Multivariable Calculus

Swayam Chube

November 8, 2022

Contents

Chapter 0

Preliminaries

Definition 0.1 (Vector Space)**.** Let F be a field and V an Abelian group. Then V is said to be an F-vector space, if

1. for all $v \in V$, $0_F v = 0_V$ and $1_F v = v$
2. for all $a \in F$ and $v \in V$, $av \in V$
3. for all $a, b \in F$ and $v \in V$, $(a+b)v = av + bv$
4. for all $a \in F$ and $u, v \in V$, $a(u+v) = au + av$

Definition 0.2 (Linear Transformation)**.** Let F be a field and V, W be F-vector spaces. A linear transformation is a function $T : V \to W$ such that

1. for all $u, v \in V$, $T(u+v) = T(u) + T(v)$
2. for all $a \in F$ and $v \in V$, $T(av) = aT(v)$

Definition 0.3 (Inner Product)**.** Let F be a field and V be a vector space over F. An inner product is a function $\langle \cdot, \cdot \rangle : V \times V \to F$ such that for all $x, y, z \in V$ and $a \in F$,

1. $\langle x + z, y \rangle = \langle x, y \rangle + \langle z, y \rangle$

2. $\langle ax, y\rangle = a\langle x, y\rangle$
3. $\langle x, y\rangle = \langle y, x\rangle$
4. $\langle x, x\rangle > 0$ if and only if $x \neq 0$

We shall mainly concern ourselves with the case where $F = \mathbb{R}$ and $V = \mathbb{R}^n$ for some positive integer n.

Lemma 0.4. *Let $T : \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation. There is a number M such that $\|T(h)\| \leq M\|h\|$ for all $h \in \mathbb{R}^m$.*

Proof. Since T is a linear transformation, it has a matrix corresponding to the chosen standard basis, say $A \in \mathbb{R}^{n\times m}$ where

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix}$$

Let $h = (h_1, \cdots, h_n)$. Then

$$\begin{aligned}
\|Ah\| &= \sqrt{\sum_{j=1}^{n}\left(\sum_{i=1}^{m} a_{ji}h_i\right)^2} \\
&\leq \sqrt{\sum_{j=1}^{n}\left(\sum_{i=1}^{m} a_{ji}^2 \sum_{i=1}^{n} h_i^2\right)} \\
&\leq \sqrt{\|A\|^2\|h\|^2} = \|A\|\|h\|
\end{aligned}$$

This completes the proof. ■

Chapter 1

Differentiation

Definition 1.1 (Differentiable)**.** A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is *differentiable* at $a \in \mathbb{R}^n$ if there is a linear transformation $\lambda : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{h \to \mathbf{0}} \frac{\|f(a+h) - f(a) - \lambda(h)\|}{\|h\|} = 0$$

where $h \in \mathbb{R}^n$. *The* linear transformation λ is then denoted by $Df(a)$ and called the *derivative* of f at a.

We shall now show that the aforementioned linear transformation is unique:

Proposition 1.2. *If $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $a \in \mathbb{R}^n$, then there is a unique linear transformation $\lambda : \mathbb{R}^n \to \mathbb{R}^m$ such that*

$$\lim_{h \to \mathbf{0}} \frac{\|f(a+h) - f(a) - \lambda(h)\|}{\|h\|} = 0$$

Proof. Define $d(h) = f(a+h) - f(a)$. Further, let $\mu : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation satisfying the differentiability condition. We then have

$$\begin{aligned} \lim_{h \to \mathbf{0}} \frac{\|\lambda(h) - \mu(h)\|}{\|h\|} &= \lim_{h \to 0} \frac{\|\lambda(h) - d(h) + d(h) - \mu(h)\|}{\|h\|} \\ &\leq \lim_{h \to \mathbf{0}} \frac{\|\lambda(h) - d(h)\|}{\|h\|} + \lim_{h \to \mathbf{0}} \frac{\|d(h) - \mu(h)\|}{\|h\|} = 0 \end{aligned}$$

Let $x \in \mathbb{R}^n$. Then, for all $t \in \mathbb{R}$, $\lambda(tx) = t\lambda(x)$ and $\mu(tx) = t\mu(x)$. Therefore,

$$0 = \lim_{t \to \mathbf{0}} \frac{\|\lambda(tx) - \mu(tx)\|}{\|tx\|} = \frac{\|\lambda(x) - \mu(x)\|}{\|x\|}$$

This completes the proof. ■

Since $Df(a)$ is a linear transformation, it can be given by a matrix multiplication in $\mathbb{R}^{m \times n}$.

