac{dg(t)}{dt} + \frac{\partial f}{\partial x^2}\Big[g(t), g(t)\Big] \frac{dg(t)}{dt}$$

Definition: (C^r) If f differentiable and f' continuous, we say that f is C^1. If f' differentiable and $f^{(2)} = (f')'$ continuous, we say that f is C^2. If the r^{th} derivative exists and is continuous, we say that f is C^r. If the r^{th}derivative $f^{(r)}$ exists for all r we say that f is C^∞ or smooth.

For each r we have a category whose objects are open subsets of finite dimensional real or complex vector spaces, and whose morphisms are C^r functions, such that $f, g\ C^r \Rightarrow f \cdot g\ C^r$. The isomorphisms in this category are called C^r diffeomorphisms. V, W are C^r diffeomorphic if $\exists V \xrightarrow{f} W, V \xleftarrow{f^{-1}} W,\ f, f^{-1}$ both C^r.

Now consider $\mathbb{R}^n \supset V \xrightarrow{f} W \subset \mathbb{R}^m$. We have that

$$f(x_1, \dots, x_n) = \Big[f^1(x_1, \dots, x_n), \dots, f^m(x_1, \dots, x_n)\Big]$$

and then $f\ C^1 \Rightarrow f' = \left(\frac{\partial f^i}{\partial x^j}\right)$ ($m \times n$ matrix) exists and is continuous, which implies $\frac{\partial f^i}{\partial x^j}$ all exist and are continuous. We want to show the converse.

Theorem 2.4.
(Just the one component case) Let $\mathbb{R}^n \supset V \xrightarrow{f} \mathbb{R}$*, V open, then f is C^1* $\Leftrightarrow$ $\frac{\partial f}{\partial x^i}$ *exists and is continuous for all i.*

Proof. Have already proved $\Rightarrow$ implication. We will prove the $\Leftarrow$ case for $n = 2$ only, i.e. let

$$\begin{aligned} \mathbb{R}^2 \supset V &\xrightarrow{f} \mathbb{R} \\ (x, y) &\mapsto f(x, y) \end{aligned}$$

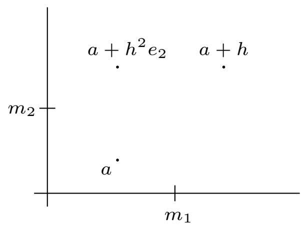
then $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$ exist and are continuous. To prove this, let $a \in V$, then

$$f(a + h) = f(a) + \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right] \begin{bmatrix} h^1 \\ h^2 \end{bmatrix} + \phi(h)$$

where the remainder is

$$\begin{aligned} \phi(h) &= f(a + h) - f(a + h^2 e_2) - \frac{\partial f}{\partial x} h^1 + f(a + h^2 e_2) - f(a) - \frac{\partial f}{\partial y} h^2 \\ &= \frac{\partial f}{\partial x}(m_1) h^1 - \frac{\partial f}{\partial x}(a) h^1 + \frac{\partial f}{\partial y}(m_2) h^2 - \frac{\partial f}{\partial y}(a) h^2 \end{aligned}$$

using the Mean Value Theorem, for some m_1 on $[a + h^2 e_2, a + h]$ and some m_2 on $[a, a + h^2 e_2]$.



Hence,

$$\begin{aligned} \frac{||\phi(h)||}{||h||} \le \frac{||h^1||}{||h||} &\left|\left|\frac{\partial f}{\partial x}(m_1) - \frac{\partial f}{\partial x}(a)\right|\right| \\ + \frac{||h^2||}{||h||} &\left|\left|\frac{\partial f}{\partial y}(m_2) - \frac{\partial f}{\partial y}(a)\right|\right| \end{aligned}$$

which $\to 0$ as $||h|| \to 0$. Therefore f' exists, and is $\left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right]$. $\square$

