Notes for Vector Calculus

Zhao Wenchuan

November 17, 2021

Contents

1	Differentiation		2
	1.1	Differentiable Mapping	2
	1.2	Directional Derivatives	6
	1.3	Mean Value Theorem in Vector Valued Functions	8

Chapter 1.

Differentiation

§1.1 Differentiable Mapping

Observation 1.1.1. Let $f: \mathbb{R}^m \to \mathbb{R}^n$, and denote f_i for the *i*-th factor of f, i.e., $f = \langle f_i \rangle_i^n$. Assume f_i is smooth at a point $\mathbf{p} \in \mathbb{R}^m$.

Intuitively, f_i is smooth at **p** iff there exists a neighbourhood N of **p** and a plane described by $P_i : \mathbb{R}^m \to \mathbb{R}$, such that

$$f_i[N] \approx P_i[N].$$

As \mathbb{R}^m is considered as a metric space, any open ball $B(\mathbf{p}, \delta) \subseteq N$ is also required neighbourhood of \mathbf{p} . In this sense, the approximation can be considered as,

$$\lim_{\delta \to 0} f_i[B(\mathbf{p}, \delta)] = \lim_{\delta \to 0} P_i[B(\mathbf{p}, \delta)].$$

In the term of elements, that is, there exists $\delta \in \mathbb{R}_{>0}$, such that for any $\mathbf{t} \in \mathbb{R}^m$ with $\mathbf{p} + \mathbf{t} \in B(\mathbf{p}, \delta)$,

$$\lim_{\mathbf{t}\to\mathbf{0}_{\mathbb{R}^m}}f_i(\mathbf{p}+\mathbf{t})=\lim_{\mathbf{t}\to\mathbf{0}_{\mathbb{R}^m}}P_i(\mathbf{p}+\mathbf{t}).$$

As P_i describes a plane, it can be considered as a translated linear mapping, and as this plane must be a tangent plane of f_i at \mathbf{p} , there exists a linear mapping $\phi_i : \mathbb{R}^m \to \mathbb{R}$ such that

$$P_i(\mathbf{p} + \mathbf{t}) = \phi_i(\mathbf{p} + \mathbf{t} - \mathbf{p}) + f_i(\mathbf{p}).$$

Thus, we have

$$\lim_{\mathbf{t}\to\mathbf{0}_{\mathbb{P}^m}} f_i(\mathbf{p}+\mathbf{t}) = \lim_{\mathbf{t}\to\mathbf{0}_{\mathbb{P}^m}} \phi_i(\mathbf{t}) + f_i(\mathbf{p}).$$

Rearrange the equation, we have

$$\lim_{\mathbf{t} \to \mathbf{0}_{\mathbb{R}^m}} \frac{f_i(\mathbf{p} + \mathbf{t}) - f_i(\mathbf{p})}{\|\mathbf{t}\|_{\mathbb{R}^m}} = \lim_{\mathbf{t} \to \mathbf{0}_{\mathbb{R}^m}} \frac{\phi_i(\mathbf{t})}{\|\mathbf{t}\|_{\mathbb{R}^m}}.$$

As ϕ_i is linear, the right hand side of the equation is a constant in \mathbb{R} . Thus, by rearrange the equation again, we have

$$\lim_{\mathbf{t} \to \mathbf{0}_{\mathbb{R}^m}} \frac{f_i(\mathbf{p} + \mathbf{t}) - f_i(\mathbf{p}) - \phi_i(\mathbf{t})}{\|\mathbf{t}\|_{\mathbb{R}^m}} = 0$$

The last equation describes the smoothness of f_i at \mathbf{p} , in calculus, f_i is said to be differentiable at \mathbf{p} .

