
Contents

1	Multivariable Calculus	3
1.1	Linear Algebra	3
1.2	Derivative	12

Chapter 1

Multivariable Calculus

1.1 Linear Algebra

1.1.1 Vector Spaces

Definition (Normed vector space): Let V be a vector space. A **norm** is a real valued function $\| \cdot \| : V \rightarrow \mathbb{R}$ which has the following properties

1. $\forall x \in V, \|x\| > 0$.
2. $\|x\| = 0 \implies x = 0$.
3. $\forall x \in V \forall \alpha \in \mathbb{F}, \|\alpha x\| = |\alpha| \|x\|$.
4. $\forall x, y \in V \quad \|x + y\| \leq \|x\| + \|y\|$.

Each normed vector space induces a metric space (V, d) where $d(x, y) = \|y - x\|$.

Theorem 1.1. *In every normed space $(V, \| \cdot \|)$ we have*

$$|||v| - |w|| \leq \|v - w\|$$

Hence the norm is Lipschitz continuous.

Definition: Assume V is a vector space and let $\| \cdot \|_1, \| \cdot \|_2$ be two norms for V . They are said to be equivalent when

$$\exists c_1, c_2 > 0 \forall x : \quad c_1 \|x\|_1 \leq \|x\|_2 \leq c_2 \|x\|_1$$

To check if the above definition is indeed an equivalence relation, we must show that following:

Reflexive $\| \cdot \|_1 \sim \| \cdot \|_1$.

Symmetric $\| \cdot \|_1 \sim \| \cdot \|_2 \implies \| \cdot \|_2 \sim \| \cdot \|_1$.

Transitive $\| \cdot \|_1 \sim \| \cdot \|_2, \| \cdot \|_2 \sim \| \cdot \|_3 \implies \| \cdot \|_1 \sim \| \cdot \|_3$.

Remark 1. Equivalent norms induce equivalent metrics, hence they induce the same topology.

Theorem 1.2. *All norms defined on a finite dimensional vector space V are equivalent.*

Proof. Let $\|\cdot\|$ be an arbitrary norm on V and $\{e_1, e_2, \dots, e_n\}$ be a basis of V . Let $\|\cdot\|_2$ be L_2 -norm (Euclidean norm). It will suffice to show $\|\cdot\| \sim \|\cdot\|_2$. Let

$$M = \max(\|e_1\|, \dots, \|e_n\|)$$

Take $x \in V$, writing $x = \sum_{i=1}^n \xi_i e_i$ we have:

$$\|x\| = \left\| \sum_{i=1}^n \xi_i e_i \right\| \leq \sum_{i=1}^n |\xi_i| \|e_i\| \leq M \sqrt{n} \|x\|_2$$

Taking $c_2 = M\sqrt{n}$ proves the right inequality. For the left inequality we need the following lemma

Lemma 1.3. *If V is a normed vector space with $\|\cdot\|_2$, as defined above, is viewed as metric space $(V, \|\cdot\|_2)$ then $\|\cdot\| : V \rightarrow \mathbb{R}$ is continuous.*

Proof. Let $x_0 \in V$ and M be defined as above. For any $\epsilon > 0$ consider $\delta = \frac{\epsilon}{M\sqrt{n}}$ then if $\|x - x_0\|_2 < \delta$

$$\left| \|x\| - \|x_0\| \right| \leq \|x - x_0\| \leq M\sqrt{n} \|x - x_0\|_2 \leq \epsilon$$

Now consider the sphere of radius $r = 1$ centered at 0, $S_1(0) = S_1 = \{x \in V : \|x\|_2 = 1\}$. One can show that S is compact (Theorem 1.4). Therefore, $\|x\|$ assumes its minimum on S . Let $a = \|x_0\|$ be the minimum. Since $0 \notin S$ then $a > 0$. By letting $y = x/\|x\|_2$, we have $y \in S$ and thus $a \leq \|y\|$ which is

$$a \|x\|_2 \leq \|x\|$$

Taking $c_1 = a$ proves the theorem. ■

Theorem 1.4. *Let $(V, \|\cdot\|)$ be a normed space over a normed complete field \mathbb{F} . The following are equivalent*

1. V is finite dimensional.
2. every bounded closed set in V is compact.
3. the closed unit ball in V is compact.

Proof. Item 1 \implies Item 2: It is similar to proving a closed set \mathbb{R}^n is compact using the fact a closed interval is compact in \mathbb{R} .

Item 2 \implies Item 3: Trivial.

Item 3 \implies Item 1: Requires the following lemma:

Lemma 1.5 (Riesz's lemma). *If V is a normed vector space and W is a closed proper subspace of V and $\alpha \in \mathbb{R}$ with $0 < \alpha < 1$, then there exists an $v \in V$ with $\|v\| = 1$ such that $\|v - w\| \geq \alpha$ for all $w \in W$*

Now suppose V were to be an infinite dimensional vector space. Then by the Lemma 1.5 there is sequence of unit vectors x_n such that $\forall m, n \in \mathbb{N}$, $\|x_n - x_m\| > \alpha$ for some $0 < \alpha < 1$. Which implies that no subsequence of $\{x_n\}$ is convergent and hence the closed unit ball can not be compact. ■

Example 1.1. The closed unit ball in the infinite dimensional vector space $C([0, 1], \mathbb{R})$ with $\|f\| = \max f(x)$ is not compact. Take $f_n(x) = x^n$. Obviously $\|f_n\| = 1$, however f_n doesn't uniformly converge and hence f_n doesn't have a limit in $C([0, 1], \mathbb{R})$ with the max norm. Consider the following norm

$$\|f\|_I = \int_0^1 |f(x)| dx$$

Note that $\|\cdot\|_I$ and $\|\cdot\|_{\max}$ are not equivalent. Let $g(x) = 0$ for all $x \in [0, 1]$. Then

$$\|f_n - g\|_I = \frac{1}{n+1} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Definition (Banach space): A normed vector space V that is complete is a **Banach space**. A **Hilbert Space** is a Banach space whose norm is induced by an inner product.