Definition 1.3 (Jacobian Matrix)**.** The matrix of $Df(a) : \mathbb{R}^n \to \mathbb{R}^m$ with respect to the standard ordered basis of $\mathbb{R}^n$ and $\mathbb{R}^m$ is called the *Jacobian Matrix* of f at a.

Theorem 1.4 (Chain Rule)**.** *If $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at a and $g : \mathbb{R}^m \to \mathbb{R}^p$ is differentiable at $f(a)$, then the composition $g \circ f : \mathbb{R}^n \to \mathbb{R}^p$ is differentiable at a and*

$$D(g \circ f)(a) = Dg(f(a)) \circ Df(a)$$

Proof. Let $b = f(a)$, $\lambda = Df(a)$ and $\mu = Dg(f(a))$. Further, define:

1. $\varphi(x) = f(x) - f(a) - \lambda(x - a)$ where $\varphi : \mathbb{R}^n \to \mathbb{R}^m$
2. $\psi(x) = g(x) - g(b) - \mu(x - b)$ where $\psi : \mathbb{R}^m \to \mathbb{R}^p$
3. $\rho(x) = g \circ f(x) - g \circ f(a) - \mu \circ \lambda(x - a)$ where $\rho : \mathbb{R}^n \to \mathbb{R}^p$

Note that

$$\begin{aligned}
\rho(x) &= g(f(x)) - g(b) - \mu(\lambda(x - a)) \\
&= g(f(x)) - g(b) - \mu(f(x) - f(a) - \varphi(x)) \\
&= \underbrace{g(f(x)) - g(b) - \mu(f(x) - b)} - \mu(\varphi(x)) \\
&= \psi(f(x)) + \mu(\varphi(x))
\end{aligned}$$

Let $\varepsilon > 0$ be given. Then, there is δ such that $\|\psi(f(x))\| < \varepsilon\|f(x) - b\|$ whenever $\|f(x) - b\| < \delta$. Now, since f is differentiable at a, there is δ_1 such that $\|f(x) - b\| < \delta$ whenever $\|x - a\| < \delta_1$.

As a result, whenever $\|x-a\| < \delta_1$,

$$\begin{aligned}\|\psi(f(x))\| &< \varepsilon\|f(x)-b\| \\ &= \varepsilon\|\varphi(x)+\lambda(x-a)\| \\ &\leq \varepsilon\left(\|\varphi(x)\| + M\|x-a\|\right)\end{aligned}$$

for some $M \in \mathbb{R}^+$.

It is now obvious that

$$\lim_{x\to a} \frac{\|\psi(f(x))\|}{\|x-a\|} = 0$$

Finally, it is not hard to see that

$$\lim_{x\to a} \frac{\|\mu(\varphi(x))\|}{\|x-a\|} = \lim_{x\to a} \frac{\|\mu(\psi(x)-\psi(a))\|}{\|\psi(x)-\psi(a)\|} \frac{\|\psi(x)-\psi(a)\|}{\|x-a\|} \leq M' \lim_{x\to a} \frac{\|\psi(x)-\psi(a)\|}{\|x-a\|} = 0$$

Therefore,

$$\lim_{x\to a} \frac{\|\rho(x)\|}{\|x-a\|} = 0$$

■

Proposition 1.5. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ with $f \equiv (f_1, \ldots, f_m)$ and $a \in \mathbb{R}^n$. Then f is differentiable at a if and only if each f_i is, and*

$$Df(a) = \begin{pmatrix} Df_1(a) \\ \vdots \\ Df_m(a) \end{pmatrix}$$

Proof. If f is differentiable at a, then $f_i \equiv \pi_i \circ f$ is trivially differentiable at a. Conversely, suppose each f_i is differentiable. Then, let $\lambda : \mathbb{R}^n \to \mathbb{R}^m$ denote the linear transformation given by the matrix

$$\lambda \equiv \begin{pmatrix} Df_1(a) \\ \vdots \\ Df_m(a) \end{pmatrix}$$

Then,

$$\begin{aligned}\lim_{h\to \mathbf{0}} \frac{\|f(a+h)-f(a)-\lambda(h)\|}{\|h\|} &\leq \lim_{h\to \mathbf{0}} \sum_{i=1}^{m} \frac{\|f_i(a+h)-f_i(a)-Df_i(a)h\|}{\|h\|} \\ &= 0\end{aligned}$$