Now, assume for any $i \in \{1, ..., n\}$, f_i is differentiable at **p**. That is, for any f_i , there exists a $\phi_i : \mathbb{R}^m \to \mathbb{R}$, such that

$$\left\langle \lim_{\mathbf{t} \to \mathbf{0}_{\mathbb{R}^m}} \frac{f_i(\mathbf{p} + \mathbf{t}) - f_i(\mathbf{p}) - \phi_i(\mathbf{t})}{\|\mathbf{t}\|_{\mathbb{R}^m}} \right\rangle_{i=1}^n = \langle 0 \rangle_{i=1}^n.$$

By vector sum, we have

$$\lim_{\mathbf{t}\to\mathbf{0}_{\mathbb{R}^m}}\frac{f(\mathbf{p}+\mathbf{t})-f(\mathbf{p})-\phi(\mathbf{t})}{\|\mathbf{t}\|_{\mathbb{R}^m}}=\mathbf{0}_{\mathbb{R}^n},$$

where $\phi = \langle \phi_i \rangle_{i=1}^n$.

In this sense, Definition 1.1.1 is introduced as following.

Definition 1.1.1 (Differentiable Mappings).

Let
$$f: \mathbb{R}^m \to \mathbb{R}^n$$
.

f is said to be differentiable at a point $\mathbf{p} \in \mathbb{R}^m$ iff for any $\mathbf{u} \in \mathbb{R}^m \setminus \{0_{\mathbb{R}^m}\}$, there exists a linear mapping $\phi : \mathbb{R}^m \to \mathbb{R}^n$ such that

$$\lim_{\mathbf{t}\to\mathbf{0}_{\mathbb{R}^m}}\frac{f(\mathbf{p}+\mathbf{t})-f(\mathbf{p})-\phi(\mathbf{t})}{\|\mathbf{t}\|_{\mathbb{R}^m}}=\mathbf{0}_{\mathbb{R}^n}.$$

Theorem 1.1.1. In Definition 1.1.1, ϕ is unique.

Proof. As Definition holds for ϕ , there exists an $\alpha: \mathbb{R}^m \to \mathbb{R}^n$ with

$$\lim_{\mathbf{t}\to\mathbf{0}_{\mathbb{R}^m}}\alpha(\mathbf{t})=\alpha(\mathbf{0}_{\mathbb{R}^m})=\mathbf{0}_{\mathbb{R}^n},$$

and a neighbourhood N of **p** such that for any $\mathbf{t} \in \mathbb{R}^m$ with $\mathbf{p} + \mathbf{t} \in N$,

$$\frac{f(\mathbf{p} + \mathbf{t}) - f(\mathbf{p}) - \phi(\mathbf{t})}{\|\mathbf{t}\|_{\mathbb{R}^m}} = \alpha(\mathbf{t}).$$

Suppose Definition 1.1.1 also holds for another $\lambda : \mathbb{R}^m \to \mathbb{R}^n$, then there exists an $\beta : \mathbb{R}^m \to \mathbb{R}^n$ with

$$\lim_{\mathbf{t}\to\mathbf{0}_{\mathbb{R}^m}}\beta(\mathbf{t})=\beta(\mathbf{0}_{\mathbb{R}^m})=\mathbf{0}_{\mathbb{R}^n},$$

and a neighbourhood N' of \mathbf{p} such that for any $\mathbf{t} \in \mathbb{R}^m$ with $\mathbf{p} + \mathbf{t} \in N'$,

$$\frac{f(\mathbf{p} + \mathbf{t}) - f(\mathbf{p}) - \lambda(\mathbf{t})}{\|\mathbf{t}\|_{\mathbb{R}^m}} = \beta(\mathbf{t}).$$

Let $\gamma = \phi - \lambda$. As ϕ and $-\lambda$ are both linear, γ is also linear. Then, we have

$$\frac{\gamma(\mathbf{t})}{\|\mathbf{t}\|_{\mathbb{R}^m}} = \alpha(\mathbf{t}) - \beta(\mathbf{t})$$

$$\iff \lim_{\mathbf{t} \to \mathbf{0}_{\mathbb{R}^m}} \gamma(\hat{\mathbf{t}}) = \lim_{\mathbf{t} \to \mathbf{0}_{\mathbb{R}^m}} (\alpha(\mathbf{t}) - \beta(\mathbf{t}))$$

$$\iff \gamma(\hat{\mathbf{t}}) = \mathbf{0}_{\mathbb{R}^n}.$$

As \mathbf{t} is arbitrarily picked from $U \cap U'$, and $U \cap U'$ is open in \mathbb{R}^m as U and U' are open, the set $\{\hat{\mathbf{t}} : \mathbf{t} \in U \cap U' - \mathbf{p}\}$ gives all possible directions in \mathbb{R}^m . And, as $\gamma(s\hat{\mathbf{t}}) = \mathbf{0}_{\mathbb{R}^n}$ for any $\mathbf{t} \in \mathbb{R}^m$ and any $s \in \mathbb{R}$, $\gamma[\mathbb{R}^m] = \{\mathbf{0}_{\mathbb{R}^m}\}$. Thus, $\phi = \lambda$.

