Real Analysis

Youming Zhao
Email: youming0.zhao@gmail.com

First draft: January 1, 2022 Last update: December 28, 2023

Contents

1 Sequences

A sequence of real numbers, denoted $(a_n)_{n\in\mathbb{N}}$, is a map from natural numbers $\mathbb{N}$ to real numbers. Note that the starting index can be any nonnegative integers, so a more general notation for a sequence is $(a_n)_{n=m}^{\infty}$.

1.1 Cauchy sequence

Definition 1 (**Cauchy sequence of reals**)**.** A sequence $(a_n)_{n=m}^{\infty}$ of real numbers is a Cauchy sequence if, for every real $\epsilon > 0$, there exists an $N \geq m$ such that $|a_n - a_{n'}| \leq \epsilon$ for all $n, n' \geq N$.

Definition 2 (bounded sequences)**.** A sequence $(a_n)_{n=m}^{\infty}$ of real numbers is bounded by a real number M iff we have $|a_n| \leq M$ for all $n \geq m$.

Proposition 1 (**Cauchy sequences are bounded**)**.** *If a sequence (s_n) is Cauchy, then (s_n) is bounded.*

Proof. Since (s_n) is Cauchy, there is an N such that for a given $\epsilon > 0$, $|s_n - s_m| \leq \epsilon$ for all $m, n > N$. Hence, $|s_n| \leq \epsilon + |s_m|$. Let $M := \max\{|s_1|, |s_2|, \cdots, |s_N|, \epsilon + |s_m|\}$, then we have two cases.

- Case 1: when $n \leq N$, we have $|s_n| \leq \max\{|s_1|, |s_1|, \cdots, |s_N|\} \leq M$;
- Case 2: when $n \geq N+1$, we have $|s_n| \leq \epsilon + |s_m| \leq M$.

Thus, (s_n) is bounded. □

1.2 Convergence and limit

Definition 3 (**Convergence of sequences**)**.** A sequence $(a_n)_{n=m}^{\infty}$ of real numbers is convergent if and only if, for a real number L and every real $\epsilon > 0$, there exists an $N \geq m$ such that $|a_n - L| \leq \epsilon$ for all $n \geq N$.

See Exercise 6.1.2 in Tao's Analysis I.

Proposition 2 (**Uniqueness of limits**)**.** *Let $(a_n)_{n=m}^{\infty}$ be a real sequence starting at some integer index m, and let $L \neq L'$ be two distinct real numbers. Then it is not possible for $(a_n)_{n=m}^{\infty}$ to converge to L while also converging to L'.*

Proof. Suppose for sake of contradiction that $(a_n)_{n=m}^{\infty}$ was converging to both L and L'. Let $\epsilon = |L - L'|/3$; note that ϵ is positive since $L \neq L'$. Since $(a_n)_{n=m}^{\infty}$ converges to L, there exists an $N \geq m$ such that $|a_n - L| \leq \epsilon$ for all $n \geq N$. Similarly, there is an $M \geq m$ such that $|a_n - L'| \leq \epsilon$ for all $n \geq M$. If we set $n := \max\{N, M\}$, then we have $|a_n - L| \leq \epsilon$ and $|a_n - L'| \leq \epsilon$. Hence, by the triangle inequality, $|L - L'| \leq |a_n - L| + |a_n - L'| \leq 2\epsilon = 2|L - L'|/3$, which contradicts the fact that $|L - L'| > 0$. Thus it is not possible to converge to both L and L'. □

Definition 4 (**Limits of sequences**)**.** If a sequence $(a_n)_{n=m}^{\infty}$ converges to some real number L, we say that $(a_n)_{n=m}^{\infty}$ is **convergent** and that its **limit** is L; we write

$$L = \lim_{n\to\infty} a_n$$

to denote this fact. If a sequence $(a_n)_{n=m}^{\infty}$ is not converging to any real number L, we say that the sequence $(a_n)_{n=m}^{\infty}$ is **divergent** and we leave $\lim_{n\to\infty} a_n$ undefined.

Remark 1. Note that, convergence means that all the terms are eventually close to **a fixed number**, whereas Cauchy means that all the terms are eventually close to **each other**.

Definition 5 (Subsequences)**.** Let $(n_k)_{k\in\mathbb{N}}$ be a sequence of natural numbers that is strictly increasing, then $(a_{n_k})_{k\in\mathbb{N}}$ is called a subsequence of $(a_n)_{n\in\mathbb{N}}$.

Definition 6 (Limit points, accumulation points)**.** x is a limit point (an accumulation point) of $(a_n)_{n=m}^{\infty}$ if, for every ϵ and every $N \geq m$, there exists an $n \geq N$ such that $|a_n - x| \leq \epsilon$.