Proposition 1.6. A normed finite dimensional vector space V over a normed complete field \mathbb{F} , is Banach space.

Proof. Let $\{v_i\} \in V$ be a Cauchy sequence, and $\{e_1, \dots, e_n\}$ be a basis for V with the norm L^1 , that is if $v = (\xi^1, \dots, \xi^n)$ then $\|v\| = \sum_{m=1}^n |\xi^m|$. Then if $v_i = (\xi_i^1, \dots, \xi_i^n)$

$$|\xi_i^m - \xi_j^m| \leq \sum_{m=1}^n |\xi_i^m - \xi_j^m| \leq \|v_i - v_j\| < \epsilon$$

then $\{\xi_i^m\}_i$ are a Cauchy sequence in \mathbb{F} and hence they converge $\xi_i^m \rightarrow \xi^m$. Then, clearly $v_i \rightarrow v = (\xi^1, \dots, \xi^n)$ as each component converges. ■

Example 1.2. \mathbb{Q} form a vector space itself over itself. It is finite dimensional as $\{1_{\mathbb{Q}}\}$ is the basis, however the sequence

$$1, 1.4, 1.41, \dots$$

does not converge even though it is Cauchy.

1.1.2 Linear Maps

Let V and W be a vector spaces over \mathbb{F} . A map $T : V \rightarrow W$ is **linear** if

$$T(x + \lambda y) = T(x) + \lambda T(y)$$

for all $x, y \in V$ and $\lambda \in \mathbb{F}$.

Definition: Let $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ be normed spaces then, a linear transformation $T : V \rightarrow W$ is **bounded** if there exists a constant $C > 0$ such that

$$\|Tv\|_W \leq C \|v\|_V$$

for all $v \in V$. We denote the set of all linear map from $V \rightarrow W$ as $\mathcal{L}(V, W)$ and the set of all bounded linear maps as $\mathcal{B}(V, W)$. If $T \in \mathcal{L}(V, W)$ is bijective such that $T^{-1} \in \mathcal{L}(V, W)$, then T is called an **isomorphism** and V, W are **isomorphic**. An operator $T \in \mathcal{L}(V, W)$ is called **isometric** if $\|Tv\|_W = \|v\|_V$ for all $v \in V$.

Definition: If $(V, \|\cdot\|_V)$, $(W, \|\cdot\|_W)$ are normed spaces then the **operator norm** of a linear transformation $T : V \rightarrow W$ is

$$\|T\| = \sup \left\{ \frac{\|Tv\|_W}{\|v\|_V} \mid v \neq 0 \right\}$$

Proposition 1.7. Let $T : U \rightarrow V$ and $T' : V \rightarrow W$ be two linear transformations.

$$\|T' \circ T\| \leq \|T\| \|T'\|$$

Proof. for an arbitrary non-zero $x \in U$

$$\|T' \circ T(x)\|_W \leq \|T'\| \|Tx\|_V \leq \|T'\| \|T\| \|x\|_U$$

which implies

$$\|T' \circ T\| \leq \|T\| \|T'\|$$

Theorem 1.8. Let $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ be normed spaces and $T : V \rightarrow W$ be a linear transformation. The following are equivalent

1. $\|T\|$ is finite.
2. T is bounded.
3. T is Lipschitz continuous.
4. T is continuous at a point.
5. $\sup_{\|v\|_V=1} \|Tv\|_W < \infty$.

Proof. item 1 \Rightarrow item 2: Obviously

$$\begin{aligned} \frac{\|Tv\|_W}{\|v\|_V} &\leq \|T\| \\ \Rightarrow \|Tv\|_W &\leq \|T\| \|v\|_V \end{aligned}$$

note that if $v = 0$ then $Tv = 0$ as well and thus the last inequality holds for all $v \in V$.

item 2 \Rightarrow item 3:

$$\|Tv - Tu\|_W = \|T(u - v)\|_W \leq C \|u - v\|_V$$

item 3 \Rightarrow item 4: Trivial.

item 4 \Rightarrow item 5: Let T be continuous at $u \in V$. Then there is a $\delta > 0$ such that

$$\|v - u\| < \delta \Rightarrow \|Tv - Tu\|_W = \|T(v - u)\|_W < 1$$

Now for an arbitrary non-zero v we have

$$\left\| \left(\frac{\delta v}{2\|v\|_V} + u \right) - u \right\|_V < \delta$$

Therefore

$$\begin{aligned} \left\| T\left(\frac{\delta v}{2\|v\|_V}\right) \right\|_W &< 1 \\ \left\| T\left(\frac{v}{\|v\|_V}\right) \right\|_W &< \frac{2}{\delta} \end{aligned}$$

item 5 \Rightarrow item 1: Let $v \in V$ be an arbitrary vector. Then

$$\begin{aligned} \sup \left\| T\left(\frac{v}{\|v\|_V}\right) \right\|_W &< \infty \\ \Rightarrow \sup \frac{\|Tv\|_W}{\|v\|_W} &< \infty \end{aligned}$$

Theorem 1.9. *If V is a finite dimensional normed vector space then any linear transformation $T : V \rightarrow W$ is continuous.*

Proof. Since V is finite dimensional, according to Theorem 1.2, any two norms are equivalent. Hence, take $\|\cdot\|_2$ to be Euclidean norm over a basis $\{e_1, \dots, e_n\}$. Let x be such that $\|x\|_2 < \delta$ for some $\delta > 0$. Therefore, $|\xi_i| < \delta^2$

$$\|Tx\|_W = \left\| \sum_{i=1}^n \xi_i T(e_i) \right\|_W \leq \sum_{i=1}^n |\xi_i| \|T(e_i)\|_W \leq \delta^2 K$$

where $K = \max \|T(e_i)\|_W$. By letting $\delta = \sqrt{\frac{\epsilon}{K}}$ we proved continuity at 0 and hence the continuity by Theorem 1.8. \blacksquare