Theorem 1.1.2. With the condition in Definition 1.1.1, if f is differentiable at \mathbf{p} , then f is continuous at \mathbf{p} .

Proof. As f is differentiable at **p**, there exists an $\alpha: \mathbb{R}^m \to \mathbb{R}^n$ with

$$\lim_{\mathbf{t}\to\mathbf{0}_{\mathbb{R}^m}}\alpha(\mathbf{t})=\alpha(\mathbf{0}_{\mathbb{R}^m})=\mathbf{0}_{\mathbb{R}^m},$$

such that

$$\frac{f(\mathbf{p} + \mathbf{t}) - f(\mathbf{p}) - \phi(\mathbf{t})}{\|\mathbf{t}\|_{\mathbb{R}^m}} = \alpha(\mathbf{t}).$$

By rearranging the equation, we observe

$$\lim_{\mathbf{t} \to \mathbf{0}_{\mathbb{R}^m}} [f(\mathbf{p} + \mathbf{t}) - \phi(\mathbf{t})] = \lim_{\mathbf{t} \to \mathbf{0}_{\mathbb{R}^m}} [\|\mathbf{t}\|_{\mathbb{R}^m} \alpha(\mathbf{t}) + f(\mathbf{p})]$$

$$\iff \lim_{\mathbf{t} \to \mathbf{0}_{\mathbb{R}^m}} f(\mathbf{p} + \mathbf{t}) = f(\mathbf{p}).$$

Thus, f is continuous at \mathbf{p} .

Theorem 1.1.3. With the condition in Definition 1.1.1, let $g: \mathbb{R}^n \to \mathbb{R}^k$.

If f is differentiable at \mathbf{p} and g is differentiable at $f(\mathbf{p})$, then $g \circ f$ is differentiable at \mathbf{p} .

Proof. As f is differentiable at \mathbf{p} , there exists a linear mapping $\phi : \mathbb{R}^m \to \mathbb{R}^n$ and a neighbourhood N of \mathbf{p} such that for any $\mathbf{t} \in \mathbb{R}^m$ with $\mathbf{p} + \mathbf{t} \in \mathbb{R}^m$,

$$f(\mathbf{p}) + \phi(\mathbf{t}) = f(\mathbf{p} + \mathbf{t}) - \|\mathbf{t}\|_{\mathbb{R}^m} \alpha(\mathbf{t}).$$

As g is differentiable at $f(\mathbf{p})$, there exists a linear mapping $\lambda : \mathbb{R}^m \to \mathbb{R}^n$ such that

$$\lim_{\mathbf{t}\to\mathbf{0}_{\mathbb{R}^m}}\frac{g(f(\mathbf{p})+\|\mathbf{t}\|_{\mathbb{R}^m}\phi(\hat{\mathbf{t}}))-f(\mathbf{p})-\lambda(\|\mathbf{t}\|_{\mathbb{R}^m}\phi(\hat{\mathbf{t}}))}{\|\|\mathbf{t}\|_{\mathbb{R}^m}\phi(\hat{\mathbf{t}})\|_{\mathbb{R}^n}}=\mathbf{0}_{\mathbb{R}^k}.$$

As ϕ is linear, we have

$$\|\mathbf{t}\|_{\mathbb{R}^m}\phi(\hat{\mathbf{t}}) = \phi(\mathbf{t}).$$

By scalar multiplication, we have

$$\left\| \|\mathbf{t}\|_{\mathbb{R}^m} \phi(\mathbf{\hat{t}}) \right\|_{\mathbb{R}^n} = \|\mathbf{t}\|_{\mathbb{R}^m} \|\phi(\mathbf{\hat{t}})\|_{\mathbb{R}^n}.$$