Proposition 3. *$a \in \mathbf{R}$ is an accumulation point of $(a_n)_{n\in\mathbb{N}}$ if and only if for all $\epsilon > 0$, the ϵ-neighborhood of a contains infinitely many sequence members of $(a_n)_{n\in\mathbb{N}}$.*

Proposition 4 (**Convergent sequences are Cauchy**)**.** *If (s_n) converges to s, (s_n) is Cauchy.*

Proof. For a given $\epsilon > 0$, there exists an N such that $|s_n - s| \leq \epsilon/2$ for all $n \geq N$. Also, we have $|s_m - s| \leq \epsilon/2$ for all $m \geq N$. By the triangle inequality, $|s_n - s_m| = |s_n - s + s - s_m| \leq |s_n - s| + |s - s_m| \leq \epsilon$. Thus, (s_n) is Cauchy. □

Corollary 1 (**Convergent sequences are bounded.**)**.** *Every convergent sequence of real numbers is bounded.*

Proof. By Proposition 4, convergent sequences are Cauchy. By Proposition 1, Cauchy sequences are bounded. Thus, convergent sequences are bounded. □

Theorem 1 (Completeness of the reals)**.** *A sequence $(a_n)_{n=1}^{\infty}$ of real numbers is a Cauchy sequence if and only if it is convergent.*

The following proof is largely taken from Terence Tao's Analysis I.

Proof. Proposition 4 has shown that every convergent sequence is Cauchy, so it suffices to prove that every Cauchy sequence is convergent.

Let $(a_n)_{n=1}^{\infty}$ be a Cauchy sequence. We know from Proposition 1 that the sequence $(a_n)_{n=1}^{\infty}$ is bounded, which implies that $L^- := \liminf_{n\to\infty} a_n$ and $L^+ := \limsup_{n\to\infty} a_n$ of the sequence are both finite. To show that the sequence converges, it will suffice to show that $L^- = L^+$.

Now let $\epsilon > 0$ be any real number. Since $(a_n)_{n=1}^{\infty}$ is a Cauchy sequence, there exists an $N \geq 1$ such that $a_N - \epsilon \leq a_n \leq a_N + \epsilon$ for all $n \geq N$. This implies that

$$a_N - \epsilon \leq \inf(a_n)_{n=N}^{\infty} \leq \sup(a_n)_{n=N}^{\infty} \leq a_N + \epsilon \tag{1}$$

and hence by the definition of L^- and L^+

$$a_N - \epsilon \leq L^- \leq L^+ \leq a_N + \epsilon. \tag{2}$$

Thus we get

$$0 \leq L^+ - L^- \leq 2\epsilon \tag{3}$$

which is true for all $\epsilon > 0$. Since L^- and L^+ do not depend on ϵ, then we must have $L^+ = L^-$. Thus, $(a_n)_{n=1}^{\infty}$ is convergent. □

Remark 2. Theorem 1 tells us Cauchy sequences and convergent sequences are equivalent. More straightforwardly,

$$\textbf{Cauchy sequences} \iff \textbf{convergent sequences}. \tag{4}$$

Theorem 2 (Limit Laws)**.** *Let $(a_n)_{n=m}^{\infty}$ and $(b_n)_{n=m}^{\infty}$ be convergent sequences of real numbers, and let x, y be the real numbers $x := \lim_{n\to\infty} a_n$ and $y := \lim_{n\to\infty} b_n$.*

(a) The sequence $(a_n + b_n)_{n=m}^{\infty}$ converges to $x + y$; in other words,

$$\lim_{n\to\infty}(a_n + b_n) = \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n. \tag{5}$$

(b) The sequence $(a_n b_n)_{n=m}^{\infty}$ converges to xy; in other words,

$$\lim_{n\to\infty}(a_n b_n) = (\lim_{n\to\infty} a_n)(\lim_{n\to\infty} b_n). \tag{6}$$

(c) For any real number c, the sequence $(ca_n)_{n=m}^{\infty}$ converges to cx; in other words,

$$\lim_{n\to\infty}(ca_n) = c\lim_{n\to\infty} a_n. \tag{7}$$

(d) The sequence $(a_n - b_n)_{n=m}^{\infty}$ converges to $x - y$; in other words,

$$\lim_{n\to\infty}(a_n - b_n) = \lim_{n\to\infty} a_n - \lim_{n\to\infty} b_n. \tag{8}$$

(e) Suppose that $y \neq 0$, and that $b_n \neq 0$ for all $n \geq m$. Then the sequence $(b_n^{-1})_{n=m}^{\infty}$ converges to y^{-1}; in other words,

$$\lim_{n\to\infty} b_n^{-1} = (\lim_{n\to\infty} b_n)^{-1}. \tag{9}$$

(f) Suppose that $y \neq 0$, and that $b_n \neq 0$ for all $n \geq m$. Then the sequence $(a_n/b_n)_{n=m}^{\infty}$ converges to x/y; in other words,

$$\lim_{n\to\infty}\frac{a_n}{b_n} = \frac{\lim_{n\to\infty} a_n}{\lim_{n\to\infty} b_n}. \tag{10}$$

(g) The sequence $(\max(a_n, b_n))_{n=m}^{\infty}$ converges to $\max(x, y)$; in other words,

$$\lim_{n\to\infty}\max(a_n, b_n) = \max(\lim_{n\to\infty} a_n, \lim_{n\to\infty} b_n). \tag{11}$$