Now let

$$\begin{aligned} M \supset V &\xrightarrow{f} N_1 \times \cdots \times N_m \\ f(x) &= \Big(f_1(x), \ldots, f_m(x)\Big) \\ &= \Big(f_j(x)\Big) \\ &\text{(}1 \times m \text{ matrix)} \end{aligned}$$

be an m component function on open subset V, where M, N_i are finite dimensional real or complex vector spaces. For any norm on M, N_i, take a norm on $N_1 \times \cdots \times N_m$ as

$$||(y_1, \ldots, y_m)|| = ||y_1|| + \cdots + ||y_m||$$

Then if $a \in V$,

$$\begin{aligned} f(a + h) &= \Big(f_j(a + h)\Big) \\ &= \Big(f_j(a) + f_j'(a)h + \psi_j(h)\Big) \\ &= \Big(f_j(a)\Big) + \Big(f_j'(a)h\Big) + \Big(\psi_j(h)\Big) \end{aligned}$$

so

$$\frac{||\psi(h)||}{||h||} = \frac{||\psi_1(h)||}{||h||} + \cdots + \frac{||\psi_m(h)||}{||h||}$$

which goes to zero as $||h|| \to 0||$ if and only if $\frac{||\psi_j(h)||}{||h||} \to 0$ as $||h|| \to 0$ for all $j = 1, \ldots, m$.

Therefore f is differentiable at a with derivative $f'(a)$ such that

$$f'(a)h = \left(f_j'(a)h\right)$$

if and only if f_j is differentiable at a with derivative $f_j'(a)$ for all $j = 1, \ldots, m$. We have that f' continuous $\Leftrightarrow f_j'$ continuous $\forall j$, and so f is $C^1 \Leftrightarrow f_j$ is $C^1 \forall j$.

Example. Let $\mathbb{R}^n \supset V \xrightarrow{f} \mathbb{R}$. Then,

$$\begin{aligned}
&f \text{ is } C^1 \Leftrightarrow f' = \left(\tfrac{\partial f}{\partial x^j}\right) \text{ exists and is continuous}\\
&f \text{ is } C^2 \Leftrightarrow f'' = \left(\tfrac{\partial^2 f}{\partial x^i \partial x^j}\right) \text{ exists and is continuous}\\
&f \text{ is } C^3 \Leftrightarrow f''' = \left(\tfrac{\partial^3 f}{\partial x^i \partial x^j \partial x^k}\right) \text{ exists and is continuous}\\
&f \text{ is } C^\infty \Leftrightarrow \text{all partial derivatives exist.}
\end{aligned}$$

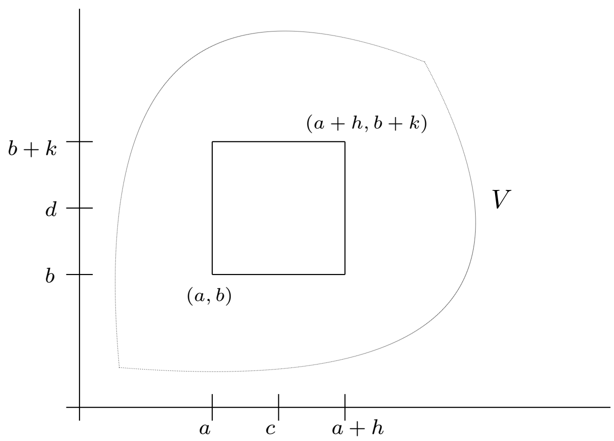
Theorem 2.5.
If $\mathbb{R}^n \supset V \xrightarrow{f} \mathbb{R}$, V open, is C^2, then

$$\frac{\partial^2 f}{\partial x^i \partial x^j} = \frac{\partial^2 f}{\partial x^j \partial x^i}$$

Proof. We will prove the case $n = 2$ only, i.e. let $\mathbb{R}^2 \supset V \xrightarrow{f} \mathbb{R}$, then

$$\frac{\partial}{\partial x}\left(\frac{\partial}{\partial y}\right) = \frac{\partial}{\partial y}\left(\frac{\partial}{\partial x}\right)$$

Let $(a, b) \in V$. Let $h \neq 0, k \neq 0$ be such that the closed rectangle $(a, b), (a+h, b), (a+h, b+h), (a, b+k)$ is contained in V.