Another proof of Proposition 1.6

Proof. Let $\{e_1, \dots, e_n\}$ be a basis for V and $\phi : V \rightarrow \mathbb{F}^n$ be the representation map for the basis. Since ϕ is a linear map and a bijection then ϕ is homeomorphism. Consider a Cauchy sequence $\{v_k\} \in V$ and let $x_k = \phi(v_k)$ then by continuity of ϕ and ϕ^{-1} we have

$$|x_i - x_j| = |\phi(v_i) - \phi(v_j)| \leq \|\phi\| \|v_i - v_j\| \leq \|\phi\| \|\phi^{-1}(x_i) - \phi^{-1}(x_j)\| \leq \|\phi\| \|\phi^{-1}\| |x_i - x_j|$$

hence $\{x_k\}$ are Cauchy in \mathbb{F}^n which by completeness of \mathbb{F} implies that they are convergent, $x_k \rightarrow x$. Let $v = \phi^{-1}(x)$ then by the right side of the inequality $v_k \rightarrow v$. \blacksquare

Remark 2. As seen in the last proof, for a bijective linear transformation T

$$1 \leq \|T\| \|T^{-1}\|$$

Theorem 1.10. *For two normed vector spaces V, W , $(\mathcal{B}(V, W), \|T\|)$ is a normed vector space. Moreover, it is a Banach space when W is a Banach space.*

Proof. Clearly $\mathcal{B}(V, W)$ is a vector space. For its norm $\|T\|$ we have

1. $\|T\| \geq 0$ by definition.

2. if $\alpha \in \mathbb{F}_W$ then

$$\|\alpha T\| = \sup \left\{ \frac{\|(\alpha T)v\|_W}{\|v\|_V} \mid v \neq 0 \right\} = |\alpha| \sup \left\{ \frac{\|Tv\|_W}{\|v\|_V} \mid v \neq 0 \right\} = |\alpha| \|T\|$$

3. for the triangle inequality

$$\begin{aligned} \|T_1 + T_2\| &= \sup \left\{ \frac{\|(T_1 + T_2)v\|_W}{\|v\|_V} \right\} \\ &\leq \sup \left\{ \frac{\|T_1v\|_W + \|T_2v\|_W}{\|v\|_V} \right\} \\ &= \sup \left\{ \frac{\|T_1v\|_W}{\|v\|_V} \right\} + \sup \left\{ \frac{\|T_2v\|_W}{\|v\|_V} \right\} \\ &= \|T_1\| + \|T_2\| \end{aligned}$$

Suppose W is a Banach space and $\{T_i\} \in \mathcal{B}(V, W)$ is a Cauchy sequence. Then for all $v \in V$

$$\forall \epsilon \exists N \text{ s.t. } m, n > N \implies \|T_m v - T_n v\|_W \leq \|T_m - T_n\| \|v\|_V < \epsilon$$

$\{T_i v\}$ is a Cauchy sequence. Since W is complete then $T_i v \rightarrow Tv$ for some function T . We claim that T is a bounded linear map and is the limit of $T_i \rightarrow T$.

$$\begin{aligned} T(v + cu) &= \lim_{i \rightarrow \infty} T_i(v + cu) = \lim_{i \rightarrow \infty} T_i v + cT_i u \\ &= Tv + cTu \end{aligned}$$

Note that $|\|T_m\| - \|T_n\|| \leq \|T_m - T_n\|$ and hence $\|T_i\|$ is a Cauchy in sequence in \mathbb{R} that has a limit t . There exists a N such that $|\|T_n\| - t| < 1$ for all $n \geq N$.

$$\frac{\|Tv\|_W}{\|v\|_V} = \lim_{i \rightarrow \infty} \frac{\|T_i v\|_W}{\|v\|_V} < t + 1$$

hence T is bounded and $T \in \mathcal{B}(V, W)$. Finally, we show that $T_i \rightarrow T$. For an arbitrary $v \neq 0$ and $\epsilon > 0$ there exist N such that

$$n \geq N \implies \|T_i v - Tv\|_W < \epsilon \|v\|_V$$

which means that

$$\|T_i - T\| = \sup \frac{\|T_i v - Tv\|_W}{\|v\|_V} < \epsilon$$

Therefore $T_i \rightarrow T$ as desired. ■

Theorem 1.11. *Let $(V, \|\cdot\|)$ be a normed space. Then any linear transformation $T : \mathbb{R}^n \rightarrow V$ is continuous. Furthermore, if T is a bijection, it is a homeomorphism.*

Proof. Since \mathbb{R}^n is finite then by Theorem 1.9, T is continuous. Assuming T is bijective, we must show that its inverse T^{-1} is continuous as well. Since T is a bijection then T is a linear isomorphism and $\dim V = \dim \mathbb{R}^n = n$ hence $T^{-1} : V \rightarrow \mathbb{R}^n$ is a continuous map. ■

Theorem 1.12. *Let V, W be two finite dimensional normed vector spaces. $T : V \rightarrow W$ linear transformation is invertible if and only if there exists a c such that:*

$$c\|v\|_V \leq \|Tv\|_W$$

Proof. If T is invertible then $T^{-1} : W \rightarrow V$ is bounded and thus

$$\|T^{-1}w\|_V \leq c \|w\|_W$$

and since T is bijective then there exists v such that $w = Tv$ which implies

$$\|y\|_V \leq c \|Ty\|_W$$

If there exists such c then $\|Tx\| > 0$ for all non-zero x and hence $\ker T = 0$ which implies that T is a bijection and is invertible. \blacksquare

Remark 3. the supremum of such c is $\|T^{-1}\|^{-1}$ which is called the **conorm** of T .