(h) The sequence $(\min(a_n, b_n))_{n=m}^{\infty}$ converges to $\min(x, y)$; in other words,

$$\lim_{n\to\infty}\min(a_n, b_n) = \min(\lim_{n\to\infty} a_n, \lim_{n\to\infty} b_n). \tag{12}$$

Proof. (a) Since $(a_n)_{n=m}^{\infty}$ converges to x, then for all $\epsilon > 0$, there exists a positive integer $N_1 > m$ such that for any $n > N_1$, $|a_n - x| \leq \frac{\epsilon}{2}$. Similarly, for all $\epsilon > 0$, there exists a positive integer $N_2 > m$ such that for any $n > N_2$, $|b_n - y| \leq \frac{\epsilon}{2}$. Then for any $n > \max(N_1, N_2)$, we have

$$|a_n + b_n - x - y| = |a_n - x + b_n - y| \leq |a_n - x| + |b_n - y| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \tag{13}$$

which implies $(a_n + b_n)_{n=m}^{\infty}$ converges to $x + y$, as desired.

(b) Since $(a_n)_{n=m}^{\infty}$ converges to x, then for all $\epsilon > 0$, there exists a positive integer $N_1 > m$ such that for any $n > N_1$, $|a_n - x| \leq \epsilon$. Also, by Corollary 1, $|a_n| \leq M \in \mathbf{R}$ for all $n \geq m$. Similarly, for all $\epsilon > 0$, there exists a positive integer $N_2 > m$ such that for any $n > N_2$, $|b_n - y| \leq \epsilon$. Then for any $n > \max(N_1, N_2)$, we have

$$\begin{aligned} |a_n b_n - xy| &= |a_n b_n - ya_n + ya_n - xy| && (14)\\ &= |a_n(b_n - y) + y(a_n - x)| && (15)\\ &\leq |a_n(b_n - y)| + |y(a_n - x)| && (16)\\ &\leq (M + |y|)\epsilon && (17) \end{aligned}$$

which implies $(a_n b_n)_{n=m}^{\infty}$ converges to xy since M and y are constants.

(c) This is the special case when $b_n = c$ for all $n \geq m$.

(d) It follows from (a) and c due to the fact $a_n - b_n = a_n + (-1) \cdot b_n$.

$\square$

Theorem 3 (squeeze test, sandwich theorem)**.** *Let $(a_n)_{n=m}^{\infty}$, $(b_n)_{n=m}^{\infty}$, and $(c_n)_{n=m}^{\infty}$ be sequences of real numbers such that $a_n \leq b_n \leq c_n$ for all $n \geq m$. Suppose also that $(a_n)_{n=m}^{\infty}$ and $(a_n)_{n=m}^{\infty}$ both converge to the same limit L. Then $(b_n)_{n=m}^{\infty}$ is also convergent to L.*

Proof. Let $(d_n)_{n=m}^{\infty}$ be a sequence with $d_n = b_n - a_n$, then $0 \leq d_n \leq c_n - a_n$. Since both $(a_n)_{n=m}^{\infty}$ and $(c_n)_{n=m}^{\infty}$ converge to L, according to the limit laws, $\lim_{n\to\infty}(c_n - a_n) = L - L = 0$. Therefore, for all $\epsilon > 0$, there exists a positive integer $N > m$ such that for all $n > N$, $|d_n| = |d_n - 0| \leq |c_n| \leq \epsilon$. This indicates that $(d_n)_{n=m}^{\infty}$ is convergent with $\lim_{n\to\infty} d_n = 0$. Furthermore,

$$\lim_{n\to\infty}(d_n + a_n) = \lim_{n\to\infty} d_n + \lim_{n\to\infty} a_n = 0 + L = L. \tag{18}$$

Since $b_n = d_n + a_n$, then $\lim_{n\to\infty} b_n = L$ as desired. $\square$

1.3 Upper bound and supremum

Definition 7 (**Upper bound**)**.** Let E be a subset of $\mathbf{R}$, and let M be a real number. We say that M is an **upper bound** for E, iff we have $x \leq M$ for every element x in E.

Definition 8 (**Least upper bound**)**.** Let E be a subset of $\mathbf{R}$, and let M be a real number. We say that M is a **least upper bound** for E, iff (a) M is an upper bound for E, and also (b) any other upper bound M' for E must be larger than or equal to M.