Put

$$g(x) = f(x, b+k) - f(x, b)$$

then

$$\begin{aligned}
f(a+h,b+h)-f(a+h,b)-f(a,b+k)+f(a,b) &= g(a+h)-g(a) \\
&= hg'(c) \\
&= h\left[\frac{\partial f}{\partial x}(c,b+k)-\frac{\partial f}{\partial x}(c,b)\right] \\
&= hk\left[\frac{\partial}{\partial y}\frac{\partial f}{\partial x}\right](c,d)
\end{aligned}$$

using the Mean Value Theorem first for some $a \leq c \leq a+h$ and then for some $b \leq d \leq b+k$. Similarly, by defining

$$\tilde{g}(x) = f(a+h,y)-f(a,y)$$

we have

$$\begin{aligned}
f(a+h,b+h)-f(a+h,b)-f(a,b+k)+f(a,b) &= \tilde{g}(b+k)-\tilde{g}(b) \\
&= k\,\tilde{g}'(d') \\
&= k\left[\frac{\partial f}{\partial y}(a+h,d')-\frac{\partial f}{\partial y}(a,d')\right] \\
&= hk\left[\frac{\partial}{\partial x}\frac{\partial f}{\partial y}\right](c',d')
\end{aligned}$$

say, and so cancelling the h and k we get

$$\frac{\partial^2 f}{\partial y \partial x}(c,d) = \frac{\partial^2 f}{\partial x \partial y}(c',d')$$

We now let $(h,k) \to (0,0)$, so $(c,d) \to (a,b)$ and $(c',d') \to (a,b)$. Hence by continuity of the second derivative we find that

$$\frac{\partial^2 f}{\partial x \partial y}(a,b) = \frac{\partial^2 f}{\partial y \partial x}(a,b)$$

□

Theorem 2.6.
(Mean Value Theorem for Functions on Finite Dimensional Normed Spaces) Let $M \supset V \xrightarrow{f} N$ *be* C^1*. Let* $x, y \in V$ *such that*

$$[x,y] = \{tx+(1-t)y \,|\, 0 \leq t \leq 1\} \subset V$$

Let

$$\Big|\Big| f'\left[tx+(1-t)y\right] \Big|\Big| \leq k \quad \forall 0 \leq t \leq 1$$

then

$$\Big|\Big| f(x)-f(y) \Big|\Big| \leq k \Big|\Big| x-y \Big|\Big|$$

Proof. We have that

$$\begin{aligned} f(x) - f(y) &= \int_0^1 \left(\frac{d}{dt} f\left[tx + (1-t)y\right] \right) dt \\ &= \int_0^1 f'\left[tx + (1-t)y\right](x-y)dt \\ \Rightarrow \left|\left|f(x) - f(y)\right|\right| &\leq \int_0^1 k\left|\left|x-y\right|\right|dt \\ &= k\left|\left|x-y\right|\right| \end{aligned}$$

as required. □

Theorem 2.7.
(Inverse Function Theorem) Let $M \supset V \xrightarrow{f} N$ be a C^r function on open V, with M, N finite dimensional real or complex vector spaces. Let $a \in V$ at which

$$M \xrightarrow{f'(a)} N$$

is invertible, then there exists an open neighbourhood W of a such that

$$W \xrightarrow{f} f(W)$$

is a C^r diffeomorphism onto open $f(W)$ in N.

Proof. i) Let T be the inverse of $f'(a)$, and put

$$F(x) = Tf(x+a) - Tf(a)$$

so

$$\begin{aligned} F(0) &= Tf(a) - Tf(a) = 0 \\ F'(0) &= Tf'(a) + 0 = 1 \end{aligned}$$

We will prove that F maps an open neighbourhood U of 0 onto an open neighbourhood $F(U)$ of 0 by a C^r diffeomorphism. It will then follow that f maps $U + a$ onto $f(U+a)$ by a C^r diffeomorphism.