Definition (General linear group): The **general linear group** of a vector space, written $\text{GL}(V)$ is the set of all bijective linear transformation.

Proposition 1.13. *If V is a finite (also works for infinite) vector space then $\text{GL}(V)$ is open in $\mathcal{L}(V, V)$, in fact, if $f \in \text{GL}(V)$ then the open ball centered at f with radius $\|f^{-1}\|^{-1}$ remains in $\text{GL}(V)$. Furthermore, the inverse operator $i : \text{GL}(V) \rightarrow \text{GL}(V)$, $i(T) = T^{-1}$ is continuous.*

Proof. First assume $f = \mathbb{1}_V$ then we prove that any linear g that $\|\mathbb{1}_V - g\| < 1$ is invertible which then implies bijectivity (true for linear maps). Let $\|v\| = 1$ then

$$\|v\| - \|gv\| \leq \|v - gv\| \leq \|\mathbb{1}_V - g\| \|v\| < 1$$

Therefore

$$0 < \|gv\| < 2$$

which means $\ker g = \{0\}$ and since V is finite then g is invertible. For a general f , we have that

$$\|1 - f^{-1} \circ g\| \leq \|f^{-1}\| \|f - g\| < 1$$

therefore $f^{-1} \circ g$ is invertible and as a consequence $g = f \circ f^{-1} \circ g$ is invertible. To prove inverse operator is continuous, fix $\epsilon > 0$ then for a $\delta > 0$ if $\|T - S\| < \delta$ then

$$\begin{aligned} \|\mathbb{1}_V - T^{-1} \circ S\| &= \|T^{-1} \circ T - T^{-1} \circ S\| \leq \|T^{-1}\| \|T - S\| < \delta \|T^{-1}\| \\ \implies \|T^{-1} - S^{-1}\| &\leq \|T^{-1} \circ S - \mathbb{1}_V\| \|S^{-1}\| < \delta \|T^{-1}\| \|S^{-1}\| \end{aligned}$$

note that by letting $\delta = \|T^{-1}\|^{-1}/2$ then

$$\|S\| > -\frac{\|T^{-1}\|^{-1}}{2} + \|T\| > \frac{\|T^{-1}\|^{-1}}{2}$$

also if for any invertible linear map R

$$\|R\| > a \implies \|Rx\| > a\|x\| \implies \frac{\|y\|}{a} = \frac{\|R \circ R^{-1}(y)\|}{a} > \|R^{-1}y\|$$

which means that $\|S^{-1}\| < 2\|T^{-1}\|$, hence by letting

$$\delta = \min \left\{ \frac{\epsilon \|T^{-1}\|^2}{2}, \frac{\|T^{-1}\|^{-1}}{2} \right\}$$

we proved the continuity. \blacksquare

Definition: Let V_1, V_2, \dots, V_n be normed vector spaces. Then $\phi : V_1 \times \dots \times V_n \rightarrow W$ is **n -linear** if by fixing any $n - 1$ component, ϕ is linear relative to the remaining component.

Proposition 1.14. *If V_1, V_2, \dots, V_n are normed vector spaces and $\phi : V_1 \times \dots \times V_n \rightarrow W$ is a n -linear then the followings are equivalent*

1. ϕ is continuous.
2. ϕ is continuous at 0.
3. ϕ is bounded, that is there exists a constant $C > 0$ such that

$$\|\phi(v_1, \dots, v_n)\|_W \leq C \|v_1\|_{V_1} \dots \|v_n\|_{V_n}$$

Remark 4. As oppose to linear transformation, n -linear function's continuity does not imply uniform continuity.

Proof. Item 1 \implies Item 2: Trivial.

Item 2 \implies Item 3: For the sake of contradiction, suppose Item 3 is false. That is, for every $k \in \mathbb{N}$ there exists a point $v_k = (v_k^1, \dots, v_k^n)$ such that

$$\|\phi(v_k^1, \dots, v_k^n)\|_W > n^n \|v_k^1\|_{V_1} \dots \|v_k^n\|_{V_n}$$

Note that v_k^m can not be zero for any k and m , otherwise $\phi(v_k) = 0$. Define

$$w_k^m = \frac{v_k^m}{n \|v_k^m\|_{V_k}} \rightarrow 0$$

which from the continuity at 0 implies that $w_k = (w_k^1, \dots, w_k^n) \rightarrow 0$. However,

$$\|\phi(w_k) - \phi(0)\|_W > n^n \frac{1}{n} \dots \frac{1}{n} = 1$$

which is a contradiction.

Item 3 \implies Item 1. Let $v_n \rightarrow v$ and define the points

$$\bar{v}_k^m = (v^1, \dots, v^m, v_k^{m+1}, \dots, v_k^n), \quad \bar{v}_k^0 = v_k$$

and $\bar{v}_k^n = v$. Note that v_k^m are bounded for sufficiently large $k \geq N_1$, therefore there exists M such that $\forall m, \|v_k^m\|_{V_m} \leq M$. Also, pick M such that $\forall m, \|v_k^m\|_{V_m} \leq M$ as well. Then

$$\begin{aligned} \|\phi(v_k) - \phi(v)\|_W &\leq \sum_{i=1}^n \|\phi(\bar{v}_k^{i-1}) - \phi(\bar{v}_k^i)\|_W \\ &= \sum_{i=1}^n \|\phi(\bar{v}_k^{i-1} - \bar{v}_k^i)\|_W \\ &\leq \sum_{i=1}^n C \|v^1\|_{V_1} \dots \|v^{i-1}\|_{V_{i-1}} \|v_k^i - v^i\|_{V_i} \|v_k^{i+1}\|_{V_{i+1}} \dots \|v_k^n\|_{V_n} \\ &\leq CM^{n-1} \sum_{i=1}^n \|v_k^i - v^i\|_{V_i} \end{aligned}$$

pick N_2 such that for all $k \geq N_2$, for each i , $\|v_k^i - v^i\|_{V_i} < \frac{\epsilon}{nCM^{n-1}}$ then

$$\|\phi(v_k) - \phi(v)\|_W < CM^{n-1} \sum_{i=1}^n \frac{\epsilon}{nCM^{n-1}} = \epsilon$$