Proposition 5 (**Uniqueness of least upper bound**)**.** *Let E be a subset of $\mathbf{R}$. Then E can have at most one least upper bound.*

Proof. Let M_1 and M_2 be two least upper bounds. Since M_1 is a least upper bound and M_2 is an upper bound, then by definition of least upper bound we have $M_2 \geq M_1$. Since M_2 is a least upper bound and M_1 is an upper bound, we similarly have $M_1 \geq M_2$. Thus $M_1 = M_2$. Thus there is at most one least upper bound. $\square$

Now we come to an important property of the real numbers:

Theorem 4 (**Existence of least upper bound**)**.** *Let E be a nonempty subset of $\mathbf{R}$. If E has an upper bound, (i.e., E has some upper bound M), then it must have exactly one least upper bound.*

Proof. This theorem will take quite a bit of effort to prove, see Page 118 of Terence Tao's "Analysis 1, 3rd Edition". $\square$

Definition 9 (**Supremum**)**.** Let E be a subset of the real numbers. If E is non-empty and has some upper bound, **we define** $\sup(E)$ **to be the least upper bound of** E. If E is non-empty and has no upper bound, we set $\sup(E) := +\infty$; if E is empty, we set $\sup(E) := -\infty$. We refer to $\sup(E)$ as the supremum of E, and also denote it by $\sup E$.

Similarly, we can define greatest lower bound and infimum.

Remark 3. We can think of Theorem 4 as saying "$\sup(E)$ always exists". Because either E is bounded above (in which case $\sup(E)$ exists), or E is unbounded(in which case $\sup(E) = \infty$). This is a fundamental theorem of analysis. Also, by Proposition 5, $\sup(E)$ or $\inf(E)$ is unique.

Important fact: $\inf(S) = -\sup(-S)$.

Theorem 5. *Let E be a subset of $\mathbf{R}^*$. Then the following statements are true.*

1. *For every $x \in E$ we have $x \leq \sup(E)$ and $x \in \inf(E)$.*
2. *Suppose that $M \in \mathbf{R}^*$ is an upper bound for E, i.e., $x \leq M$ for all $x \in E$. Then we have $\sup(E) \leq M$.*
3. *Suppose that $M \in \mathbf{R}^*$ is a lower bound for E, i.e., $x \geq M$ for all $x \in E$. Then we have $\inf(E) \geq M$.*

Proof. □

Proposition 6 (least upper bound property)**.** *Let $(a_n)_{n=m}^{\infty}$ be a sequence of real numbers, and let x be the extended real number $x := \sup(a_n)_{n=m}^{\infty}$. Then we have $a_n \leq x$ for all $n \geq m$. Also, whenever $M \in \mathbf{R}^*$ is an upper bound for a_n (i.e., $a_n \leq M$ for all $n \geq m$), we have $x \leq M$. Finally, for every extended real number y for which $y < x$, there exists at least one $n \geq m$ for which $y < a_n \leq x$.*

Proof. □

Proposition 7 (Monotone bounded sequences converge)**.** *Let $(a_n)_{n=m}^{\infty}$ be a sequence of real numbers which has some finite upper bound $M \in \mathbf{R}$, and which is also increasing (i.e., $a_{n+1} \geq a_n$ for all $n \leq m$). Then $(a_n)_{n=m}^{\infty}$ is convergent, and in fact*

$$\lim_{n\to\infty} a_n = \sup(a_n)_{n=m}^{\infty} \leq M.$$

Proof. Since $(a_n)_{n=m}^{\infty}$ is upper bounded by M, then by Theorem 5, $l = \sup(a_n)_{n=m}^{\infty} \leq M$. By the definition of supremum, for every $\epsilon > 0$, there exists N such that $l - \epsilon < x_N \leq x_n \leq l$ for all $n \geq N$. Therefore, we have $|l - x_n| < \epsilon$. Since ϵ is arbitrary, then $\lim_{n\to\infty} a_n = l$. □

Example: The sequence $(a_n)_{n=1}^{\infty}$ given by $a_n = (1 + \frac{1}{n})^n$ is convergent.

Proof. First, we show that $(a_n)_{n=1}^{\infty}$ is an increasing sequence. In order to do this, we employ GM-Am inequality to get

$$a_n = (1 + \frac{1}{n})^n \cdot 1 \leq \left(\frac{n(1+\frac{1}{n})+1}{n+1}\right)^{n+1} = \left(\frac{n+1+1}{n+1}\right)^{n+1} = a_{n+1} \tag{19}$$

which indicates $(a_n)_{n=1}^{\infty}$ is an increasing sequence. Now we show that $(a_n)_{n=1}^{\infty}$ is bounded from above as follows.

$$\begin{aligned}
a_n =& (1 + \frac{1}{n})^n && (20)\\
=& \sum_{k=0}^{n} \binom{n}{k} \cdot 1^{n-k} \cdot (\frac{1}{n})^k && (21)\\
=& \binom{n}{0} \cdot 1^n \cdot (\frac{1}{n})^0 + \binom{n}{1} \cdot 1^{n-1} \cdot (\frac{1}{n})^1 + \sum_{k=2}^{n} \binom{n}{k} \cdot 1^{n-k} \cdot (\frac{1}{n})^k && (22)\\
=& 1 + 1 + \sum_{k=2}^{n} \binom{n}{k} \cdot \frac{1}{n^k} && (23)\\
=& 2 + \sum_{k=2}^{n} \frac{n(n-1)(n-2)\cdots(n-k+1)}{n \cdot n \cdot n \cdot \cdots \cdot n} \cdot \frac{1}{k!} && (24)\\
\leq& 2 + \sum_{k=2}^{n} \frac{1}{k!} \leq 2 + \sum_{k=2}^{n} \frac{1}{k(k-1)} && (25)\\
=& 2 + 1 - \frac{1}{n} < 3. && (26)
\end{aligned}$$

Hence, $(a_n)_{n=1}^{\infty}$ is increasing and bounded above. By Proposition 7, it converges. $\square$

$\square$

Remark 4. By convention we use e to denote the limit of the above sequence, namely

$$\lim_{n\to\infty}(1+\frac{1}{n})^n = e = 2.718281828459\cdots \tag{27}$$

which is an irrational number.