This completes the proof. ∎

Proposition 1.6. *If $f, g : \mathbb{R}^n \to \mathbb{R}$ are differentiable at a, then*

$$D(f+g)(a) = Df(a) + Dg(a)$$
$$D(f \cdot g)(a) = g(a)Df(a) + f(a)Dg(a)$$

Moreover, if $g(a) \neq 0$, then

$$D(f/g)(a) = \frac{g(a)Df(a) - f(a)Dg(a)}{g(a)^2}$$

Proof. Let $h = f + g$ and $\lambda = Df(a) + Dg(a)$. Then,

$$\lim_{h \to \mathbf{0}} \frac{\|f(a+h) - f(a) + g(a+h) - g(a) - Df(a)(h) - Dg(a)(h)\|}{\|h\|}$$
$$\leq \lim_{h \to \mathbf{0}} \frac{\|f(a+h) - f(a) - Df(a)(h)\|}{\|h\|} + \lim_{h \to \mathbf{0}} \frac{\|g(a+h) - g(a) - Dg(a)(h)\|}{\|h\|}$$
$$= 0$$

Let $h = f \cdot g$ and $\lambda = g(a)Df(a) + f(a)Dg(a)$. There is a significant amount of algebraic manipulation involved here:

$$\|f(a+h)g(a+h) - f(a)g(a) - g(a)Df(a)(h) - f(a)Dg(a)(h)\|$$
$$\leq \|(g(a+h) - g(a))(f(a+h) - f(a))\| + \|f(a)(g(a+h) - g(a) - Dg(a)(h))\|$$
$$+ \|g(a)(f(a+h) - f(a) - Df(a)(h))\|$$

This immediately implies the desired conclusion.

TODO: Write up the thing for f/g ∎

Definition 1.7 (Partial Derivative)**.** Let $f : \mathbb{R}^n \to \mathbb{R}$ and $a \in \mathbb{R}^n$. Then, the limit

$$\lim_{h \to \mathbf{0}} \frac{f(a_1, \ldots, a_i + h, \ldots, a_n) - f(a_1, \ldots, a_n)}{h}$$

if it exists is called the i-th partial derivative of f at a and is denoted by $D_i f(a)$.

We denote $D_j(D_if)(x)$ as $D_{i,j}f(x)$. It is important to remember that this notation reverses the order of i and j. Derivatives such as $D_{i,j}f$ are called *second-order* or *mixed* partial derivatives of f.

Theorem 1.8. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable. If $D_{i,j}f$ and $D_{j,i}f$ exist and are continuous in an open set containing a, then*

$$D_{i,j}f(a) = D_{j,i}f(a)$$

Proof. This proof requires Fubini's Theorem. Suppose $D_{i,j}f(a) - D_{j,i}f(a) \neq 0$. Without loss of generality, suppose $D_{i,j}f(a) - D_{j,i}f(a) > 0$. Let $\varepsilon > 0$. Using continuity, there is an open set U containing a such that $D_{i,j}f(x) - D_{j,i}f(x) > \varepsilon$ for all $x \in U$. Let A be a closed rectangle contained in A.

But, due to Fubini's Theorem,

$$\int_A D_{i,j}f(x) - D_{j,i}f(x) = \int \cdots \int D_{i,j}f(x) - \int \cdots \int D_{j,i}f(x) = 0$$

A contradiction. ■

Proposition 1.9. *If $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $a \in \mathbb{R}^n$, then $D_jf_i(a)$ exists for $1 \leq i \leq m$ and $1 \leq j \leq n$ and $f'(a)$ is the $m \times n$ matrix $\left[D_jf_i(a)\right]_{m\times n}$*

Proof. That each partial $D_jf_i(a)$ exists is immediate from the fact that f is differentiable. Define the function $h_i : \mathbb{R} \to \mathbb{R}^n$ given by $h_i(x) = (a_1, \ldots, x, \ldots, a_n)$ where x occurs at the i-th position. Then, by definition, we have that $D_if(a) = D(f \circ h_i)(a_i)$. And due to the chain rule,

$$D_if(a) = D(f \circ h_i)(a) = Df(h_i(a_i))Dh(a_i) = Df(a)\begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

This immediately implies the desired conclusion. ■

In conclusion,

$$\text{Differentiable at } a \Longrightarrow \text{Partials exist at } a$$

Theorem 1.10 (Differentiability Theorem)**.** *If $f : \mathbb{R}^n \to \mathbb{R}^m$, then $Df(a)$ exists if all $D_j f_i(x)$ exist in an open set containing a and if each function $D_j f_i$ is continuous at a. Such a function f is called* ***continuously differentiable****.*