Now, we have

$$\begin{split} &\lim_{\mathbf{t}\to\mathbf{0}_{\mathbb{R}^m}}\frac{g(f(\mathbf{p})+\phi(\mathbf{t}))-g(f(\mathbf{p}))-\lambda(\phi(\mathbf{t}))}{\|\mathbf{t}\|_{\mathbb{R}^m}}=\mathbf{0}_{\mathbb{R}^k}\\ &\iff &\lim_{\mathbf{t}\to\mathbf{0}_{\mathbb{R}^m}}\frac{g(f(\mathbf{p}+\mathbf{t}))-g(f(\mathbf{p}))-(\lambda\circ\phi)(\mathbf{t})}{\|\mathbf{t}\|_{\mathbb{R}^m}}=\mathbf{0}_{\mathbb{R}^k}. \end{split}$$

As λ and ϕ are both linear, $\lambda \circ \phi$ are also linear.

By Definition 1.1.1, $g \circ f$ is differentiable at **p**.

Observation 1.2.1. Let $f: \mathbb{R}^m \to \mathbb{R}^n$, and let $g: \mathbb{R} \to \mathbb{R}^m$ be defined as

$$g(t) := \mathbf{p} + t\mathbf{u},$$

where $\mathbf{p}, \mathbf{u} \in \mathbb{R}^m$ and $\mathbf{u} \neq \mathbf{0}_{\mathbb{R}^m}$.

Let $h = f \circ g$ and define $h' : D_{h'} \subseteq \mathbb{R} \to \mathbb{R}^n$ as

$$h'(t) := \lim_{t \to t_0} \frac{h(t) - h(t_0)}{t - t_0},$$

where for any $t \in D_{h'}$, the this limit exists in \mathbb{R}^n . Thus,

$$h'(0) = \lim_{t \to 0} \frac{f(\mathbf{p} + t\mathbf{u}) - f(\mathbf{p})}{t}$$

describes the instantaneous rate of change of f along the straight line $\{\mathbf{p} + t\mathbf{u} : t \in \mathbb{R}\}$ with $\|\mathbf{u}\|_{\mathbb{R}^m}$ as the unit length. h'(0) is so-called the **u**-directional derivative of f at \mathbf{p} (See Definition 1.2.1).

Definition 1.2.1 (Directional Derivatives).

Let $f: \mathbb{R}^m \to \mathbb{R}^n$. Let $\mathbf{u} \in \mathbb{R}^m \setminus \{\mathbf{0}_{\mathbb{R}^m}\}$ and let $\mathbf{p} \in \mathbb{R}^m$. The \mathbf{u} -directional derivative of f at \mathbf{p} is defined as

$$\nabla_{\mathbf{u}} f(\mathbf{p}) := \lim_{t \to 0} \frac{f(\mathbf{p} + t\mathbf{u}) - f(\mathbf{p})}{t},$$

if the limit exists in \mathbb{R}^n .

Theorem 1.2.1. With the condition in Definition 1.2.1, for any $s \in \mathbb{R} \setminus \{0\}$,

$$\nabla_{\mathbf{s}\mathbf{u}}f(\mathbf{p}) = s\nabla_{\mathbf{u}}f(\mathbf{p}).$$

Proof. Let $\theta = ts^{-1}$, then, by Definition 1.2.1, we have

$$s\nabla_{\mathbf{u}}f(\mathbf{p}) = s \lim_{t \to 0} \frac{f(\mathbf{p} + t\mathbf{u}) - f(\mathbf{p})}{t}$$
$$= \lim_{t \to 0} \frac{f(\mathbf{p} + t\mathbf{u}) - f(\mathbf{p})}{ts^{-1}}$$
$$= \lim_{t \to 0} \frac{f(\mathbf{p} + t\mathbf{u}) - f(\mathbf{p})}{\theta}$$
$$= \nabla_{s\mathbf{u}}f(\mathbf{p}).$$

Theorem 1.2.2. With the condition in Definition 1.2.1, if $\nabla_{\mathbf{u}} f(\mathbf{p})$ exists, then there exists an open subset $U \subseteq \mathbb{R}^m$ with $\mathbf{p} \in U$ such that f is relative continuous on the line described by $\mathbf{p} + t\mathbf{u}$ for some $t \in \mathbb{R}$.