1.4 Bolzano-Weirstrass Theorem

Theorem 6 (Bolzano-Weierstrass theorem: every bounded sequence has a convergent subsequence)**.** *Let $(a_n)_{n=0}^{\infty}$ be a bounded sequence (i.e., there exists a real number $M > 0$ such that $|a_n| \leq M$ for all $n \in \mathbb{N}$). Then there is at least one subsequence of $(a_n)_{n=0}^{\infty}$ which converges.*

Proof. Since $(a_n)_{n=0}^{\infty}$ is a bounded sequence, we can find an interval $[c_0, d_0] \subset [-M, M]$ such that every member of $(a_n)_{n=0}^{\infty}$ resides in $[c_0, d_0]$. Now we bisect $[c_0, d_0]$ at $(c_0 + d_0)/2$. If the left half contains infinitely many members of $(a_n)_{n=0}^{\infty}$, let $c_1 := c_0$ and $d_1 := (c_0 + d_0)/2$, otherwise let $c_1 := (c_0 + d_0)/2$ and $d_1 := d_0$. In this way, we can construct the following nested intervals.

$$[c_0, d_0] \supset [c_1, d_1] \supset \cdots \supset [c_n, d_n] \supset \cdots \tag{28}$$

which yields two monotone bounded sequences $(c_n)_{n=0}^{\infty}$ and $(d_n)_{n=0}^{\infty}$. Specifically, $(c_n)_{n=0}^{\infty}$ is increasing and bounded above, and $(d_n)_{n=0}^{\infty}$ is decreasing and bounded below. Therefore, by Proposition 7, $(c_n)_{n=0}^{\infty}$ and $(d_n)_{n=0}^{\infty}$ are both convergent. Since $d_n - c_n = \frac{1}{2^n}(d_0 - c_0)$,

$$\lim_{n\to\infty}(d_n - c_n) = \lim_{n\to\infty}\frac{1}{2^n}(d_0 - c_0) = 0, \tag{29}$$

which implies $\lim_{n\to\infty} d_n = \lim_{n\to\infty} c_n = 0$. For each interval $[c_n, d_n]$, we choose a point b_n which is a member of $(a_n)_{n=0}^{\infty}$. Thus, we get a subsequence $(b_n)_{n=0}^{\infty}$ where $c_n \leq b_n \leq d_n$. By the sandwich theorem, $(b_n)_{n=0}^{\infty}$ is a convergent subsequence which converges to $\lim_{n\to\infty} d_n$. $\square$

1.5 The upper limit and lower limit

Let E denote the set of limit points,

$$E = \{\xi | \xi \text{ is a limit point of } \{x_n\}\}.$$

$H = \max E$ is called the upper limit of the sequence $\{x_n\}$, denoted as

$$H = \overline{\lim}_{n\to\infty} x_n;$$

$H = \min E$ is called the lower limit of the sequence $\{x_n\}$, denoted as

$$H = \underline{\lim}_{n\to\infty} x_n;$$

Proposition 8 (Limits are limit points)**.** *Let $(a_n)_{n=m}^{\infty}$ be a sequence which converges to a real number c. Then c is a limit point of $(a_n)_{n=m}^{\infty}$, and in fact it is the only limit point of $(a_n)_{n=m}^{\infty}$.*

Proof. See Exercise 6.4.1 in Tao's Analysis I. $\square$

1.6 Limit superior and limit inferior

This subsection is largely taken from Definition 6.4.6 and Example 6.4.7 of Tao's Analysis I. Now we look at two special types of limit points: the limit superior (lim sup) and limit inferior (lim inf).

Definition 10 (Limit superior and limit inferior)**.** Suppose that $(a_n)_{n=m}^{\infty}$ is a sequence. We define a new sequence $(a_N^+)_{N=m}^{\infty}$ by the formula

$$a_N^+ := \sup(a_n)_{n=N}^{\infty}.$$

More informally, a_N^+ is the supremum of all the elements in the sequence from a_N onwards. We then define the limit superior of the sequence $(a_n)_{n=m}^{\infty}$, denote $\limsup_{n\to\infty} a_n$, by the formula

$$\lim \sup_{n\to\infty} a_n := \inf(a_N^+)_{N=m}^{\infty}.$$

Notice the sequence $(a_N^+)_{N=m}^{\infty}$ is decreasing.