Proof. Recall that due to a preceeding theorem, f is differentiable if and only if each f_i is differentiable and its derivative is given by stacking the derivatives of all the respective functions. Therefore, it suffices to show this in the case when $m = 1$. Now,

$$\begin{aligned} f(a+h) - f(a) &= f(a_1 + h_1, a_2, \ldots, a_n) - f(a_1, \ldots, a_n) \\ &\quad + f(a_1 + h_1, a_2 + h_2, a_3, \ldots, a_n) - f(a_1 + h_1, a_2, \ldots, a_n) \\ &\quad + f(a_1 + h_1, \ldots, a_n + h_n) - f(a_1 + h_1, \ldots, a_{n-1} + h_{n-1}, a_n) \end{aligned}$$

Now, due to the Mean Value Theorem, there are real numbers b_i such that

$$\begin{aligned} &f(a_1 + h_1, \ldots, a_i + h_i, a_{i+1}, \ldots, a_n) - f(a_1 + h_1, \ldots, a_i, \ldots, a_n) \\ &= h_i D_i f(a_1 + h_1, \ldots, b_i, \ldots, a_n) \end{aligned}$$

We now have

$$\begin{aligned} &\lim_{h \to \mathbf{0}} \frac{|f(a+h) - f(a) - \sum_{i=1}^n D_i f(a) h_i|}{\|h\|} \\ &= \lim_{h \to \mathbf{0}} \frac{|\sum_{i=1}^n (D_i f(c_i) - D_i f(a)) h_i|}{\|h\|} \\ &\leq \lim_{h \to \mathbf{0}} \sum_{i=1}^n \frac{|D_i f(c_i) - D_i f(a)||h_i|}{\|h\|} \\ &\leq \lim_{h \to \mathbf{0}} \sum_{i=1}^n |D_i f(c_i) - D_i f(a)| = 0 \end{aligned}$$

This completes the proof. ■

Definition 1.11 (Directional Derivative)**.** Let $f : \mathbb{R}^n \to \mathbb{R}$. For $v \in \mathbb{R}^n$, the limit

$$\lim_{h \to 0} \frac{f(a + hv) - f(a)}{h}$$

if it exists, is denoted by $D_v f(a)$ and is called the *directional derivative* of f at a in the direction v.

Lemma 1.12. *Let $f : \mathbb{R}^n \to \mathbb{R}$ and $a \in \mathbb{R}^n$. If f is differentiable at a, then $D_v f(a) = Df(a)(v)$.*

Proof. By definition, we have that

$$0 = \lim_{h \to 0} \frac{f(a + hv) - f(a) - Df(a)(hv)}{h} = \lim_{h \to 0} \frac{f(a + hv) - f(a)}{h} - Df(a)(v)$$

and we have the desired conclusion. ■

1.1 Counterexamples

Partials but not Differentiable

Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ given by

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

We first show that f is continuous at $(0, 0)$. Indeed,

$$\lim_{(h,k) \to \mathbf{0}} \frac{h^2 k}{h^2 + k^2} \leq \lim_{(h,k) \to \mathbf{0}} k = 0$$

Next, note that both the partials exist and

$$D_x f(\mathbf{0}) = \begin{cases} \frac{2xy^3}{(x^2 + y^2)^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

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

We shall now show that the function is not differentiable at $(0, 0)$. Suppose it were. Then, the derivative $Df(a)$ would be given by

$$Df(\mathbf{0}) = \begin{pmatrix} D_x f(0) & D_y f(0) \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix}$$

Therefore, one would expect,

$$\lim_{h\to\mathbf{0}} \frac{f(h)}{\|h\|} = 0$$

But note that

$$\lim_{h\to 0} \frac{f(h,h)}{h} = \frac{1}{2}$$

a contradiction.

Directional Derivatives but not Continuous

Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ given by

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

We shall first show that the directional derivatives exist at $\mathbf{0}$. Indeed, for $\mathbf{v} = (a,b)$, with $a \neq 0$,

$$\lim_{h\to 0} \frac{f(ha,hb)}{h\sqrt{a^2+b^2}} = \lim_{h\to 0} \frac{ab^2}{(a^2+h^2b^4)} = \frac{b^2}{a}$$

On the other hand,

$$\lim_{h\to 0} \frac{f(0,h)}{h} = 0$$

But, f is not continuous at $\mathbf{0}$, since it takes the value $\frac{1}{2}$ for each point on the parabola $y^2 = x$.