Choose a closed ball B, centre 0, radius $r > 0$ such that, by continuity of F'

$$\left.\begin{aligned} &||1_M - F'(x)|| \leq \tfrac{1}{2} \\ &\det F'(x) \neq 0 \end{aligned}\right\} \forall x \in B$$

then, using $||u|| - ||v|| \leq ||u - v||$,

$$\begin{aligned} \forall x, y \in B, ||x-y|| - ||F(x) - F(y)|| &\leq ||(1-F)x - (1-F)y|| \\ &\leq \frac{1}{2}||x-y|| \end{aligned}$$

by the Mean Value Theorem. Therefore

$$||F(x) - F(y)|| \geq \frac{1}{2}||x-y||$$

so $F(x) = F(y) \Rightarrow x = y$, so F is injective on B. Also, $F(x) \to F(y) \Rightarrow x \to y$, so F^{-1} is continuous. Hence

$$B \xrightarrow{F} F(B)$$

is a homeomorphism.

ii) We will now show $\frac{1}{2}B \subset F(B)$. Take $a \in \frac{1}{2}B$. Put

$$g(x) = x - F(x) + a$$

then

$$g'(x) = 1 - F'(x) + 0$$

so

$$||g'(x)|| \leq \frac{1}{2} \ \forall x \in B$$

and therefore

$$||g(x) - g(y)|| \leq \frac{1}{2}||x - y|| \ \forall x, y \in B$$

by the Mean Value Theorem, and so g is a contraction. Also,

$$\begin{aligned} ||g(x)|| &= ||g(x) - g(0) + a|| && \text{since } g(0) = a \\ &\leq ||g(x) - g(0)|| + ||a|| \\ &\leq \frac{1}{2}||x|| + \frac{1}{2}r && \text{by MVT} \\ &\leq \frac{1}{2}r + \frac{1}{2}r \\ &= r \ \ \forall x \in B \end{aligned}$$

Therefore $x \in B \Rightarrow g(x) \in B$, so $B \xrightarrow{g} B$, and is a contraction.

Consider now a sequence of points $x_0, x_1, x_2 \dots$ in B where $g(x_0) = x_1$, $g(x_1) = x_2$ and so on, or $x_r = g^r(x_0)$. Let

$$\lim_{r \to \infty} x_r = z \in B$$

$$\Rightarrow g(z) = g \lim_{r \to \infty} x_r = \lim_{r \to \infty} g(x_r) = \lim_{r \to \infty} x_{r+1} = z$$

hence z is a fixed point. Therefore

$$z = z - F(z) + a$$

$$\Rightarrow F(z) = a$$

so $a \in F(B)$.

iii) Now let B^0 be the interior of B, and let $U = B^0 \cap F^{-1}\left(\frac{1}{2}B^0\right)$, all open sets, then $U \xrightarrow{F} F(U)$ is a C^r homeomorphism of open U onto open $F(U)$.

We want to show that the inverse map $U \xleftarrow{G} F(U)$ is C^r. Let $x, x+h \in F(U)$, then $G(x) = y$, $G(x+h) = y + l$, say. Then

$$F(y + l) = F(y) + Sl + \phi(l)$$

where $S = F'(y)$ and $\frac{||\phi(l)||}{||l||} \to 0$ as $||l|| \to 0$, and $||h|| \geq ||l||$ (as $||F(x) - F(y)|| \geq \frac{1}{2}||x - y||$), so $||l|| \to 0$ as $||h|| \to 0$. Therefore,

$$\begin{aligned} x + h &= x + Sl + \phi(l) \\ \Rightarrow l &= S^{-1}h - S^{-1}\phi(l) \end{aligned}$$

so

$$G(x + h) = y + l = G(x) + S^{-1}h - S^{-1}\phi(l)$$

and

$$\frac{||S^{-1}\phi(l)||}{||h||} \leq ||S^{-1}|| \frac{||\phi(l)||}{||l||} \frac{||l||}{||h||}$$

with $\frac{||l||}{||h||} \leq 2$, so this all goes to zero as $||h|| \to 0$, hence G is differentiable at x and

$$G'(x) = S^{-1} = \Big[F'(y)\Big]^{-1} = \Big[F'(G(x))\Big]^{-1}$$

therefore if G is C^s for some $0 \leq s \leq r$, then G' is the composition of the C^s functions G, F' and $[\,\cdot\,]^{-1}$, and hence is $C^s \Rightarrow G$ is C^{s+1}. So G $C^s \Rightarrow G$ C^{s+1} $\forall\, 0 \leq s \leq r$, therefore G is C^r as required. □

3 Exercises

1. Let f be a constant function. Show $f'(a)$ is zero $\forall\, a \in M$.

   Solution: Let $f(x) = c$ $\forall\, x \in M$. Then

   $$f(a + h) = c = f(a) + 0h + 0$$

   and therefore $f'(a) = 0$ $\forall\, x \in M$

2. Let f be a linear function. Show that $f'(a)$ is equal to f $\forall\, x \in M$. Show $f''(a) = 0$ $\forall\, x \in M$.