Similarly, we can define

$$a_N^- := \inf(a_n)_{n=N}^{\infty}.$$

and define the limit inferior of the sequence $(a_n)_{n=m}^{\infty}$, denote $\liminf_{n\to\infty} a_n$, by the formula

$$\lim \inf_{n\to\infty} a_n := \sup(a_N^-)_{N=m}^{\infty}.$$

Notice the sequence $(a_N^-)_{N=m}^{\infty}$ is increasing.

Remark 5. To make the above definition more straightforward, we express a_N^+ in an equivalent way, namely,

$$a_N^+ = \sup\{a_N, a_{N+1}, \cdots\}.$$

Likewise,

$$a_N^- = \inf\{a_N, a_{N+1}, \cdots\}.$$

Example: Let $a_1, a_2, a_3, \cdots$ denote the sequence

$$1.1, -1.01, 1.001, -1.0001, 1.00001, \ldots.$$

Then $a_1^+, a_2^+, a_3+, \cdots$ is the sequence

$$1.1, 1.001, 1.001, 1.00001, 1.00001, \ldots$$

and its infimum is 1. Hence the limit superior of this sequence is 1. Similarly, $a_1^-, a_2^-, a_3-, \cdots$ is the sequence

$$-1.01, -1.01, -1.0001, -1.0001, -1.000001, \ldots$$

and its supremum is -1. Hence the limit inferior of this sequence is 1. One should compare this with the supremum and infimum of the sequence, which are 1.1 and -1.01 respectively. □

Here we provide another definition for the limit superior and limit inferior.

Definition 11 (limit superior and limit inferior)**.** Let $(a_n)_{n\in\mathbb{N}}$ be a sequence of real numbers. $a \in \mathbf{R} \cup \{-\infty, \infty\}$ is called the limit superior of $(a_n)_{n\in\mathbb{N}}$ if a is the largest accumulation point of $(a_n)_{n\in\mathbb{N}}$, denoted $a := \limsup_{n\to\infty} a_n$. Likewise, $b \in \mathbf{R} \cup \{-\infty, \infty\}$ is called the limit inferior of $(a_n)_{n\in\mathbb{N}}$ if b is the smallest accumulation point of $(a_n)_{n\in\mathbb{N}}$, denoted $a := \liminf_{n\to\infty} a_n$.

From this definition, we have the following facts.

$$\limsup_{n\to\infty} a_n = \lim_{n\to\infty} \sup\{a_k \mid k \geq n\} \tag{30}$$

$$\liminf_{n\to\infty} a_n = \lim_{n\to\infty} \inf\{a_k \mid k \geq n\} \tag{31}$$

1.7 The relation between upper (lower) limit and limit superior (inferior)

Mathematically speaking, they are the same thing. Now we put it in a theorem and prove its correctness.

Theorem 7. *Given a bounded sequence $\{x_n\}$, we have $\overline{X} = \overline{\lim}_{n\to\infty} x_n$, and $\underline{X} = \underline{\lim}_{n\to\infty} x_n$ iff the following holds*

$$\overline{X} = \lim\sup_{n\to\infty} x_n; \quad \underline{X} = \lim\inf_{n\to\infty} x_n.$$

Proof. Let $b_n = \sup_{k\geq n}\{x_k\}$. From $\sup_{k\geq n}\{x_k\} \geq \sup_{k\geq n+1}\{x_k\}$, we have $b_n \geq b_{n+1}$. So $\{b_n\}$ is decreasing and bounded. Then $\lim_{n\to\infty} b_n$ exists. Since $\overline{X} = \overline{\lim}_{n\to\infty} x_n$, there exists a subsequence $\lim_{n_k\to\infty} x_{n_k} = \overline{X}$. Since $b_n \geq x_n$, there also exists a subsequence b_{n_k} such that $b_{n_k} \geq x_{n_k}$. Thus, $\lim_{n_k\to\infty} b_{n_k} \geq \lim_{n_k\to\infty} x_{n_k} = \overline{X}$. By the uniqueness of limits, $\lim_{n\to\infty} b_n = \lim_{n_k\to\infty} b_{n_k} \geq \overline{X}$.

Since $b_n = \sup_{k\geq n}\{x_k\}$, for every $\epsilon > 0$, there exists $x_{k_n} \in \{x_k\}$ such that $x_{k_n} \geq b_n - \epsilon$. Then $\lim_{n\to\infty} x_{k_n} > \lim_{n\to\infty} b_n - \epsilon$. Since $\{x_{k_n}\}$ is a subsequence of $\{x_k\}$, we always have $\overline{X} \geq \lim_{n\to\infty} x_{k_n}$. Thus, $\overline{X} > \lim_{n\to\infty} b_n - \epsilon$. Since ϵ is arbitrary, $\overline{X} \geq \lim_{n\to\infty} b_n$. Hence, $\lim_{n\to\infty} b_n = \overline{X}$.

