

12 Differentiability in one variable

Definition 12.1. (differentiability)

Let $f : [a, b] \rightarrow \mathbb{R}$ be a function and $x \in [a, b]$.

The function f' is defined by

$$f'(x) = \begin{cases} \lim_{t \rightarrow x} \frac{f(t) - f(x)}{t - x}, & \text{if } x \in (a, b), \\ \lim_{t \rightarrow x^+} \frac{f(t) - f(x)}{t - x}, & \text{if } x = a, \\ \lim_{t \rightarrow x^-} \frac{f(t) - f(x)}{t - x}, & \text{if } x = b, \end{cases} \quad \text{provided the limit exists.}$$

More precisely, the latter two are often denoted by $f'(a^+)$ and $f'(b^-)$ respectively.

The domain of f' is the set of points x at which the limit exists.

The function f' is called the **derivative of f** .

If f' is defined at x , we say that f is **differentiable at x** .

If f is differentiable at every point of a set $E \subseteq [a, b]$,

we say that f is **differentiable on E** .

Definition 12.2. (higher order derivatives)

If f has a derivative f' on an interval, and f' is differentiable,

we denote the derivative of f' by f'' or $f^{(2)}$ and call f'' the **second derivative of f** .

By the same way, we can define functions $f^{(2)}, f^{(3)}, \dots, f^{(n-1)}, f^{(n)}$

by each of which is the derivative of the previous one.

We call $f^{(n)}$ the **n -th derivative of f** .

Theorem 12.3. (little o)

Let $f : [a, b] \rightarrow \mathbb{R}$ be a function and $x \in [a, b]$.

Then f is differentiable at $x \in (a, b)$, $x = a$, or $x = b$

\iff there exists $m \in \mathbb{R}$ such that $f(x+h) = f(x) + mh + o(h)$ for all h close to $0, 0^+$, or 0^- , in **Landau notation**,

where $o(h)$, called **little o**, is a real-valued function of h which satisfies $\lim_{h \rightarrow 0, 0^+, \text{ or } 0^-} \frac{o(h)}{h} = 0$, respectively;

in this case, $f'(x) = m$.

Theorem 12.4. (differentiability implies continuity)

Let $f : [a, b] \rightarrow \mathbb{R}$ be a function.

If f is differentiable at $x \in [a, b]$, then f is continuous at x .

Theorem 12.5. (basic rules of differentiability)

Let $f, g : [a, b] \rightarrow \mathbb{R}$ be functions and let $x \in [a, b]$.

If f, g are differentiable at x , then $f + g$, fg , and $\frac{f}{g}$ are differentiable at x , and

1. **the sum and difference rules:** $(f \pm g)'(x) = f'(x) \pm g'(x)$;
2. **the product rule:** $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$; and
3. **the quotient rule:** $\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g^2(x)}$ provided $g(x) \neq 0$.

Example 12.6. (basic differentiable functions)

Every polynomial is differentiable and so is every rational function,

except at the points where the denominator is zero.

Theorem 12.7 (chain rule).

Let $g : [a, b] \rightarrow \mathbb{R}$ be a function, let I be an interval with $I \supset g([a, b])$,

and let $f : I \rightarrow \mathbb{R}$ be a function.

If g is differentiable at $x \in [a, b]$ and f is differentiable at $g(x)$,

then $f \circ g$ is differentiable at x and $(f \circ g)'(x) = f'(g(x))g'(x)$.

Proposition 12.8. (for exercise)

Let (a, b) be an interval in \mathbb{R} and $f : (a, b) \rightarrow \mathbb{R}$ be a differentiable function.

1. If $\lim_{x \rightarrow a^+} f'(x)$ exists, as a finite number, and $a \in \mathbb{R}$, then $\lim_{x \rightarrow a^+} f(x)$ exists.
2. If, in addition to item 1, f is continuous on $[a, b)$, then f is differentiable at a and $f'(a^+) = \lim_{x \rightarrow a^+} f'(x)$.
3. If, in addition to item 1, f is continuous on $(a, b]$, then f is uniformly continuous on $(a, b]$.
4. In item 1, the condition $a \in \mathbb{R}$ is needed, see the example $f(x) = \sin \sqrt{-x}$ on $(-\infty, 0)$.
In item 3, the function might have unbounded derivative, for instance $f(x) = \sqrt{-x}$ on $(-1, 0]$.

Introduction to Real Analysis

13 The generalized MVT, Rolle's theorem and the monotonicity theorem

Definition 13.1. (local extreme)

Let (X, d) be a metric space and $f : X \rightarrow \mathbb{R}$ be a function.