Proof. Let U be an open subset of \mathbb{R}^m , and let $g: \mathbb{R} \to \mathbb{R}^m$ be defined as

$$g(t) := \mathbf{p} + t\mathbf{u}.$$

Then f is relative continuous on the line defined by $\mathbf{p} + t\mathbf{u}$ for some $t \in \mathbb{R}$ iff $f \upharpoonright_{q[\mathbb{R}]}$ is continuous on $U \cap g[\mathbb{R}]$.

Let $h = f \circ g$, then

$$\nabla_{\mathbf{u}} f(\mathbf{p}) = \lim_{t \to 0} \frac{h(t) - h(0)}{t} = \mathbf{v} \in \mathbb{R}^n.$$

Then, there exists an $\alpha : \mathbb{R} \to \mathbb{R}^n$ with $\alpha(t) \to \mathbf{0}_{\mathbb{R}^n}$ as $t \to 0$, such that there exists an open subset $I \subseteq \mathbb{R}$ with $0 \in I$, such that for any $t \in I$,

$$h(t) = t\mathbf{v} + t\alpha(t) + h(0).$$

Then we have

$$\lim_{t \to 0} h(t) = \lim_{t \to 0} [t\mathbf{v} + t\alpha(t) + h(0)]$$

$$\iff \lim_{t \to 0} h(t) = h(0).$$

Thus, h is continuous at 0.

As it is easy to show g is bijective, $g \circ g^{-1}$ is an identity mapping on $g[\mathbb{R}] \subseteq \mathbb{R}^m$. As composition of mappings is associative, we have

$$h = f \circ g \iff h \circ g^{-1} = f \circ g \circ g^{-1}$$
$$\iff h \circ g^{-1} = f \circ (g \circ g^{-1})$$
$$\iff h \circ g^{-1} = f \upharpoonright_{g[\mathbb{R}]}.$$

It is also easy to find that g^{-1} is continuous everywhere, thus, as h is continuous at 0, $f \upharpoonright_{g[\mathbb{R}]}$ is continuous on $U \cap g[\mathbb{R}]$. Thus, f is relative continuous on the line defined by $\mathbf{p} + t\mathbf{u}$ for some $t \in \mathbb{R}$.

§1.3 Mean Value Theorem in Vector Valued Functions

Theorem 1.3.1 (Mean Value Theorem on $\mathbb{R}^m \to \mathbb{R}^n$).

Let $f: \mathbb{R}^m \to \mathbb{R}^n$. Let $\mathbf{p}, \mathbf{q} \in \mathbb{R}^m$, for convenience, let $g: \mathbb{R} \to \mathbb{R}^m$ be defined as

$$g(t) := \mathbf{p} + t(\mathbf{q} - \mathbf{p}).$$

If $f \upharpoonright_{g[\mathbb{R}]}$ is continuous on g[(0,1)], and differentiable on g[[0,1]], then

$$||f(\mathbf{q}) - f(\mathbf{p})||_{\mathbb{R}^n} \le \sup_{\mathbf{x} \in g[(a,b)]} ||\nabla_{\mathbf{u}} f(\mathbf{x})||_{\mathbb{R}^n}.$$

Lemma 1.3.1. Let $f : \mathbb{R} \to \mathbb{R}$, and let $a, b \in \mathbb{R}$ with a < b. Suppose f is continuous on [a, b] and differentiable on (a, b), and $0 \notin f[(a, b)]$.

Then, f is monotonic.

Lemma 1.3.2 (Rolle's Theorem).

Let $f: \mathbb{R}^m \to \mathbb{R}^n$. Let $\mathbf{a}, \mathbf{b} \in \mathbb{R}^m$ with $f(\mathbf{a}) = f(\mathbf{b})$. Suppose f is relative continuous on $\ell[\mathbf{a}, \mathbf{b}]$, and relative differentiable on $\ell(\mathbf{a}, \mathbf{b})$.