Continuous and Directional Derivatives but not Differentiable

Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ given by

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

The continuity at $\mathbf{0}$ is obvious. Let $\mathbf{v} = (a,b)$ with $b \neq 0$. Then, the directional derivative of f at $\mathbf{0}$ in the direction of v is given by

$$\lim_{h\to 0} \frac{f(ah,bh)}{h} = \frac{bh}{|bh|}\frac{|h|}{h}\sqrt{a^2+b^2} = \frac{b}{|b|}\sqrt{a^2+b^2}$$

On the other hand, if $b = 0$, then $D_{\mathbf{v}} f(\mathbf{0})$ is just 0.

We shall now show that teh function is not differentiable at $(0,0)$. Suppose it were. Then, the derivative $Df(\mathbf{0})$ would be given by

$$Df(\mathbf{0}) = \begin{pmatrix} 0 & 1 \end{pmatrix}$$

Therefore, one would expect

$$0 = \lim_{(h,k)\to\mathbf{0}} \frac{f(h,k) - k}{\sqrt{h^2 + k^2}}$$

But note that

$$\lim_{h\to 0} \frac{f(h,h) - h}{h\sqrt{2}} = 1 - \frac{1}{\sqrt{2}}$$

a contradiction.

Chapter 2

Integration

2.1 Multiple Integrals

Definition 2.1 (Partition)**.** A *partition* of a rectangle $\mathcal{R} = [a_1, b_1] \times \cdots \times [a_n, b_n]$ is a collection $P = (P_1, \ldots, P_n)$ where each P_i is a partition of the interval $[a_i, b_i]$. A partition divides $\mathcal{R}$ into *subrectangles*.

Definition 2.2 (Refinement)**.** A partition P' of $\mathcal{R}$ is said to refine P if every subrectangle of P' is contained in a subrectangle of P.

Definition 2.3 (Upper and Lower Sums)**.** Let $\mathcal{R}(P)$ denote the subrectangles induced by P. Further, for each subrectangle S, let $m_S(f)$ and $M_S(f)$ denote the minimum and maximum values taken by f on S respectively. Then, we define

$$L(f, P) = \sum_{S \in \mathcal{R}(P)} m_S(f) v(S) \qquad U(f, P) = \sum_{S \in \mathcal{R}(P)} M_S(f) v(S)$$

as the lower and upper sums of f for P respectively.

Lemma 2.4. *Suppose the partition P' refines P over $\mathcal{R}$. Then,*

$$L(f,P) \leq L(f,P') \qquad \text{and} \qquad U(f,P') \leq U(f,P)$$

Proof. Obviously, for each $S \in \mathcal{R}(P')$, there is a unique $\mathscr{C}(S) \in \mathcal{R}(P)$ containing it. Therefore,

$$L(f,P') = \sum_{S \in \mathcal{R}(P')} m_S(f) v(S) \geq \sum_{S \in \mathcal{R}(P')} m_{\mathscr{C}(S)}(f) v(S) = L(f,P)$$

and similarly, for the upper sum. ■

Corollary. *If P and P' are any two partitions, then $L(f,P') \leq U(f,P)$.*

Proof. Let P'' be a partition which refines both P and P'. Then,

$$L(f,P') \leq L(f,P'') \leq U(f,P'') \leq U(f,P)$$

■

Definition 2.5. Let $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ be a function where A is a closed rectangle. Let $\mathscr{P}$ be the set of all possible partitions of A. Then f is said to be integrable over A if

$$\sup_{P \in \mathscr{P}} L(f,P) = \inf_{P \in \mathscr{P}} U(f,P)$$

and this common value is said to be the *integral* of f over A and is often denoted by $\int_A f \, d\mathbf{x}$.

Definition 2.6 (Measure Zero)**.** A suset A of $\mathbb{R}^n$ has measure 0 if for every $\varepsilon > 0$, there is a countable cover $\{U_1, U_2, \ldots\}$ of A by closed/open rectangles such that $\sum_{i=1}^{\infty} v(U_i) < \varepsilon$.

Proposition 2.7. *If $A = \bigcup_{i=1}^{\infty} A_i$ and each A_i has measure 0, then A has measure 0.*

Proof. Straightforward. ■

Definition 2.8 (Content Zero)**.** A subset A of $\mathbb{R}^n$ has content0 if for every $\varepsilon > 0$, there is a finite cover $\{U_1, \dots, U_n\}$ of A by closed/open rectangles such that $\sum_{i=1}^{n} v(U_i) < \varepsilon$.

A set A with content 0 obviously has measure 0.