We say that f has a **local maximum** (resp. **local minimum**) at $p \in X$

if there exists $\delta > 0$ such that

if $x \in X$ with $d(x, p) < \delta$ then $f(x) \leq f(p)$ (resp. $f(x) \geq f(p)$).

Theorem 13.2 (the critical point theorem in one variable).

Let $f : [a, b] \rightarrow \mathbb{R}$ be a function, where $a < b$.

If f is differentiable at $p \in (a, b)$, and f has a local maximum or a local minimum at p , then $f'(p) = 0$.

Theorem 13.3 (Rolle's theorem).

Let $f : [a, b] \rightarrow \mathbb{R}$ be function, where $a < b$.

1. If f is continuous on $[a, b]$ and differentiable on (a, b) , and $f(a) = f(b)$ then there exists a point $c \in (a, b)$ such that $f'(c) = 0$.
2. If $f^{(n)}$ is differentiable on $[a, b]$, $f(a) = f(b)$, and $f^{(i)}(a) = 0$ for all $1 \leq i \leq n$, then there exists a point $c \in (a, b)$ such that $f^{(n+1)}(c) = 0$.

Theorem 13.4 (the generalized mean value theorem).

Let $f, g : [a, b] \rightarrow \mathbb{R}$ be functions, where $a < b$.

If f and g are continuous on $[a, b]$ and differentiable on (a, b) ,

then there exists a point $c \in (a, b)$ such that $(f(b) - f(a))g'(c) = f'(c)(g(b) - g(a))$.

The special case when $g(x) = x$ is referred to as the **mean value theorem**.

Definition 13.5. (monotonicity of functions)

Let $f : A \subseteq \mathbb{R} \rightarrow \mathbb{R}$ be a function.

1. We say that f is **increasing** (resp. **strictly increasing**) on A if $f(x_1) \leq f(x_2)$ (resp. $f(x_1) < f(x_2)$) for all $x_1, x_2 \in A$ with $x_1 < x_2$.
2. We say that f is **decreasing** (resp. **strictly decreasing**) on A if $f(x_1) \geq f(x_2)$ (resp. $f(x_1) > f(x_2)$) for all $x_1, x_2 \in A$ with $x_1 < x_2$.
3. We say that f is **monotonic** (resp. **strictly monotonic**) on A if f is either increasing or decreasing (resp. either strictly increasing or strictly decreasing) on A .

Theorem 13.6 (the monotonicity theorem).

Let $f : (a, b) \rightarrow \mathbb{R}$ be a differentiable function.

1. If there exists $M > 0$ such that $|f'(x)| \leq M$ for all $x \in (a, b)$, then $|f(x_1) - f(x_2)| \leq M|x_1 - x_2|$ for all $x_1, x_2 \in (a, b)$.
2. If $f'(x) = 0$ for all $x \in (a, b)$, then f is constant on (a, b) .
3. If $f'(x) \geq 0$ (resp. $f'(x) > 0$) for all $x \in (a, b)$, then f is increasing (resp. strictly increasing) on (a, b) .
4. If $f'(x) \leq 0$ (resp. $f'(x) < 0$) for all $x \in (a, b)$, then f is decreasing (resp. strictly decreasing) on (a, b) .
5. If $f'(x) > 0$ for all $x \in (a, b)$ except finitely many points at which $f'(x) = 0$, then f is strictly increasing on (a, b) . (exercise)

Proposition 13.7.

There exists a non-constant function from \mathbb{R} to \mathbb{R}

such that $f^{(n)}(0) = 0$ and $f^{(n)}(x)$ exists for all $x \in \mathbb{R}$ and for all $n \in \mathbb{N}$. (exercise)

Theorem 13.8 (universal chord theorem).

Let $f : [a, b] \rightarrow \mathbb{R}$ be continuous such that $f(a) = f(b)$.

Then for any number $n \in \mathbb{N}$, there exists $c \in [a, b]$ such that $f(c) = f(c + \frac{b-a}{n})$.

Moreover, the result is not true if $n \notin \mathbb{Z}$.

Introduction to Real Analysis

14 L'Hôpital's rule in one variable and Darboux's theorem

Theorem 14.1 (L'Hôpital's rule).

Let $f, g : (a, b) \rightarrow \mathbb{R}$ be differentiable on $(a, b) \setminus \{c\}$ satisfying that $g'(x) \neq 0$ for all $x \in (a, b) \setminus \{c\}$, where $-\infty \leq a < b \leq +\infty$ and $a \leq c \leq b$; here we use the ordering $<$ on the extended real line.