   Solution: $f(a + h) = f(a) + f(h) + 0$, so therefore $f'(a) = f$ $\forall\, x \in M \Rightarrow f' =$ constant $\Rightarrow f''(a) = 0$ $\forall\, x \in M$.

3. Let $f : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ be defined by $f(A) = A^2$. Prove that f is differentiable and find its derivative.

   Solution: $f(A + H) = A^2 + AH + HA + H^2$ and

   $$\frac{||H^2||}{||H||} = \frac{||H\,H||}{||H||} \leq \frac{||H||\,||H||}{||H||} \to 0 \text{ as } ||H|| \to 0$$

   hence f is differentiable, with derivative given by

   $$f'(A)H = AH + HA$$

4. Let B be a fixed $n\times n$ matrix. Let $f : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ be defined by $f(A) = ABA$. Prove that f is differentiable and find its derivative.

   Solution: $f(A + H) = (A + H)B(A + H) = ABA + ABH + HBA + HBH$ and

   $$\frac{||HBH||}{||H||} \leq \frac{||H||||B||||H||}{||H||} \to 0 \text{ as } ||H|| \to 0$$

hence f is differentiable, with derivative given by

$$f'(A)H = ABH + HBA$$

5. Let $f : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ be defined by $f(A) = A^t A$. Prove that f is differentiable and find its derivative. Prove that if A is an orthogonal matrix then the image of the linear operator $f'(A)$ is the space of real symmetric $n\times n$ matrices.

   Solution: $f(A+H) = (A+H)^t(A+H) = A^tA + A^tH + H^tA + H^tH$ and

$$\frac{||H^tH||}{||H||} \leq \frac{||H^t||||H||}{||H||} \to 0 \text{ as } ||H|| \to 0$$

   hence f is differentiable, with derivative given by

$$f'(A)H = A^tH + H^tA$$

   If A orthogonal, we need to show that there exists $H \in \mathbb{R}^{n\times n}$ such that $f'(A)H = S$, ie such that $A^tH + H^tA = S$ and such a matrix H is given by $H = \frac{1}{2}AS$.

6. Let V be the space of real non-singular $n\times n$ matrices. Let $f : V \to \mathbb{R}^{n\times n}$ be given by $f(A) = A^{-1}$. Find $f'(A)$.

   Solution: We want $f(A+H) = (A+H)^{-1} = A^{-1}+$ linear in $H+$ remainder. Now,

$$(A+H)^{-1} - A^{-1} = (A+H)^{-1}\Big[\mathbb{I} - (A+H)A^{-1}\Big] = (A+H)^{-1}\Big(- HA^{-1}\Big) \approx -A^{-1}HA^{-1}$$

   for small H. The remainder term is then

$$\begin{aligned}\phi(H) &= (A+H)^{-1} - A^{-1} + A^{-1}HA^{-1} \\ &= (A+H)^{-1}\Big[\mathbb{I} - (A+H)A^{-1} + (A+H)A^{-1}HA^{-1}\Big] \\ &= (A+H)^{-1}\Big[HA^{-1}HA^{-1}\Big]\end{aligned}$$

   then

$$\frac{||\phi(H)||}{||H||} \leq \big|\big|(A+H)^{-1}\big|\big| \frac{||H||^2}{||H||} \big|\big|A^{-1}\big|\big|^2 \to 0 \text{ as } ||H|| \to 0$$

   as required, hence

$$f'(A)H = -A^{-1}HA^{-1}$$

7. A function F of n real variables is called homogeneous of degree r if it satisfies

$$F(tx_1, \ldots, tx_n) = t^r F(x_1, \ldots, x_n)$$

   By differentiation with respect to t show that such a function F is an eigenfunction of the operator

$$x_1\frac{\partial}{\partial x^1} + \cdots + x_n\frac{\partial}{\partial x^n}$$

   and find the eigenvalue.

