Contents

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 $\|\cdot\| : V \to \mathbb{R}$ 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. $\forall x, y \in V\ \ \|x + y\| \leq \|x\| + \|y\|$.

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

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

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

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

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

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

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 $\|\cdot\| : V \to \mathbb{R}$ 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 $\|x - x_0\|_2 < \delta$

$$|\|x\| - \|x_0\|| \leq \|x - x_0\| \leq M\sqrt{n} \, \|x - x_0\| \leq \epsilon$$

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. 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. ∎

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

$$\|f\|_I = \int_0^1 |f(x)| \, \mathrm{d}x$$

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

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

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.

Corollary 1.4. *Any finite dimensional normed vector space V over a normed complete field $\mathbb{F}$ is a Banach space.*

1.1.2 Linear Maps

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

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

Proposition 1.5. *Let $T : U \to V$ and $T' : V \to W$ 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 1.6. *Let $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ be normed spaces and $T : V \to W$ 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

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

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

item 2 ⇒ item 3:

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

item 3 ⇒ item 4: Trivial.

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

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

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

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

Theorem 1.7. *If V is a finite dimensional normed vector space then any linear transformation $T : V \to W$ 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 $\{e_1, \ldots, e_n\}$. Let x be such that $\|x\|_2 < \delta$ for some $\delta > 0$. Therefore, $|\xi_i| < \delta^2$

$$\|Tx\|_W = \left\| \sum \xi_i T(e_i) \right\|_W \leq \sum |\xi_i| \|T(e_i)\|_W \leq \delta^2 K$$

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 1.6. ∎

Theorem 1.8. *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

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

3. for the triangle inequality

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

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

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

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

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

Note that $|\|T_m| - \|T_n\|| \le \|T_m - T_n\|$ ans 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 $n \ge N$.

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

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

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

which means that

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

Therefore $T_i \to T$ as desired. ■

Theorem 1.9. *Let $(V, \|\cdot\|)$ be a normed space. Then any linear transformation $T : \mathbb{R}^n \to V$ is continuous. Furthermore, if T is a bijection, it is a homeomorphis.*

Proof. Since $\mathbb{R}^n$ is finite then by Theorem 1.7, 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} is a continuous map. ■

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

Proposition 1.10. *If V is a finite (also works for infinite) vector space then $\mathrm{GL}(V)$ is open in $\mathcal{L}(V, V)$, in fact, if $f \in \mathrm{GL}(V)$ then the open ball centered at f with radius $\|f^{-1}\|^{-1}$ remains in $\mathrm{GL}(V)$. Furthermore, the inverse operator $i : \mathrm{GL}(V) \to \mathrm{GL}(V)$, $i(T) = T^{-1}$ is continuous.*

Theorem 1.11. *$T : \mathbb{R}^n \to \mathbb{R}^n$ linear transformation is invertible if and only if there exists a c such that:*

$$c\,\|x\| \leq \|Tx\|$$

Proof. If T is invertible then $T^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ is bounded and thus

$$\left\|T^{-1}x\right\| \leq c\,\|x\|$$

and since T is bijective then there exists y such that $x = Ty$ which implies

$$\|y\| \leq c\,\|Ty\|$$

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 hence T is invertible. ■

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

Proposition 1.12. *If $V_1, V_2, \ldots, V_n$ are normed vector spaces and $\phi : V_1 \times \cdots \times V_n \to W$ is a n-linear then the followings are equivalent*

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

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

Proof. Item 1 $\implies$ Item 2: Trivial.
Item 2 $\implies$ Item 3 ■

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 $f : U \subset V \to W$ where U is open. Then f is differentiable at x_0 when a linear transformation $T : V \to W$ such that

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

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

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

T if it exists is unique, represented by $f'(x_0)$, $\mathrm{D}f$, or $\mathrm{d}f(x)$ and called the **total derivative** or **Fréchet derivative**.