If $\frac{f(x)}{g(x)}$ has the **indeterminate form** $\frac{0}{0}$ or $\frac{\infty}{\infty}$ at $x = c^+$,

i.e., $\lim_{x \rightarrow c^+} f(x) = \lim_{x \rightarrow c^+} g(x) = 0$ or $\lim_{x \rightarrow c^+} g(x) = \pm\infty$ (in this case, we do not need to control $\lim_{x \rightarrow c^+} f(x)$),

and $\lim_{x \rightarrow c^+} \frac{f'(x)}{g'(x)}$ exists or is $\pm\infty$,

then $\lim_{x \rightarrow c^+} \frac{f(x)}{g(x)} = \lim_{x \rightarrow c^+} \frac{f'(x)}{g'(x)}$.

Notice that the result may fail if the latter condition is not satisfied.

The analogous statement is true if $x \rightarrow c^-$, $x \rightarrow +\infty$, or $x \rightarrow -\infty$.

Corollary 14.2.

For any $a, b > 0$ we have

$$\lim_{x \rightarrow +\infty} \frac{x^a}{e^{bx}} = \lim_{x \rightarrow +\infty} \frac{(\ln x)^b}{x^a} = \lim_{x \rightarrow 0^+} \frac{(\ln x)^b}{x^{-a}} = 0 \text{ and } \lim_{x \rightarrow 0} (1 \pm ax)^{\frac{\pm b}{x}} = e^{\operatorname{sgn}((\pm a)(\pm b))ab} \text{ (exercise).}$$

Remark 14.3.

There is a function f on \mathbb{R} such that f is differentiable on \mathbb{R} but f' is not continuous at some point. (exercise)

Theorem 14.4 (Darboux's theorem or intermediate value property for derivatives).

Let $f : [a, b] \rightarrow \mathbb{R}$ be a differentiable function, where $a < b$.

(Notice that f' might not be continuous on $[a, b]$.)

If $s \in \mathbb{R}$ with $f'(a) < s < f'(b)$ or $f'(a) > s > f'(b)$,

then there exists $c \in (a, b)$ such that $f'(c) = s$. (exercise)

Definition 14.5 (little o and big O).

Let $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ be a function and $g : \mathbb{R}^m \rightarrow \mathbb{R}^k$ be a function, and let $p \in \mathbb{R}^m$.

We say that f is **little o of g** at p , and write $f(x) = o(g(x))$ as $x \rightarrow p$,

if for any $M > 0$ there exists $\delta > 0$ such that $\|f(x)\| \leq M\|g(x)\|$ for all x with $\|x - p\| < \delta$,

(equivalently, $\lim_{x \rightarrow p} \frac{f(x)}{\|g(x)\|} = 0$ provided that $g(x) \neq 0$).

We say that f is **big O of g** at p , and write $f(x) = O(g(x))$ as $x \rightarrow p$,

if there exist $M > 0$ and $\delta > 0$ such that $\|f(x)\| \leq M\|g(x)\|$ for all x with $\|x - p\| < \delta$.

If $m = 1$, then the similar definitions work for $p = \pm\infty$.

15 Differentiability of vector-valued functions

Definition 15.1. (differentiability of vector-valued functions)

Let $f : [a, b] \rightarrow \mathbb{R}^k$ be a vector-valued function and $x \in [a, b]$.

The function f' , still vector-valued, is defined by

$$f'(x) = \begin{cases} \lim_{t \rightarrow x} \frac{f(t) - f(x)}{t - x}, & \text{if } x \in (a, b), \\ \lim_{t \rightarrow x^+} \frac{f(t) - f(x)}{t - x}, & \text{if } x = a, \\ \lim_{t \rightarrow x^-} \frac{f(t) - f(x)}{t - x}, & \text{if } x = b, \end{cases} \quad \text{provided the limit exists.}$$

One can define $f'(x)$ to be the vector such that $f(x+h) = f(x) + f'(x)h + o(h)$

for all h near 0 in \mathbb{R} , where $o(h)$ is a vector-valued function with $\lim_{h \rightarrow 0} \frac{o(h)}{h} = 0$.

The domain of f' is the set of points x at which the limit exists.

The function f' is called the **derivative of f** .

If f' is defined at x , we say that f is **differentiable at x** .

If f is differentiable at every point of a set $E \subseteq [a, b]$,

we say that f is **differentiable on E** .

Higher-order derivatives are defined in a similar way.

Remark 15.2. (differentiability of components)

Let $f = (f_1, \dots, f_k)$. Then f is differentiable at x

\iff each f_i is differentiable at x for $1 \leq i \leq k$; in this case $f'(x) = (f'_1(x), \dots, f'_k(x))$.