Now we show the necessity. Since $\overline{X} = \overline{\lim}_{n\to\infty} x_n$, for $\forall\epsilon > 0$, there exists N such that $x_n < H+\epsilon, \forall n > N$. Thus, $\lim\sup_{n\to\infty} x_n \leq H+\epsilon$. Then $b_n \leq H+\epsilon$. Since ϵ is arbitrary, $\lim_{n\to\infty} b_n \leq H$. Since $\overline{X} = \overline{\lim}_{n\to\infty} x_n$, there exist infinitely many $x_n > H-\epsilon, \forall\epsilon > 0$. $b_n = \sup_{k\geq n}\{x_k\}$, we always have $b_n \geq x_n > H - \epsilon$. Then $\lim_{n\to\infty} b_n \geq \lim_{n\to\infty} x_n > H - \epsilon$. Since ϵ is arbitrary, $\lim_{n\to\infty} b_n \geq H$. Thus, $H = \lim_{n\to\infty} b_n = \lim\sup_{n\to\infty} x_n$.

We can show the part for $\underline{X}$ in a similar way. This completes our proof. □

Remark 6. This result can be generalized to unbounded sequences. If $\{x_n\}$ is unbounded from above, then $b_n = \sup_{k\geq n}\{x_k\} = +\infty$. Then there exists a unbounded subsequence $\{x_{n_k}\}$ such that $\lim_{k\to\infty} x_{n_k} = +\infty$. Thus, $\overline{X} = +\infty$. If $\lim_{n\to\infty} b_n = \lim\sup_{n\to\infty} x_n = -\infty$, it is obvious that $\overline{X} = -\infty$. On the other hand, if $\{x_n\}$ is unbounded from below, then $a_n = \inf_{k\geq n}\{x_k\} = -\infty$ and $\underline{X} = -\infty$. If $a_n = \inf_{k\geq n}\{x_k\} = +\infty$, then $a_n \leq x_n$ and $\underline{X} = +\infty$.

1.8 Theorem of a nested sequence of closed intervals

Definition 12 (Nested sequence of closed intervals). If a sequence of closed intervals, $\{[a_n, b_n]\}$, satisfies

1. $[a_{n+1}, b_{n+1}] \subset [a_n, b_n]$;
2. $\lim_{n\to\infty}(b_n - a_n) = 0$.

then $\{[a_n, b_n]\}$ is called a nested sequence of closed intervals.

Theorem 8 (Theorem of a nested sequence of closed intervals). *If $\{[a_n, b_n]\}$ is a nested sequence of closed intervals, then there exists a unique x that lies in all $[a_n, b_n]$ and $x = \lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n$. In other words, no other y can lie in all $[a_n, b_n]$.*

Proof. From the first condition of Definition 12, $\{[a_n, b_n]\}$ is a closed nested sequence of intervals, we have

$$a_1 \leq \cdots a_n \leq a_{n+1} \leq \cdots \leq b_{n+1} \leq b_n \cdots \leq b_1$$

Thus, a_n is monotone increasing and upper bounded by b_1, and b_n is monotone decreasing and lower bounded by a_1. By Proposition 7, both $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ converge. Suppose $\lim_{n\to\infty} a_n = x$,

$$\lim_{n\to\infty} b_n = \lim_{n\to\infty}(b_n - a_n + a_n) = 0 + x = x.$$

this means x is the supremum of $(a_n)_{n=1}^{\infty}$ and the infimum of $(b_n)_{n=1}^{\infty}$. Thus,

$$a_n \leq x \leq b_n, \; n = 1, 2, 3, \ldots$$

Suppose there is another x' that satisfies

$$a_n \le x' \le b_n,\ n = 1, 2, 3, \ldots$$

By the squeeze test,

$$x' = \lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = x$$

Thus, x is unique that is in all $[a_n, b_n]$. □

The following remark is largely from https://www.zhihu.com/question/40106847.

Remark 7. Does this theorem hold for open intervals? No. A counter example is $\{(0, \frac{1}{n})\}$. Since $\lim_{n\to\infty} \frac{1}{n} = 0$, then x(if it exists) is supposed to be very small and near 0. However, we can always find a sufficiently large integer $N = [1/x] + 1 > 1/x$ where $[1/x]$ denotes the integer part of $1/x$. For any $n > N$, we have $x \ge 1/N > 1/n > 0$, i.e., $x \notin (0, 1/n)$. Hence, we can not find an x in all (a_n, b_n). Following this procedure, the only possible real number which is in all (a_n, b_n) is 0, but $0 \notin (0, \frac{1}{n})$. Now we give the conditions under which this theorem holds for a nested sequence of open intervals: a_n is **strictly** monotone increasing and b_n is **strictly** monotone decreasing, i.e., $a_1 < \cdots a_n < a_{n+1} < \cdots < b_{n+1} < b_n \cdots < b_1$. The proof is similar with the one for closed intervals.

2 Functions

2.1 An important lemma

A vector whose dot product with every unit vector is no larger than λ has norm $\le \lambda$. We introduce a lemma to formulate it as follows.