Example 1.2. Every linear function $f : V \to W$ with $f(v) = Tv + b$ where $b \in W$ is differentiable and $\mathrm{D}f(v) = T$. Since

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

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

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

Example 1.3. Let $\mu : \mathbb{R} \times V \to V$ with $\mu(r, x) = rx$. Then μ is differentiable and $(\mathrm{D}\mu(r, x))(t, h) = rh + tx$ as

$$\|\mu((r+t), (x+h)) - \mu(r, x) - (\mathrm{D}\mu(r, x))(t, h)\| = \|rx + rh + tx + th - rx - rh - tx\| = |t| \, \|h\| \le \epsilon \, \|(t,$$

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

Proposition 1.13. *Differentiability of f at x implies continuity at x.*

Proof.

$$\|f(x+h) - f(x)\| = \|(\mathrm{D}f(x))(h) + R(h)\| \le \|\mathrm{D}f(x)\| \, \|v\| + \|R(v)\| \to 0$$

as $v \to 0$. ■

Proposition 1.14. *Assume $f : U \subset V \to W$ is differentiable at x_0 and let $u \in V$ be a non-zero vector then*

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

Proof. Let $h = tu$ then

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

Definition (Directional derivative): If we let $\|u\| = 1$ then the limit in Propostion 1.14 becomes the **directional derivative** of f in the direction of u and is denoted by $\mathrm{D}_u f$.

Remark 2. The existence of all directional derivatives of f doesnt imply its differentiability or even its continuity.

Theorem 1.15. *$f : V \to W$ has all of its partial derivative and they're continuous then f is differentiable.*

Proof. Let $\{e_1, \ldots, e_n\}$ be a unit basis for V.

$$\left\| f(x_1 + h_1, \ldots, x_n + h_n) - f(x_1, \ldots, x_n) - \sum_{i=1}^{n} (\mathrm{D}_{e_i} f(x))(h) \right\|$$
$$\leq \sum_{i=1}^{n} \|f(x_1, \ldots, x_{i-1}, x_i + h_i, \ldots, x_n + h_n) - f(x_1, \ldots, x_i, x_{i+1} + h_{i+1}, \ldots x_n + h_n) - (\mathrm{D}_{e_i} f(x))(h)\|$$
$$\leq$$

Proposition 1.16. *Let $f, g : V \to W$ be differentiable at x and $h : W \to U$ be differentiable at $y = f(x)$. Furthermore, let c be an scalar then*

1. $\mathrm{D}(f + cg) = \mathrm{D}f + c\,\mathrm{D}g$.
2. *$h \circ f$ is differentiable at x and*
$$\mathrm{D}(h \circ f) = (\mathrm{D}(h) \circ f) \circ \mathrm{D}(f)$$
3. *For a bilinear function β*
$$(\mathrm{D}\beta(f, g))(v) = \beta((\mathrm{D}f)(v)\,, g) + \beta(f, (\mathrm{D}g)(v))$$

Proposition 1.17. *$f : U \subset V \to W_1 \times \cdots \times W_n$ is differentiable at x_0 if and only if all its component is differentiable at x_0. Furthermore, $\mathrm{D}f = (\mathrm{D}f_1, \ldots, \mathrm{D}f_n)$.*

Theorem 1.18 (Leibnitz rule). *Let $V_1, V_2, \ldots, V_n$ be finite dimensional vector spaces and $f : V_1 \times \cdots \times V_n \to W$ is a n-linear function. f is differentiable at $a = (a_1, \ldots, a_n)$ and*

$$(\mathrm{D}f(a))(h_1, \ldots h_n) = f(h_1, a_2, \ldots, a_n) + f(a_1, h_2, \ldots, a_n) + \cdots + f(a_1, a_2, \ldots, h_n)$$

Proof.

Example 1.4. Let $Z : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ with $Z(u, v) = u \times v$ be a bilinear function, $f, g : \mathbb{R} \to \mathbb{R}^3$ and $h(t) = f(t) \times g(t)$. $h = Z \circ \phi$ where $\phi(t) = (f(t)\,, g(t))$. Then we have:

$$\begin{aligned}
\mathrm{D}h(t) &= (\mathrm{D}Z)(\phi(t)) \circ \mathrm{D}\phi(t) \\
&= (\mathrm{D}Z)(\phi(t)) \circ (\mathrm{D}f(t)\,, \mathrm{D}g(t)) \\
&= Z(\mathrm{D}f(t)\,, g(t)) + Z(f(t)\,, \mathrm{D}g(t)) \\
&= \mathrm{D}f(t) \times g(t) + f(t) \times \mathrm{D}g(t)
\end{aligned}$$

Example 1.5. Consider $A = [f_{ij}(x_1, \ldots, x_n)]$ where each f_{ij} is differentiable. Then

$$\mathrm{D}\det(A)$$

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