Proposition 15.3. (derivative of scalar product, inner product and cross product)

Let $f, g : \mathbb{R} \rightarrow \mathbb{R}^k$ and $\alpha : \mathbb{R} \rightarrow \mathbb{R}$ be differentiable. Then the followings hold (exercise).

1. $(\alpha f)' = \alpha' f + \alpha f'$;
2. $(f \cdot g)' = f' \cdot g + f \cdot g'$;
3. if $k = 3$, then $(f \times g)' = f' \times g + f \times g'$.

Remark 15.4. (tangent line and others)

Let $f : \mathbb{R} \rightarrow \mathbb{R}^k$ be differentiable.

1. If $f(t)$ is the position of a particle moving in \mathbb{R}^k at time t , then $f'(t)$ represents the **velocity** of the particle at time t .
2. If $f(t) = a + t(b - a)$, where $a, b \in \mathbb{R}^k$, then f gives a parametric representation of the **straight line** through the points a and b in \mathbb{R}^k .
3. If f gives a parametric representation of a curve in \mathbb{R}^k and $f'(a) \neq 0$, where $a \in \mathbb{R}$ and 0 is the zero vector, then the function $g(t) = f(a) + tf'(a)$ gives a parametric representation of the **tangent line** to the curve at the point $f(a)$.
If $f'(a) = 0$, then the curve may not have a tangent line at $f(a)$.

Proposition 15.5.

Let $f : \mathbb{R} \rightarrow \mathbb{R}^k$ be differentiable.

1. There are vector-valued functions for which the ordinary mean value theorem dose not hold.
2. If there exists $M > 0$ such that $\|f'(x)\| \leq M$ for all $x \in [a, b]$, then $\|f(b) - f(a)\| \leq M|b - a|$.
3. Later, we will have **variant versions of the mean value theorem** in Theorems 19.1 (for $f : \mathbb{R}^k \rightarrow \mathbb{R}$) and Theorem 27.1 (for $f : \mathbb{R}^k \rightarrow \mathbb{R}^m$).

16 Differentiability in several variables and gradient

Definition 16.1. (partial derivatives)

Let $D \subseteq \mathbb{R}^k$ be open and $f : D \rightarrow \mathbb{R}$ be real-valued.

For any $1 \leq i \leq k$, the *i-th partial derivative of f at a point $a = (a_1, a_2, \dots, a_k) \in D$* is defined to be the limit

$$\partial_i f(a) = \lim_{h \rightarrow 0, h \in \mathbb{R}} \frac{f(a_1, \dots, a_{i-1}, a_i + h, a_{i+1}, \dots, a_k) - f(a_1, \dots, a_k)}{h} \quad \text{provided the limit exists.}$$

Other notations for partial derivatives are

$$\partial_i f(a) = \frac{\partial f(a)}{\partial x_i} = \frac{\partial f}{\partial x_i}(a) = \frac{\partial}{\partial x_i} f(a) = \partial_{x_i} f(a) = f_{x_i}(a) = f_i(a).$$

We say that f is **differentiable** at $a \in D$

if there exists a vector $m \in \mathbb{R}^k$ such that $\lim_{h \rightarrow 0, h \in \mathbb{R}^k} \frac{f(a+h) - f(a) - m \cdot h}{\|h\|} = 0$, where \cdot denotes the inner product.

In this case, the vector m is uniquely determined, and

we call it the **gradient** of f at a and denote it by $\nabla f(a)$.

One can also define $\nabla f(a)$ to be the vector

such that $f(a+h) = f(a) + \nabla f(a) \cdot h + o(h)$ for all vectors h close to 0 in \mathbb{R}^k ,

where $o(h)$ is a real-valued function of h with $\lim_{h \rightarrow 0, h \in \mathbb{R}^k} \frac{o(h)}{\|h\|} = 0$ and is said to be **little o** .

Remark 16.2. (gradient, tangent plane, and tangent hyperplane)

Let $f : \mathbb{R}^k \rightarrow \mathbb{R}$ be differentiable at $a \in \mathbb{R}^k$. In Corollary 17.5 and Theorem 18.4, we will have the followings.

1. The negative gradient $-\nabla f(a)$ gives a vector in the domain pointing in the direction of the water flow (the steepest descent) at the point, which is also orthogonal to the level curve at the point, and its magnitude $\|\nabla f(a)\|$ gives the speed (the steepness) of the water flow at that point; while the directional derivative gives the steepness in a given direction.
2. If $k = 2$, then $z = f(x)$ represents a **surface** in \mathbb{R}^3 and $z = f(a) + \nabla f(a) \cdot (x - a)$ represents the **tangent plane** to the surface $z = f(x)$ at $x = a$, that is, the plane $\{(x, z) \in \mathbb{R}^2 \times \mathbb{R} : (-\nabla f(a), 1) \cdot ((x, z) - (a, f(a))) = 0\}$ in \mathbb{R}^3 through the point $(a, f(a))$ with the normal vector $(-\nabla f(a), 1)$.
3. If $k > 2$, then $z = f(x)$ represents a **hypersurface** in \mathbb{R}^{k+1} and $z = f(a) + \nabla f(a) \cdot (x - a)$ represents the **tangent hyperplane** to the hypersurface $z = f(x)$ at $x = a$.