Exercises

1. Show that for a linear transformation T , $\|T\| = \sup_{\|v\|_V \leq 1} \|Tv\|_W$.

1.2 Derivative

Let V, W be finite dimensional vector spaces and $f : U \subset V \rightarrow W$ where U is open. Then f is differentiable at x_0 when a linear transformation $T : V \rightarrow W$ such that

$$\lim_{\|h\| \rightarrow 0} \frac{\|f(x_0 + h) - f(x_0) - T(h)\|}{\|h\|} = 0$$

Or equivalently there exists a sublinear function $R(h)$ such that

$$f(x_0 + h) - f(x_0) - Th = R(h) \quad \frac{R(h)}{\|h\|} \rightarrow 0$$

T if it exists is unique, represented by $f'(x_0)$, Df , or $df(x)$ and called the **total derivative** or **Fréchet derivative**.

Example 1.3. Any linear function $f : V \rightarrow W$ with $f(v) = Tv + b$ where $b \in W$ is differentiable and $Df(v) = T$. Since

$$\|h\|_V < \delta \implies \|f(v + h) - f(v) - (Df(v))(h)\|_W = \|T(v + h) - Tv - Th\|_W = 0 < \epsilon \|h\|_V$$

Hence, the derivative of any linear function is constant. Consider $S : V \times V \rightarrow V$ with $S(v, u) = v + u$. S is differentiable because S is linear (why?). We claim that $DS = S$ as

$$\|S((v + h), (u + k)) - S(v, u) - S(h, k)\| = 0$$

Example 1.4. Let $\mu : \mathbb{R} \times V \rightarrow V$ with $\mu(r, x) = rx$. Then μ is differentiable and $(D\mu(r, x))(t, h) = rh + tx$ as

$$\begin{aligned} \|\mu((r + t), (x + h)) - \mu(r, x) - (D\mu(r, x))(t, h)\| &= \|rx + rh + tx + th - rx - rh - tx\| \\ &= |t| \|h\| \leq \epsilon \|(t, h)\| \end{aligned}$$

by letting $\|(t, h)\| = \sqrt{t^2 + \|h\|^2}$ and $\delta = \epsilon$.

Proposition 1.15. *Differentiability of f at x implies continuity at x .*

Proof.

$$\|f(x + h) - f(x)\| = \|(Df(x))(h) + R(h)\| \leq \|Df(x)\| \|h\| + \|R(h)\| \rightarrow 0$$

as $h \rightarrow 0$. ■

Proposition 1.16. *Assume $f : U \subset V \rightarrow W$ is differentiable at x_0 and let $u \in V$ be a non-zero vector then*

$$f'(x_0)(u) = \lim_{t \rightarrow 0} \frac{f(x_0 + tu) - f(x_0)}{t}$$

Proof. Let $h = tu$ then

$$\begin{aligned} R(tu) &= f(x_0 + tu) - f(x_0) - T(tu) \\ &= f(x_0 + tu) - f(x_0) - tT(u) \\ \implies \frac{R(tu)}{t} &= \frac{f(x_0 + tu) - f(x_0)}{t} - T(u) \\ \implies \lim_{t \rightarrow 0} \frac{f(x_0 + tu) - f(x_0)}{t} &= T(u) \end{aligned}$$

Definition (Directional derivative): If we let $\|u\| = 1$ then the limit in Proposition 1.16 becomes the **directional derivative** of f in the direction of u and is denoted by $D_u f$.

Remark 5. The existence of all directional derivatives of f doesn't imply its differentiability or even its continuity.

Remark 6. If $Df : U \rightarrow \mathcal{L}(V, W)$ is continuous then each $\frac{\partial f_i}{\partial x_j}$ is continuous. Since

$$Df(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x) & \dots & \frac{\partial f_1}{\partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \dots & \frac{\partial f_m}{\partial x_n}(x) \end{bmatrix}$$

and the reverse is true as well.

Theorem 1.17. $f : V \rightarrow W$ has all of its partial derivative in a neighbourhood of $u \in U$ and they're continuous at u then f is differentiable at u . Especially, if $\frac{\partial f_i}{\partial x_j}$ exist and are continuous at every point of u then $f \in \mathcal{C}^1$.

Proof. We prove that each f_i is differentiable. Let $\{e_1, \dots, e_n\}$ be a basis for V and take $\|x\| = \sum |\xi_j|$. Consider a convex neighbourhood E of a . Then, for a given $\epsilon > 0$ we will show there exists a $\delta > 0$ such that

$$\|h\| < \delta \implies \left\| f_i(a+h) - f_i(a) - \sum_{j=1}^n (D_{e_j} f_i(a))(h_j) \right\| \leq \epsilon \|h\|$$

Consider the point sequence $a^k = \sum_{j < k} a_j e_j + \sum_{j \geq k} (a_j + h_j) e_j$ where $a^1 = a + h$ and $a^{n+1} = a$ then

$$\left\| f_i(a+h) - f_i(a) - \sum_{j=1}^n (D_{e_j} f_i(a))(h_j) \right\| \leq \sum_{k=1}^n \|f_i(a^k) - f_i(a^{k+1}) - (D_{e_k} f_i(a))(h_k)\|$$

hence we are done if

$$\|f_i(a^k) - f_i(a^{k+1}) - (D_{e_k} f_i(a))(h_k)\| \leq \epsilon |h_k|$$

for $k = n$

$$\|f_i(a^n) - f_i(a) - (D_{e_n} f_i(a))(h_n)\|$$

which is equivalent to the existence of n th partial derivative of a . and for $k < n$

$$\begin{aligned} & \|f_i(a^k) - f_i(a^{k+1}) - (D_{e_k} f_i(a))(h_k)\| \\ & \leq \|f_i(a^k) - f_i(a^{k+1}) - (D_{e_k} f_i(a^k))(h_k)\| + \|(D_{e_k} f_i(a^k))(h_k) - (D_{e_k} f_i(a))(h_k)\| \end{aligned}$$

which uses the existence of partial derivatives in neighbourhood and its continuity. ■

Proposition 1.18. Let $f, g : V \rightarrow W$ be differentiable at x and $h : W \rightarrow U$ be differentiable at $y = f(x)$. Furthermore, let c be a scalar then

1. $D(f + cg) = Df + cDg$.
2. $h \circ f$ is differentiable at x and

$$D(h \circ f) = ((Dh) \circ f) \circ Df$$

Proof.