Then, there exits $\mathbf{c} \in \ell(\mathbf{a}, \mathbf{b})$ such that $\nabla_{\mathbf{u}} f(\mathbf{c}) = \mathbf{0}_{\mathbb{R}^n}$, where $\mathbf{u} = \mathbf{b} - \mathbf{a}$.

Proof. First, consider $f = \langle f_i \rangle_{i=1}^n$.

Suppose for any $\mathbf{c} \in \ell(\mathbf{a}, \mathbf{b}), \ \nabla_{\mathbf{u}} f(\mathbf{c}) \neq \mathbf{0}_{\mathbb{R}^n}$, then there exists $i \in \{1, \ldots, n\}$ such that $\nabla_{\mathbf{u}} f_i(\mathbf{c}) \neq 0$.

Let $g: \mathbb{R} \to \mathbb{R}^m$ be defined as

$$g(t) = \mathbf{b} - t\mathbf{a},$$

and let $h_i = f_i \circ g$. Then, for any $t \in (0,1)$, $h'_i(t) \neq 0$.

As f_i is differentiable on g[(0,1)], and g is differentiable on (0,1), by Theorem 1.1.3, h_i is differentiable on (0,1). In this case, $0 \notin h'_i[(0,1)]$ implies h_i is monotonic (Lemma 1.3.1). This implies

$$h_i(0) = f_i(\mathbf{a}) \neq f_i(\mathbf{b}) = h_i(1).$$

As $f(\mathbf{a}) = \langle f_i(\mathbf{a}) \rangle_{i=1}^n$ and $f(\mathbf{b}) = \langle f_i(\mathbf{b}) \rangle_{i=1}^n$, we have $f(\mathbf{a}) \neq f(\mathbf{b})$. This contradicts the assumption that $f(\mathbf{a}) = f(\mathbf{b})$.

Thus, there has to be a $\mathbf{c} \in \ell(\mathbf{a}, \mathbf{b})$ such that $\nabla_{\mathbf{u}} f_i(\mathbf{c})$.

Change into $f \circ g$ as $g : \mathbb{R} \to \mathbb{R}^m$ below.

Lemma 1.3.3 (Cauchy Mean Value Theorem).

Let $f, g : \mathbb{R}^m \to \mathbb{R}^n$ where f and g are differentiable on an open subset $U \subseteq \mathbb{R}^m$, and continuous on U^- . Suppose $\mathbf{0}_{\mathbb{R}^n} \notin \nabla_{\mathbf{u}} g[U]$ for any $\mathbf{u} \in \mathbb{R}^m$.

Then, for any $\mathbf{p}, \mathbf{q} \in U$ with $\mathbf{p} \neq \mathbf{q}$, there exits a **c** such that

$$\frac{\nabla_{\mathbf{u}} f(\mathbf{c})}{\|\nabla_{\mathbf{u}} g(\mathbf{c})\|_{\mathbb{R}^n}} = \frac{f(\mathbf{b}) - f(\mathbf{a})}{\|g(\mathbf{b}) - g(\mathbf{a})\|}.$$

Proof.

Lemma 1.3.4. Let $f: \mathbb{R} \to \mathbb{R}^n$. Let I = (a, b) be an open interval of \mathbb{R} . If f is continuous on I and differentiable on I^- , then there exists a $c \in I$ such that

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Proof. As f is differentiable on I, by Cauchy Mean Value Theorem, there exists a

Proof for Theorem 1.3.1. Let $h = f \circ g$. As f is continuous on g[(0,1)] and g is continuous everywhere, h is continuous on (0,1). By Theorem 1.1.3, as f is differentiable on g[[0,1]] and g is differentiable on [0,1], then, by Theorem 1.1.3, h is differentiable on [0,1].

Let $h':D\subseteq\mathbb{R}\to\mathbb{R}^n$ be defined as

$$h'(t) := \lim_{t \to 0} \frac{h(c+t) - h(t)}{t},$$

where D is the set of all points in \mathbb{R} such that the limit exists in \mathbb{R}^n .

By Lemma 1.3.4, there exists a $c \in (0,1)$ such that

$$h'(c) = \frac{h(1) - h(0)}{1 - 0}.$$

10