Theorem 16.3. (differentiability, continuity, and partial derivatives)

Let $f : \mathbb{R}^k \rightarrow \mathbb{R}$ be a function and $a \in \mathbb{R}^k$.

If f is differentiable at a ,

then f is continuous at a , the partial derivatives $\partial_i f(a)$ all exists for $1 \leq i \leq k$,

and $\nabla f(a) = (\partial_1 f(a), \dots, \partial_k f(a))$.

Theorem 16.4. ($C^1 \Rightarrow$ differentiable \Rightarrow partial derivatives exist)

Let $S \subseteq \mathbb{R}^k$ be open, $a \in S$, and $f : S \rightarrow \mathbb{R}$.

1. If the partial derivatives $\partial_i f(x)$ all exist for all x in some neighborhood of a and all are continuous at a (in this case we say that f is C^1 at a or $f \in C^1(\{a\})$), then f is differentiable at a and $\nabla f(a) = (\partial_1 f(a), \dots, \partial_k f(a))$.
2. If the partial derivatives $\partial_i f(x)$ all exist and are bounded for all $x \in S$, then f is continuous on S ; however, it is possible that f is not differentiable at a point in S when $k \geq 2$. (exercise)

17 Differential and directional derivative

Definition 17.1.

Let $f : \mathbb{R}^k \rightarrow \mathbb{R}$ be differentiable and $a \in \mathbb{R}^k$.

Let $h = (h_1, h_2, \dots, h_k)$ be the given increment from a to $a + h$.

Then the **increment** of f at a is given by $\Delta f(a; h) := f(a + h) - f(a) = \nabla f(a) \cdot h + o(h)$,

which represents the changes in the value of f when a changes to $a + h$.

If the increment $h = (dx_1, dx_2, \dots, dx_k)$ is infinitesimal,

then the increment of f at a can be linearly approximated by $\nabla f(a) \cdot h$ by neglecting the error term $o(h)$.

This leads us to define the **differential** of f at x as $df(x) := \nabla f(x) \cdot (dx_1, dx_2, \dots, dx_k) = \partial_1 f(x) dx_1 + \dots + \partial_k f(x) dx_k$.

As applications, the following questions can be answered about a circular cylinder with base radius r and height h .

(i) About (i.e., approximately) how much does the volume increase if (r, h) is increased from $(1, 2)$ to $(1.02, 2.01)$?

(ii) About how much should r be decreased from 1 to keep the volume constant, if h is increased from 2 to 2.01?

Remark 17.2. (sum, product, quotient rules of differentials)

Let $f, g : \mathbb{R}^k \rightarrow \mathbb{R}$ be differentiable.

Then $d(f \pm g) = df \pm dg$, $d(fg) = f dg + g df$,

and $d\left(\frac{f}{g}\right) = \frac{g df - f dg}{g^2}$ at which g does not vanish.

Definition 17.3. (directional derivative)

Let $f : \mathbb{R}^k \rightarrow \mathbb{R}$ be a function and $a \in \mathbb{R}^k$ be a point.

Let u be a **unit vector** in \mathbb{R}^k , that is, $\|u\| = 1$.

The (**two-sided**) **directional derivative of f at a in the direction of u** is defined to be

$$\partial_u f(a) = \left. \frac{d}{dt} f(a + tu) \right|_{t=0} = \lim_{t \rightarrow 0, t \in \mathbb{R}} \frac{f(a + tu) - f(a)}{t} \text{ provided the limit exists.}$$

Theorem 17.4. (differentiability and directional derivative)

Let $f : \mathbb{R}^k \rightarrow \mathbb{R}$ be differentiable at $a \in \mathbb{R}^k$.

Then the directional derivative $\partial_u f(a)$ exists for all unit vectors $u \in \mathbb{R}^k$,

and $\partial_u f(a) = \nabla f(a) \cdot u$.