1. we have

$$\begin{aligned} & \|f + cg(x + k) - f + cg(x) - (Df(x) + cDg(x))(k)\| \\ & \leq \|f(x + k) - f(x) - (Df(x))(h)\| + |c| \|g(x + k) - g(x) - (Dg(x))(h)\| \end{aligned}$$

2. we know that

$$\begin{cases} f(x + k) - f(x) - (Df(x))(k) = R(k) \\ h(y + l) - h(y) - (Dh(y))(l) = S(l) \end{cases}$$

and we wish to prove that

$$h \circ f(x + k) - h \circ f(x) - (Dh(f(x)) \circ Df(x))(k) = T(k)$$

where $\|T(k)\| \leq \epsilon \|k\|$ whenever $\|k\| < \delta$. Let $l = f(x + k) - f(x)$ and substituting into the second equation

$$\begin{aligned} & h(f(x + k)) - h(f(x)) - (Dh(y))(f(x + k) - f(x)) \\ & = h(f(x + k)) - h(f(x)) - (Dh(y))((Df(x))(k) + R(k)) \\ & = h(f(x + k)) - h(f(x)) - (Dh(y) \circ Df(x))(k) - (Dh(y))(R(k)) \\ & = T(k) - (Dh(y))(R(k)) = S(l) \\ \implies & T(k) = S(l) + (Dh(y))(R(k)) \end{aligned}$$

Proposition 1.19. $f : U \subset V \rightarrow W_1 \times \dots \times W_n$ is differentiable at x_0 if and only if all its component is differentiable at x_0 . Furthermore, $Df = (Df_1, \dots, Df_n)$.

Proof. Define the following norm on $W_1 \times \dots \times W_n$

$$\|(w_1, \dots, w_n)\| = \sum_{i=1}^n \|w_i\|_{W_i} \quad (1.1)$$

then

$$\|f(x_0 + h) - f(x_0) - (Df(a))(h)\| = \sum_{i=1}^n \|f_i(x_0 + h) - f_i(x_0) - (Df_i(a))(h)\|$$

which is what was what was wanted. ■

Theorem 1.20 (Leibnitz rule). Let V_1, V_2, \dots, V_n be finite dimensional vector spaces and $f : V_1 \times \dots \times V_n \rightarrow W$ is a n -linear function. f is differentiable at $a = (a_1, \dots, a_n)$ and

$$(Df(a))(h_1, \dots, h_n) = f(h_1, a_2, \dots, a_n) + f(a_1, h_2, \dots, a_n) + \dots + f(a_1, a_2, \dots, h_n)$$

Proof. we have that

$$f(a + h) = \sum_{\xi_i \in \{a_i, h_i\}} f(\xi_1, \dots, \xi_n)$$

therefore

$$f(a + h) - f(a) - \sum_{i=1}^n f(a_1, \dots, a_{i-1}, h_i, a_{i+1}, \dots, a_n) = \sum_{\substack{\xi_i \in \{a_i, h_i\} \\ \text{at least two } h_i}} f(\xi_1, \dots, \xi_n)$$

Let $\delta = 1$ then $\|h\| = \sum \|h_i\| < 1$ also i, j , $\|h_i\| \|h_j\| \leq \|h\|^2$. Hence if we define

$$A = \max \left\{ \prod_{i \in I} \|a_i\| \mid I \subset \mathbb{N}_n \right\}$$

then

$$\sum_{\substack{\xi_i \in \{a_i, h_i\} \\ \text{at least two } h_i}} f(\xi_1, \dots, \xi_n) \leq (2^n - n - 1)A \|h\|^2$$

and letting $\delta = \min \left\{ 1, \frac{\epsilon}{(2^n - n - 1)(A + 1)} \right\}$ we arrive at the conclusion. ■

Example 1.5. Let $Z : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$ with $Z(u, v) = u \times v$ be a bilinear function, $f, g : \mathbb{R} \rightarrow \mathbb{R}^3$ and $h(t) = f(t) \times g(t)$. $h = Z \circ \phi$ where $\phi(t) = (f(t), g(t))$. Then we have:

$$\begin{aligned} Dh(t) &= (DZ)(\phi(t)) \circ D\phi(t) \\ &= (DZ)(\phi(t)) \circ (Df(t), Dg(t)) \\ &= Z(Df(t), g(t)) + Z(f(t), Dg(t)) \\ &= Df(t) \times g(t) + f(t) \times Dg(t) \end{aligned}$$

Example 1.6. Consider $A = [f_{ij}(x_1, \dots, x_n)]$ where each f_{ij} is differentiable. Then

$$D\det(A)$$

can be calculated using the Leibnitz rule, since determinant is n -linear function.

1.2.1 Mean value theorem

in general doesnt work $f(t) = (t^2, t^3)$ however it works on a convex domain to reals.