Proposition 2.9. *If A is compact and has measure 0, then A has content 0.*

Proof. Trivial. ■

Definition 2.10 (Oscillation)**.** Let $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ be a function. Define

$$M(a, f, \delta) = \sup\{f(x) : x \in A,\ \|x - a\| < \delta\}$$
$$m(a, f, \delta) = \inf\{f(x) : x \in A,\ \|x - a\| < \delta\}$$

The *oscillation* $o(f, a)$ of f at a is defined by

$$o(f, a) = \lim_{\delta \to 0} M(a, f, \delta) - m(a, f, \delta)$$

Lemma 2.11. *Let A be a closed rectangle and let $f : A \to \mathbb{R}$ be a bounded function such that $o(f, x) < \varepsilon$ for all $x \in A$. Then there is a partition P of A with $U(f, P) - L(f, P) < \varepsilon \cdot v(A)$.*

Proof. Using the definition of a limit, for each $x \in A$, there is an open set U_x containing x such that $M_{U_x}(f) - m_{U_x}(f) < \varepsilon$. Since A is compact, there is $\{U_1, \dots, U_m\} \subseteq \{U_x \mid x \in A\}$ that forms a finite subcover for A.

Now, choose a partition P of A such that for all subrectangles $S \in P$, there is an open set U_j containing it. Then, we may trivially note that

$$M_S(f) - m_S(f) < M_{U_j}(f) - m_{U_j}(f) < \varepsilon$$

The conclusion is now obvious. ■

Theorem 2.12. *Let A be a closed rectangle and $f : A \to \mathbb{R}$ a bounded function. Let $B = \{x \mid f \text{ is not continuous at } x\}$. Then f is integrable if and only if B is a set of measure 0.*

Proof. ■

Definition 2.13 (Characteristic Function)**.** If $C \subseteq \mathbb{R}^n$, the *characteristic function* of C is defined as

$$\chi_C = \begin{cases} 1 & x \in C \\ 0 & x \notin C \end{cases}$$

Theorem 2.14. *Let A be a closed rectangle and $C \subseteq A$. Then, $\chi_C : A \to \mathbb{R}$ is integrable if and only if ∂C has measure 0.*

Proof. It is not hard to argue that

$$\partial C = \{x \mid \chi_C \text{ is not continuous at } x\}$$

This implies the desired conclusion. ■

Definition 2.15 (Jordan Measurable, Content)**.** A bounded set C whose boundary has measure 0 is called *Jordan-measurable*. The integral $\int_C 1$ is called the *content* of C or the *volume* of C.

2.1.1 Properties of Integrals

In what follows $A \subseteq \mathbb{R}^n$ is a closed rectangle.

Proposition 2.16. *Let $f, g : A \to \mathbb{R}$ be bounded integrable functions. Then $f + g$ and cf are integrable for all $c \in \mathbb{R}$.*

Proof. Obviously, $f+g$ is bounded. Using the integrability of f and g, there are partitions P_f and P_g of A such that $U(f,P_f) - L(f,P_f) < \varepsilon/2$ and $U(g,P_g) - L(g,P_g) < \varepsilon/2$. Let P be a partition of A that refines both P_f and P_g. Then,

$$\begin{aligned}
U(f+g,P) - L(f+g,P) &= \sum_{S\in P} (M_S(f+g) - m_S(f+g))v(S) \\
&\leq \sum_{S\in P} (M_S(f) + M_S(g) - m_S(f) - m_S(g))v(S) \\
&= (U(f,P) - L(f,P)) - (U(g,P) - L(g,P))\, v(S) \\
&< \varepsilon
\end{aligned}$$

Therefore, $f+g$ is integrable.

The assertion regarding cf is trivial. ■

The converse of the statement regarding $f+g$ is obviously not true. Since $\mathbf{0} = \chi_{\mathbb{Q}\cap[0,1]} + (-\chi_{\mathbb{Q}\cap[0,1]})$ but neither oef the functions on the right hand side of the equality are integrable.

Lemma 2.17. *Let $f : A \to \mathbb{R}$ be a bounded integrable function. Then $|f|$ is integrable.*

Proof. Obviously $|f|$ is bounded. Since f is integrable, there is a partition P of A with $U(f,P) - L(f,P) < \varepsilon$. Then,

$$\begin{aligned}
U(|f|,P) - L(|f|,P) &= \sum_{S\in P} (M_S(|f|) - m_S(|f|))v(S) \\
&\leq \sum_{S\in P} (M_S(f) - m_S(f))v(S) \\
&< \varepsilon
\end{aligned}$$

Therefore $|f|$ is integrable. ■

Lemma 2.18. *Let $f : A \to \mathbb{R}$ be a bounded integrable function. Then f^2 is integrable.*

Proof. Let $T = \sup_{x \in A} f(x)$. Using the integrability of f, there is a partition P of A such that $U(f,P) - L(f,P) < \varepsilon/2T$. As a result, we have

$$\begin{aligned} U(f^2,P) - L(f^2,P) &= \sum_{S \in P} (M_S(f^2) - m_S(f^2))v(S) \\ &\leq \sum_{S \in P} 2T(M_S(f) - m_S(f))v(S) \\ &< 2T \cdot \frac{\varepsilon}{2T} = \varepsilon \end{aligned}$$

This completes the proof. ■

The converse is not true. Consider the function $\theta : [0,1] \to \mathbb{R}$ given by

$$\theta(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ -1 & \text{otherwise} \end{cases}$$

Then f is not integrable but $f^2 = \mathbf{1}$ is.