But the existence of all directional derivatives does not imply differentiability. (exercise)

Corollary 17.5. (gradient and extreme values of directional derivatives)

Let $f : \mathbb{R}^k \rightarrow \mathbb{R}$ be differentiable at the point $a \in \mathbb{R}^k$ where $\nabla f(a) \neq 0$,

then the maximal (resp. minimal) value of $\partial_u f(a)$ over all unit vectors u

is $\|\nabla f(a)\|$ (resp. $-\|\nabla f(a)\|$) and it occurs when $u = \frac{\nabla f(a)}{\|\nabla f(a)\|}$ (resp. $u = -\frac{\nabla f(a)}{\|\nabla f(a)\|}$).

18 The chain rule

Theorem 18.1 (the chain rule for one variable).

Let $g: \mathbb{R} \rightarrow \mathbb{R}^k$ be a differentiable at $a \in \mathbb{R}$ and

let $f: \mathbb{R}^k \rightarrow \mathbb{R}$ be a differentiable at $g(a) \in \mathbb{R}^k$.

Then the composition function $f \circ g$ is differentiable at a and $(f \circ g)'(a) = \nabla f(g(a)) \cdot g'(a)$.

In Leibniz's notation, with $t \mapsto z(x_1(t), x_2(t), \dots, x_k(t))$,

we have $\frac{dz}{dt}(a) = \sum_{i=1}^k \frac{\partial z}{\partial x_i}(x_1(a), x_2(a), \dots, x_k(a)) \frac{dx_i}{dt}(a)$.

Writing explicit variables is to avoid confusion of z' and $\partial_i z$ on both sides.

Theorem 18.2 (the chain rule for several variables).

Let $g = (g_1, \dots, g_k): \mathbb{R}^m \rightarrow \mathbb{R}^k$ be a function such that

for each $1 \leq i \leq k$, the map $g_i: \mathbb{R}^m \rightarrow \mathbb{R}$ is differentiable at $a \in \mathbb{R}^m$ (resp. of C^1 near a),

and let $f: \mathbb{R}^k \rightarrow \mathbb{R}$ be a differentiable at $g(a) \in \mathbb{R}^k$ (resp. of C^1 near $g(a)$).

Then the composition function $f \circ g$ is differentiable at a (resp. of C^1 near a)

and $\partial_j(f \circ g)(a) = \nabla f(g(a)) \cdot (\partial_j g_1(a), \dots, \partial_j g_k(a))$ for all $1 \leq j \leq m$.

In Leibniz's notation, with $(t_1, \dots, t_m) \mapsto z(x_1(t_1, \dots, t_m), \dots, x_k(t_1, \dots, t_m))$,

we have $\frac{\partial z}{\partial t_j}(a) = \sum_{i=1}^k \frac{\partial z}{\partial x_i}(x_1(a), \dots, x_k(a)) \frac{\partial x_i}{\partial t_j}(a)$ for all $j = 1, \dots, m$.

Writing explicit variables is to avoid confusion of $\partial_j z$ and $\partial_i z$ on both sides.

Theorem 18.3 (Euler's theorem).

Let $f: \mathbb{R}^k \rightarrow \mathbb{R}$ be a function which is **positively homogeneous of degree** $r \in \mathbb{R}$,

that is, $f(tx) = t^r f(x)$ for all $t > 0$ and $x \neq 0$, e.g., $f(x_1, x_2, x_3) = x_1^3 x_2^2 x_3^1 + x_1^6$.

If, in addition, f is differentiable at $a = (a_1, \dots, a_k) \in \mathbb{R}^k$ with $a \neq 0$,

then $rf(a) = a \cdot \nabla f(a) = \sum_{i=1}^k a_i \partial_i f(a)$.

Theorem 18.4. (gradient, normal, and orthogonality)

1. Let $U \subseteq \mathbb{R}^3$ be open and $f: U \rightarrow \mathbb{R}$ be differentiable on U such that
 $S = \{x \in U : f(x) = 0\}$ is a **two-dimensional smooth surface** in \mathbb{R}^3 , refer to Def. 29.2,
that is, locally S is an image of a C^1 one-to-one function (x_1, x_2, x_3) on an open subset of \mathbb{R}^2 .
If $a \in S$ and $\nabla f(a) \neq 0$, then the vector $\nabla f(a)$ is normal to the surface S at a ,
and the equation of the tangent plane to S at a is $\nabla f(a) \cdot (x - a) = 0$.
2. Let $V \subseteq \mathbb{R}^2$ be open, $g: V \rightarrow \mathbb{R}$ be differentiable on V ,
 $T = \{(x, z) \in \mathbb{R}^2 \times \mathbb{R} : x \in V \text{ and } z = g(x)\}$, and $a \in V$.
Then the tangent plane to T at $(a, g(a))$ is given by the equation $z = g(a) + \nabla g(a) \cdot (x - a)$,
that is, the plane in $\mathbb{R}^2 \times \mathbb{R}$ through the point $(a, g(a))$ with the normal vector $(-\nabla g(a), 1)$.
3. Let $V \subseteq \mathbb{R}^2$ be open, $g: V \rightarrow \mathbb{R}$ be differentiable on V , and $(a_1, a_2) \in V$.
if a differentiable function $t \mapsto (x_1(t), x_2(t)) \in V$ for $-1 < t < 1$
parameterizes **a part of the level curve** of g at (a_1, a_2) ,
that is, $\{(x_1(t), x_2(t)) \in V : -1 < t < 1\} \subseteq \{(x_1, x_2) \in V : g(x_1, x_2) = g(a_1, a_2)\}$ and $(x_1(0), x_2(0)) = (a_1, a_2)$,
then $\nabla g(a_1, a_2) \cdot (x'_1(0), x'_2(0)) = 0$,
that is, the gradient is orthogonal to the level curve at the given point.