Theorem 1.21. Let V, W be normed finite dimensional vector spaces and $f : U \rightarrow W$ is differentiable and $A, B \in U$ are such that the line connecting in completely contained in U and for each p on that line

$$\|Df(p)\| \leq M$$

then

$$\|f(B) - f(A)\|_W \leq M \|B - A\|_V$$

First consider the following lemma: Assume the following lemma

Lemma 1.22. *If $\phi : [0, 1] \rightarrow W$ is continuous, differentiable on $]0, 1[$ and $\|\phi'(t)\| \leq M$ for all $t \in]0, 1[$ then*

$$\|\phi(1) - \phi(0)\|_W \leq M$$

Proof. We provide three proofs for the lemma

1. Assuming the norm on W is induced by an inner product. Then, let $e = \frac{\phi(1) - \phi(0)}{\|\phi(1) - \phi(0)\|}$ be a unit vector in W then $\psi : [0, 1] \rightarrow \mathbb{R}$, $\psi(t) = e \cdot \phi(t)$ is continuous and differentiable on $]0, 1[$. By the mean value theorem

$$\begin{aligned} |\psi(1) - \psi(0)| &= |\psi'(t_0)| \\ |e \cdot (\phi(1) - \phi(0))| &= |e \cdot \phi'(t)| \\ \|\phi(1) - \phi(0)\| &\leq M \end{aligned}$$

2. Using the Hahn-Banach theorem, that is for a finite dimensional vector space V and $e \in V$ with $\|e\| = 1$ then there exists a linear function $\theta : V \rightarrow \mathbb{R}$ such that $\|\theta\| = 1$ and $\theta(e) = 1$. Now let $\psi(t) = \theta(\phi(t))$ then

$$\begin{aligned} |\psi(1) - \psi(0)| &= |\psi'(t_0)| \\ |\theta(\phi(1) - \phi(0))| &= (D\theta(\phi(t_0)))(\phi'(t_0)) \\ \|\phi(1) - \phi(0)\| &= \|\theta(\phi'(t_0))\| \leq \|\theta\| \|\phi'(t_0)\| \leq M \end{aligned}$$

3. From Haiman. For any ϵ consider the set T_ϵ .

$$T_\epsilon = \{t \in [0, 1] \mid \forall s, 0 \leq s \leq t, \|\phi(s) - \phi(0)\| \leq (M + \epsilon)s + \epsilon\}$$

first note that $T_\epsilon = [0, c]$ and $c > 0$ because for $s = 0$ the inequality is strict and both sides are continuous with respect to s . We claim that $c = 1$ because otherwise $c < 1$ we have, by differentiability of ϕ , there exists a $\delta < 1 - c$ such that if

$$\begin{aligned} \|h\| < \delta &\implies \|\phi(c + h) - \phi(c) - (D\phi(c))(h)\| \leq \epsilon \|h\| \\ &\implies \|\phi(c + h) - \phi(c)\| \leq \|h\| (\epsilon + \|D\phi(c)\|) \\ &\leq \|h\| (\epsilon + M) \end{aligned}$$

also since $c \in T_\epsilon$

$$\begin{aligned} \|\phi(c) - \phi(0)\| &< (M + \epsilon)c + \epsilon \\ \implies \|\phi(c + h) - \phi(0)\| &< (M + \epsilon)(c + h) + \epsilon \quad 0 < h < \delta \end{aligned}$$

hence $c + h \in T_\epsilon$ which is a contradiction and thus $c = 1$. \square

Proof. Let $\sigma : [0, 1] \rightarrow U$ is the parameterization of the line connecting A to B , $\sigma(t) = (1 - t)A + tB$. Let $\phi = f \circ \sigma$ then clearly ϕ is continuous on $[0, 1]$ and differentiable on $]0, 1[$ and we have

$$\begin{aligned} \phi'(t) &= (Df(\sigma(t)))(\sigma'(t)) \\ &= (Df(\sigma(t)))(B - A) \\ \implies \|\phi'(t)\| &\leq \|Df(\sigma(t))\| \|B - A\|_V \leq M \|B - A\|_V \end{aligned}$$

therefore by the Lemma 1.22

$$\|f(B) - f(A)\|_W = \|\phi(1) - \phi(0)\|_W \leq M \|B - A\|_V$$

Corollary 1.23. *Let $U \subset V$ be connected and open and $f : U \rightarrow W$ be differentiable and $Df(u) = 0$ for all $u \in U$ then f is constant.*

Proof. closedness easy, openness from the MVT. ■

Corollary 1.24. *Let V_1, V_2, W be finite dimensional normed vector space and $U \subset V_1 \times V_2$ is open such that for every $y \in V_2$ the intersection $(V_1 \times \{y\}) \cap U$ is connected. Assume $f : U \rightarrow W$ is differentiable and $D_{V_1}f(x, y) = 0$ for all $(x, y) \in U$ then for any two point $(x_1, y), (x_2, y) \in U, f(x_1, y) = f(x_2, y)$.*

1.2.2 Fundamental theorem of calculus

Theorem 1.25. *Let U be an open set of V such that for every $A, B \in U$ the line segment connecting A and B remains in U and let $\sigma : [0, 1] \rightarrow U$ be that line, $\sigma(t) = (1 - t)A + tB$, and lastly let $f : U \rightarrow W$ be continuously differentiable. Then*

$$f(B) - f(A) = T(B - A)$$

where T is

$$T = \int_0^1 Df \circ \sigma(t) dt$$

Proof. look at a matrix, integrate with respect to each element and apply ■

Theorem 1.26. *Consider the $T : U \times U \rightarrow \mathcal{L}(V, W)$ is continuous and such that*

$$f(B) - f(A) = (T(A, B))(B - A)$$

then $f \in C^1$ and $Df(A) = T(A, A)$

Proof. only need to prove f is differentiable and equals to that shit. ■

Corollary 1.27. *Let V be a normed finite dimensional vector space and U is open subset of V . If*

$$f : [a, b] \times U \rightarrow \mathbb{R}$$

is continuous then

$$F(y) = \int_a^b f(x, y) dx$$

is continuous. Furthermore, if $\frac{\partial f}{\partial y_i}$ exists and is continuous then $\frac{\partial F}{\partial y_i}$ exists and is continuous as well.

$$\frac{\partial F}{\partial y_i} = \int_a^b \frac{\partial f}{\partial y_i}(x, y) dx$$

Proof. continuity implies there are balls, compactness implies there are finite balls, take minimum ■

1.2.3 Inverse function theorem

Definition (Local convergence): A functional sequence f_n is **locally convergent** if for each $x \in U$ there exists a open set $x \in V \subset U$ such that $f_n|_V$ is uniformly convergent.