Proposition 2.19. *Let $A \subseteq \mathbb{R}$ be a closed rectangle and $f, g : A \to \mathbb{R}$ be bounded and integrable. Then, $f \cdot g$ is integrable.*

Proof. Obviously, $f \cdot g$ is bounded. Since we may write

$$fg = \frac{1}{4}\left((f+g)^2 - (f-g)^2\right)$$

it follows that fg is integrable. ■

Theorem 2.20 (Fubini)**.** *Let $A \subseteq \mathbb{R}^n$ and $B \subseteq \mathbb{R}^m$ be closed rectangles, and let $f : A \times B \to \mathbb{R}$ be integrable. For $x \in A$, let $g_x : B \to \mathbb{R}$ be defined by $g_x(y) = f(x,y)$ and let*

$$\begin{aligned} \mathfrak{L}(x) &= \mathbf{L}\int_B g_x = \mathbf{L}\int_B f(x,y)\,dy \\ \mathfrak{U}(x) &= \mathbf{U}\int_B g_x = \mathbf{U}\int_B f(x,y)\,dy \end{aligned}$$

Then $\mathfrak{L}$ and $\mathfrak{U}$ are integrable on A and

$$\int_{A\times B} f = \int_A \mathfrak{L} = \int_A \left(\mathbf{L}\int_B f(x,y)\,dy\right)\,dx$$
$$\int_{A\times B} f = \int_A \mathfrak{U} = \int_A \left(\mathbf{U}\int_B f(x,y)\,dy\right)\,dx$$

Proof. Let P_A be a partition of A and P_B be a partition of B. Then $P = P_A \times P_B$ is a partition of $A \times B$. Now, note that

$$\begin{aligned} L(f,P) &= \sum_{S\in P} m_S(f)v(S) \\ &= \sum_{S_A\in P_A}\sum_{S_B\in P_B} m_{S_A\times S_B}(f)v(S_A\times S_B) \\ &\leq \sum_{S_A\in P_A}\left(\sum_{S_B\in P_B} m_{S_A\times S_B}(f)v(S_B)\right)v(S_A) \end{aligned}$$

Now, for any $x \in S_A$,

$$\sum_{S_B\in P_B} m_{S_A\times S_B}(f)v(S_B) \leq \sum_{S_B\in P_B} m_{S_B}(g_x)v(S_B) \leq \mathbf{L}\int_B g_x = \mathfrak{L}(x)$$

Therefore,

$$L(f,P) \leq L(\mathfrak{L}, P_A)$$

Similarly, we obtain

$$U(f,P) \geq U(\mathfrak{U}, P_A)$$

Consequently, we obtain

$$L(f,P) \leq L(\mathfrak{L},P_A) \leq U(\mathfrak{L},P_A) \leq U(\mathfrak{U},P_A) \leq U(f,P)$$

and we have the desired conclusion. ∎

1. In the case when f is a continuous function over the closed rectangle $A \times B$, we trivially note that g_x is also a continuous function from over B and is therefore integrable. As a result,

   $$\mathbf{L}\int_B g_x = \mathbf{U}\int_B g_x$$

   and we may write

   $$\int_{A\times B} f = \int_A\int_B f(x,y)$$

2. Consider the function:
$$f(x,y) = \begin{cases} \frac{x^2-y^2}{(x^2+y^2)^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$$
and let us attempt to calculate $\int_{[0,1]\times[0,1]} f$.

We have
$$\mathfrak{L}(x) = \mathfrak{U}(x) = \begin{cases} \frac{1}{1+x^2} & x \neq 0 \\ \text{undefined} & x = 0 \end{cases}$$
and similarly,
$$\mathfrak{L}(y) = \mathfrak{U}(y) = \begin{cases} -\frac{1}{1+y^2} & y \neq 0 \\ \text{undefined} & y = 0 \end{cases}$$
Therefore,
$$\int_0^1 \int_0^1 f(x,y)\, dy\, dx = \frac{\pi}{4} \qquad \int_0^1 \int_0^1 f(x,y)\, dx\, dy = -\frac{\pi}{4}$$
As a result, both the integrals exist but the function is not integrable over $[0,1] \times [0,1]$.