Proposition 18.5.

Let $F, G: \mathbb{R}^m \rightarrow \mathbb{R}$ be differentiable.

Derive the formulas for the partial derivatives of $F + G$, FG , F/G , and F^G , by using Theorem 18.2. (exercise)

Introduction to Real Analysis

19 The mean value theorem and Rolle's theorem in several variables

Theorem 19.1 (the mean value theorem in several variables).

Let $a, b \in \mathbb{R}^k$ with $a \neq b$ and let $L = \{a + t(b - a) : 0 \leq t \leq 1\}$, which is the line segment from a to b .

Let $S \subseteq \mathbb{R}^k$ be an open set with $L \subseteq S$.

If $f : S \rightarrow \mathbb{R}$ is continuous on L and is differentiable on $L \setminus \{a, b\}$,

then there exists $c \in L$ such that $f(b) - f(a) = \nabla f(c) \cdot (b - a)$, where a, b are written as row vectors in \mathbb{R}^k .

Corollary 19.2.

Let $S \subseteq \mathbb{R}^k$ be open and **convex**, that is, $\{a + t(b - a) : 0 \leq t \leq 1\} \subseteq S$ for any $a, b \in S$.

Let $f : S \rightarrow \mathbb{R}$ be differentiable on S .

1. If there exists $M > 0$ such that $\|\nabla f(x)\| \leq M$ for every $x \in S$,
then $|f(b) - f(a)| \leq M\|b - a\|$ for all $a, b \in S$.
The convexity is needed here.
2. If $\nabla f(x) = 0$ for all $x \in S$, then f is a constant function on S .
This result is generalized in the following theorem.

Theorem 19.3.

Let $S \subseteq \mathbb{R}^k$ be open and connected.

If $f : S \rightarrow \mathbb{R}$ is differentiable on S and $\nabla f(x) = 0$ for all $x \in S$,

then f is a constant function on S .

The connectedness is needed here.

Theorem 19.4. (the critical point theorem in several variables)

Let $S \subseteq \mathbb{R}^k$ be open and $f : S \rightarrow \mathbb{R}$ be differentiable at $c \in S$.

If f has a local maximum or a local minimum at c ,

then c is a **critical point** of f , i.e., $\nabla f(c) = 0$.

Theorem 19.5 (Rolle's theorem in several variables).

1. Let $a, b \in \mathbb{R}^k$ with $a \neq b$, $L = \{a + t(b - a) : 0 \leq t \leq 1\}$, and $S \subseteq \mathbb{R}^k$ be an open set with $L \subseteq S$.
If $f : S \rightarrow \mathbb{R}$ is continuous on L and is differentiable on $L \setminus \{a, b\}$ with $f(a) = f(b)$,
then there exists $c \in L \setminus \{a, b\}$ such that $\partial_u f(c) = 0$, where $u = \frac{b-a}{\|b-a\|}$. (exercise)
2. Let $S \subseteq \mathbb{R}^k$ be nonempty, open and bounded.
Let $f : \mathbb{R}^k \rightarrow \mathbb{R}$ be continuous on \overline{S} , be differentiable on S , and be constant on ∂S .
Then there exists $c \in S$ such that $\nabla f(c) = 0$. (exercise)

Proposition 19.6.

Let $S \subseteq \mathbb{R}^k$ be open and connected. Let $f : S \rightarrow \mathbb{R}$ be differentiable on S . (exercise)

1. If $k = 1$ and $f'(x) = 0$ for all $x \in S$, then f is a constant function on S .
2. If $k \geq 2$, S is convex and $\partial_1 f(x) = 0$ for all $x \in S$,
then f is a **constant function in the first variable**, that is, if $(x_1, a), (x_2, a) \in S$ then $f(x_1, a) = f(x_2, a)$.
3. The convexity is needed in Item 2.