Theorem 1.28. *Let V, W be normed finite dimensional spaces, $U \subset V$ is open and connected, $x_0 \in U$ and $f_n : U \rightarrow W$ is a sequence of differentiable function that*

1. $f_n(x_0)$ is convergent.
2. $Df_n : U \rightarrow \mathcal{L}(V, W)$ is locally convergent to some function $g : U \rightarrow \mathcal{L}(V, W)$

then the sequence f_n is locally convergent to $f : U \rightarrow W$ and $Df = g$. Furthermore, because of connectedness of U for each $x \in U$, $f_n(x)$ is convergent.

Proof. take open ball W around x_0 such that $Df_n|_W$ is uniformly convergent. then prove the first statement.

$$\|f_m(x) - f_n(x)\| \leq \|(f_m - f_n)(x) - (f_m - f_n)(x_0)\| + \|f_m(x_0) - f_n(x_0)\|$$

apply MVT here and make the bounds smaller using (2). Then prove the differentiability with e/3. To prove (3) use open/close argument. ■

contraction fixed point theorem.

Theorem 1.29 (Inverse function theorem). *Let V, W be finite dimensional normed vector space such that $\dim V = \dim W$ and $U \subset V$ is open. If $f : U \rightarrow W$ is continuously differentiable and for some $a \in U$, $Df(a)$ is invertible. Then, there are open set $S \subset V$ and $T \subset W$ that $a \in S \subset U$ and $f(a) \in T$ such that $f|_S$ is bijective and $(f|_S)^{-1} = g$ where $g \in \mathcal{C}^1$ and*

$$Dg(f(x)) = (Df(x))^{-1}$$

Proof. Let S be an open convex set around a such that for all $x \in S$

$$\|Df(x) - Df(a)\| < \frac{1}{2} \|Df^{-1}(a)\|^{-1}$$

hence $Df(x)$ is invertible. Let $T = f(S)$ then we shall prove the following

1. $f|_S$ is bijective.

Let $\psi : S \rightarrow V$ with

$$\begin{aligned} \psi(x) &= x - (Df(a))^{-1} (f(x)) \\ \implies D\psi(x) &= \mathbb{1}_V - (Df(a))^{-1} \circ Df(x) \\ &= (Df(a))^{-1} \circ (Df(a) - Df(x)) \\ \implies \|D\psi(x)\| &\leq \|(Df(a))^{-1}\| \|Df(a) - Df(x)\| \\ &< \frac{1}{2} [(Df(a))^{-1}] [(Df(a))^{-1}]^{-1} = \frac{1}{2} \end{aligned}$$

therefore by mean value theorem

$$\|\psi(x_1) - \psi(x_2)\| \leq \frac{1}{2} \|x_1 - x_2\|$$

then by Theorem 1.12 if

$$\|x_1 - x_2\| \leq K \|f(x_1) - f(x_2)\|$$

we are done. To do so, note that

$$\begin{aligned} & \| (x_1 - x_2) - (Df(a))^{-1} (f(x_1) - f(x_2)) \| \leq \|x_1 - x_2\| \\ \implies & \|x_1 - x_2\| - \| (Df(a))^{-1} (f(x_1) - f(x_2)) \| \leq \frac{1}{2} \|x_2 - x_1\| \\ \implies & \|x_1 - x_2\| \leq 2 \| (Df(a))^{-1} \| \|f(x_1) - f(x_2)\| \end{aligned}$$

2. T is open.

For each $y \in W$ define

$$f_y(x) = x + (Df(a))^{-1} (y - f(x))$$

Since $f_y(x) \in \mathcal{C}^1$ then

$$\begin{aligned} Df_y(x) &= 1_V - (Df(a))^{-1} \circ Df(x) \\ \implies & \|Df_y(x)\| \leq \| (Df(a))^{-1} \| \|Df(a) - Df(x)\| < \frac{1}{2} \\ \implies & \|f_y(x_1) - f_y(x_2)\| \leq \frac{1}{2} \|x_1 - x_2\| \end{aligned}$$

Now for each $f(x_0) = y_0 \in T$ we wish to prove there exist a $\sigma > 0$ such that $B_\sigma(y_0)$ is contained in T . In other words, $\forall y \in B_\sigma(y_0)$

$$\exists x \in S, f(x) = y \iff f_y(x) = x$$

now to apply the contraction fixed point we must find complete metric space X such that $f(X) = X$. Consider $\overline{B_\rho(x_0)} \in S$ which is a complete metric space, and y such that $\|y - y_0\| < \frac{r\rho}{2}$ where $r = \| (Df(a))^{-1} \|^{-1}$. Lastly, we show that $x \in \overline{B_\rho(x_0)} \implies f_y(x) \in \overline{B_\rho(x_0)}$.

$$\begin{aligned} \|f_y(x) - x_0\| &\leq \|f_y(x) - f_y(x_0)\| + \|f_y(x_0) - x_0\| \\ &\leq \frac{1}{2} \|x - x_0\| + \frac{\rho}{2} \leq \rho \end{aligned}$$

therefore there exist a unique $x \in \overline{B_\rho(x_0)}$ such that $f_y(x) = x$.

3. $g = (f|_V)^{-1} : T \rightarrow T$ is continuously differentiable.