   Solution: Fix $x_1, \ldots, x_n$ and differentiate with respect to t, hence

$$x^1\frac{\partial}{\partial x^1}F(tx_1, \ldots, tx_n) + \cdots + x^n\frac{\partial}{\partial x^n}F(tx_1, \ldots, tx_n) = rt^{r-1}F(x_1, \ldots, x_n)$$

Let $t = 1$, then

$$\left(x_1 \frac{\partial}{\partial x^1} + \cdots + x_n \frac{\partial}{\partial x^n}\right) F = rF$$

hence F an eigenfunction of the the operator $x_1 \frac{\partial}{\partial x^1} + \cdots + x_n \frac{\partial}{\partial x^n}$ with eigenvalue r.

8. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be the function

$$f = \begin{cases} 2xy\frac{x^2-y^2}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = 0 \end{cases}$$

Show that

$$\frac{\partial^2 f}{\partial x \partial y} \neq \frac{\partial^2 f}{\partial y \partial x}$$

at $(0,0)$.

Solution: We have that

$$\left(\frac{\partial^2 f}{\partial x \partial y}\right)_{(0,0)} = \left[\frac{\partial}{\partial x}\frac{\partial f}{\partial y}\right]_{(0,0)} = \lim_{t \to 0}\left[\frac{\frac{\partial f}{\partial y}(t,0) - \frac{\partial f}{\partial y}(0,0)}{t}\right]$$

and away from the origin

$$\frac{\partial f}{\partial y} = \frac{\partial}{\partial y}\left[\frac{2x^3y - 2xy^3}{x^2+y^2}\right] = \frac{(x^2+y^2)(2x^3 - 6xy^2) - (2x^3y - 2xy^3)(2y)}{(x^2+y^2)^2}$$

$$\Rightarrow \frac{\partial f}{\partial y}(t,0) = 2t$$

and as $f(x,y) = 0$ at $(0,0)$ and along the y-axis, then $\frac{\partial f}{\partial y}(0,0) = 0$. Hence

$$\left(\frac{\partial^2 f}{\partial x \partial y}\right)_{(0,0)} = \lim_{t\to 0}\frac{2t}{t} = 2$$

By symmetry, $\left(\frac{\partial^2 f}{\partial y \partial x}\right)_{(0,0)} = -2$, hence

$$\frac{\partial^2 f}{\partial x \partial y} \neq \frac{\partial^2 f}{\partial y \partial x}$$

at $(0,0)$.

9. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be the function

$$f = \begin{cases} \frac{2xy^2}{x^2+y^4} & (x,y) \neq (0,0) \\ 0 & (x,y) = 0 \end{cases}$$

Calculate $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ at $(0,0)$. Calculate $\frac{d}{dt}f(ta, tb)$ at $t = 0$ for $a, b \in \mathbb{R}$. Deduce that the chain rule does not hold at $(0,0)$.

Solution: As $f(x,y)=0$ along the x and y axes and at the origin, we have

$$\frac{\partial f}{\partial x}(0,0)=\frac{\partial f}{\partial y}(0,0)=0$$

We also have $\frac{d}{dt}f(ta,tb)=a\frac{\partial f}{\partial x}(ta,tb)+b\frac{\partial f}{\partial y}(ta,tb)$.
Now, away from the origin f is C^∞, so the chain rule holds and

$$\begin{aligned}\left.\frac{d}{dt}f(ta,tb)\right|_{t=0} &= \lim_{t\to 0}\frac{f(ta,tb)-f(0,0)}{t}\\ &= \lim_{t\to 0}\frac{2ab^2t^3}{t(a^2t^2+b^4t^4)}\\ &= \begin{cases}\frac{2b^2}{a} & a\neq 0\\ 0 & a=0\end{cases}\end{aligned}$$

while the chain rule would give

$$\left.\frac{d}{dt}f(ta,tb)\right|_{t=0}=a\frac{\partial f}{\partial x}(0,0)+b\frac{\partial f}{\partial y}(0,0)=0$$

so the chain rule does not hold at $(0,0)$ and indeed f is not differentiable at $(0,0)$.