20 Implicit differentiation

Theorem 20.1. (*implicit differentiation*)

1. If a differentiable function $y = y(x_1, \dots, x_k)$ is defined implicitly by the equation $F(x_1, \dots, x_k, y) = 0$, where F is differentiable, then for every $1 \leq j \leq k$,

$$\frac{\partial y}{\partial x_j}(x_1, \dots, x_k) = -\frac{\frac{\partial F}{\partial x_j}(x_1, \dots, x_k, y)}{\frac{\partial F}{\partial y}(x_1, \dots, x_k, y)} \quad \text{provided} \quad \frac{\partial F}{\partial y}(x_1, \dots, x_k, y) \neq 0.$$

2. If two differentiable functions $u = u(x, y)$ and $v = v(x, y)$ are defined implicitly by the equations $F(x, y, u, v) = 0$ and $G(x, y, u, v) = 0$, where F and G are differentiable, then

$$\frac{\partial u}{\partial x} = -\frac{\det \begin{bmatrix} \frac{\partial F}{\partial x} & \frac{\partial F}{\partial v} \\ \frac{\partial G}{\partial x} & \frac{\partial G}{\partial v} \end{bmatrix}}{\det \begin{bmatrix} \frac{\partial F}{\partial u} & \frac{\partial F}{\partial v} \\ \frac{\partial G}{\partial u} & \frac{\partial G}{\partial v} \end{bmatrix}} \quad \text{and} \quad \frac{\partial v}{\partial x} = -\frac{\det \begin{bmatrix} \frac{\partial F}{\partial u} & \frac{\partial F}{\partial x} \\ \frac{\partial G}{\partial u} & \frac{\partial G}{\partial x} \end{bmatrix}}{\det \begin{bmatrix} \frac{\partial F}{\partial u} & \frac{\partial F}{\partial v} \\ \frac{\partial G}{\partial u} & \frac{\partial G}{\partial v} \end{bmatrix}},$$

provided the denominator is not zero.

3. If there are k differentiable functions $y_i = y_i(x_1, \dots, x_m)$, $1 \leq i \leq k$ defined implicitly by the equations $F_i(x_1, \dots, x_m, y_1, \dots, y_k) = 0$, where F_i is differentiable for all $1 \leq i \leq k$, then for $1 \leq i \leq k$ and $1 \leq j \leq m$, we have

$$\frac{\partial y_i}{\partial x_j} = -\frac{\det \begin{bmatrix} \frac{\partial F_1}{\partial y_1} & \dots & \frac{\partial F_1}{\partial x_j} & \dots & \frac{\partial F_1}{\partial y_k} \\ \vdots & & \vdots & & \vdots \\ \frac{\partial F_i}{\partial y_1} & \dots & \frac{\partial F_i}{\partial x_j} & \dots & \frac{\partial F_i}{\partial y_k} \\ \vdots & & \vdots & & \vdots \\ \frac{\partial F_k}{\partial y_1} & \dots & \frac{\partial F_k}{\partial x_j} & \dots & \frac{\partial F_k}{\partial y_k} \end{bmatrix}}{\det \begin{bmatrix} \frac{\partial F_1}{\partial y_1} & \dots & \dots & \dots & \frac{\partial F_1}{\partial y_k} \\ \vdots & & \vdots & & \vdots \\ \frac{\partial F_i}{\partial y_1} & \dots & \dots & \dots & \frac{\partial F_i}{\partial y_k} \\ \vdots & & \vdots & & \vdots \\ \frac{\partial F_k}{\partial y_1} & \dots & \dots & \dots & \frac{\partial F_k}{\partial y_k} \end{bmatrix}},$$

provided the denominator is not zero.

Remark 20.2.

1. When practicing implicit differentiations, we should go through the proof of the above theorems, instead of directly applying the formulas.
2. If y is implicitly defined by a function $F(x_1, \dots, x_k, y) = 0$ and the variables x_i 's are constrained to satisfy a equation $H(x_1, \dots, x_k) = 0$, then the differentiation of y with respect to x_1 , depends critically on which variable is fixed and the other $k-1$ variables, including x_1 , are considered as independent variables, e.g., $\frac{\partial y}{\partial x_1}|_{\{x_k \text{ is fixed}\}}$ and $y(x_1, x_2, \dots, x_{k-1})$ or $\frac{\partial y}{\partial x_1}|_{\{x_2 \text{ is fixed}\}}$ and $y(x_1, x_3, \dots, x_k)$.