The above theorem is best elucidated using an example. Define
$$D = \{(x,y) \in [0,1] \times [0,1] \mid y \leq x^2\}$$
Let us attempt to evaluate $\int_D xy$. First, we must rewrite this as an integral over a rectangle. Let $A = [0,1] \times [0,1]$. Then, we wish to evaluate the integral
$$\int_A xy\chi_D(x,y)$$
Then, $g_x : [0,1] \to \mathbb{R}$ is given by
$$g_x(y) = \begin{cases} xy & y \leq x^2 \\ 0 & y > x^2 \end{cases}$$
It is not hard to see that $g_x(y)$ is integrable. Therefore,
$$\mathfrak{L}(x) = \mathbf{L}\int_{[0,1]} g_x = \mathbf{U}\int_{[0,1]} g_x = \mathfrak{U}(x) = x\int_0^{x^2} y\, dy = \frac{1}{2}x^5$$
Finally,
$$\int_A xy\chi_D(x,y) = \int_{[0,1]} \frac{1}{2}x^5 = \frac{1}{12}$$

2.2 Partitions of Unity

Theorem 2.21. *Let $A \subseteq \mathbb{R}^n$ and let $\mathcal{O}$ be an open cover of A. Then there is a collection Φ of C^∞ functions φ defined in an open set containing A with the following properties*

1. *For each $x \in A$ we have $0 \leq \varphi(x) \leq 1$*
2. *For each $x \in A$ there is an open set V containing x such that all but finitely many $\varphi \in \Phi$ are 0 on V*
3. *For each $x \in A$ we have $\sum_{\varphi \in \Phi} \varphi(x) = 1$*
4. *For each $\varphi \in \Phi$ there is an open set U in $\mathcal{O}$ such that $\varphi = 0$ outside of some closed set contained in U.*

A collection Φ satisfying 1 to 3 is called a C^∞ partition of unity for A. If Φ also satisfies 4, it is said to be subordinate to the cover $\mathcal{O}$.

Proof. ∎

Theorem 2.22 (Change of Variables)**.** *Let $A \subseteq \mathbb{R}^n$ be an **open** set and $g : A \to \mathbb{R}^n$ be a $1-1$, continuously differentiable function such that $\deg g'(x) \neq 0$ for all $x \in A$. If $f : g(A) \to \mathbb{R}$ is integrable, then*

$$\int_{g(A)} f = \int_A (f \circ g) |\det g'|$$

Proof. TODO: Add in proof ∎

Chapter 3

Integration on Chains

Definition 3.1 (Tensor)**.** Let V be an $\mathbb{R}$-vector space. A function $T : V^k \to \mathbb{R}$ is called *multilinear* if for each i with $1 \leq i \leq k$ we have

$$\begin{aligned} T(v_1, \ldots, v_i + v_i', \ldots, v_k) &= T(v_1, \ldots, v_i, \ldots, v_k) + T(v_1, \ldots, v_i', \ldots, v_k) \\ T(v_1, \ldots, av_i, \ldots, v_k) &= aT(v_1, \ldots, v_i, \ldots, v_k) \end{aligned}$$

A multilinear function is called a k-tensor on V and the set of all k-tensors, denoted by $\mathfrak{J}^k(V)$ becomes an $\mathbb{R}$-vector space

Definition 3.2 (Tensor Product)**.** Let V be an $\mathbb{R}$-vector space. For $S \in \mathfrak{J}^k(V)$ and $T \in \mathfrak{J}^l(V)$, we define the *tensor product* $S \otimes T \in \mathfrak{J}^{k+l}(V)$ in the most obvious sense.

Obviously, $S \otimes T \neq T \otimes S$. Furthermore, one notes that $\mathfrak{J}^1(V) = V^*$, the dual space. Recall that for finite dimensional spaces, $\dim V^* = \dim V$.

Theorem 3.3. *Let $v_1, \ldots v_n$ be a basis for V and let $\varphi_1, \ldots \varphi_n$ be the dual basis, $\varphi_i(v_j) = \delta_{ij}$. Then the set of all K-fold tensor products*

$$\varphi_{i_1} \otimes \cdots \otimes \varphi_{i_k} \qquad 1 \leq i_1, \ldots, i_k \leq n$$

is a basis for $\mathfrak{I}^k(v)$ which therefore has dimension n^k.

Proof. Straightforward ∎