Lemma 1. *Given a vector x, if $x^T u \le \lambda$ holds for arbitrary u with $\|u\| = 1$, then $\|x\| \le \lambda$.*

Proof. Suppose $\|x\| > \lambda$, we let $u = \frac{x}{\|x\|}$,then we get

$$x^T u = \frac{x^T x}{\|x\|} = \frac{\|x\|^2}{\|x\|} = \|x\| > \lambda$$

which contradicts the given condition. Hence the assumption does not hold. This completes the proof. □

2.2 Smooth function

Definition 13 (Smooth function)**.** We say that a function f: dom$f \to \mathbf{R}$ is β-smooth over a convex set $\mathcal{K} \subset$ dom$f$ with respect to $\|\cdot\|$ if for all $x, y \in \mathcal{K}$ it holds that[3]

$$f(y) \le f(x) + \langle y - x, \nabla f(x)\rangle + \frac{\beta}{2}\|y - x\|^2.$$

After some algebraic manipulations, we can get the equivalent form

$$\frac{\beta}{2} y^T y - f(y) \ge \frac{\beta}{2} x^T x - f(x) + \langle \beta x - \nabla f(x), y - x\rangle$$

Hence, f is β-smooth over a convex set $\mathcal{K} \subset$ dom$f$ iff $\frac{\beta}{2} x^T x - f(x)$ is convex on the same set.

2.3 Smoothness classes

A map $f : U \to \mathbf{R}^m$ is of class C^r if it is r^{th}-order differentiable at each $p \in U$ and its derivatives continuously on p. Since differentiability implies continuity, all the derivatives of order less than r are automatically continuous. Only the r^{th} derivative is in question. **If f is of class C^r for all r then it is smooth or of class C^∞.**

2.4 Local Lipschitz property

This proposition is taken from section 6.3 of Marsden & Hoffman's Elementary Classical Analysis, second edition. The proof is not provided there, so I write a proof here.

Proposition 9. *Suppose $A \in \mathbf{R}^n$ is open and $f : A \in \mathbf{R}^m$ is differentiable on A. Then f is continuous. In fact, for each $x_0 \in A$, there are a constant $M > 0$ and a $\delta_0 > 0$ such that $\|x - x_0\| < \delta_0$ implies $\|f(x) - f(x_0)\| \leq M\|x - x_0\|$. (This is called the local Lipschitz property.)*

Proof. Since $f : A \in \mathbf{R}^m$ is differentiable on the open set A, all of its partial derivatives $D_k f_j$ at any point in A exist and are bounded. According to the Mean-Value theorem for vector-valued functions (see Theorem 12.9 in [1]), we have

$$\mathbf{a} \cdot \{f(\mathbf{y}) - f(\mathbf{x})\} = \mathbf{a} \cdot \{f'(\mathbf{z})(\mathbf{y} - \mathbf{x})\}. \tag{32}$$

where $\mathbf{a}$ is an arbitrary unit vector and $\mathbf{z}$ is a convex combination of $\mathbf{x}, \mathbf{y}$. Let $\mathbf{a} = \frac{f(\mathbf{y}) - f(\mathbf{x})}{\|f(\mathbf{y}) - f(\mathbf{x})\|}$, then

$$\begin{aligned}
\frac{f(\mathbf{y}) - f(\mathbf{x})}{\|f(\mathbf{y}) - f(\mathbf{x})\|} \cdot (f(\mathbf{y}) - f(\mathbf{x})) &= \frac{f(\mathbf{y}) - f(\mathbf{x})}{\|f(\mathbf{y}) - f(\mathbf{x})\|} \cdot \{f'(\mathbf{z})(\mathbf{y} - \mathbf{x})\} \\
\Rightarrow \|f(\mathbf{y}) - f(\mathbf{x})\| &\leq \|f'(\mathbf{z})(\mathbf{y} - \mathbf{x})\| \\
&\leq \|\nabla f_1(\mathbf{z}) \cdot (\mathbf{y} - \mathbf{x}), \cdots, \nabla f_m(\mathbf{z}) \cdot (\mathbf{y} - \mathbf{x})\| \\
&\leq \sum_{i=1}^{m} |\nabla f_i(\mathbf{z}) \cdot (\mathbf{y} - \mathbf{x})| \\
&\leq \sum_{i=1}^{m} \|\nabla f_i(\mathbf{z})\| \|\mathbf{y} - \mathbf{x}\| \\
&= \sum_{i=1}^{m} \sqrt{\sum_{j=1}^{n} D_j f_i(\mathbf{z})^2} \|\mathbf{y} - \mathbf{x}\| \\
&\leq \sum_{i=1}^{m} \sqrt{\sum_{j=1}^{n} M_{ji}^2} \|\mathbf{y} - \mathbf{x}\| = M\|\mathbf{y} - \mathbf{x}\|
\end{aligned}$$

where $M = \sum_{i=1}^m \sqrt{\sum_{j=1}^n M_{ji}^2}$ and M_{ji} is the upper bound of $D_j f_i$ for all i, j and any convex combination of $\mathbf{x}, \mathbf{y}$. $\nabla f_k(\mathbf{z}) = (D_1 f_k(\mathbf{z}), \ldots, D_n f_k(\mathbf{z}))$ and $f'(\mathbf{z})(\mathbf{v}) = (\nabla f_1(\mathbf{z}) \cdot \mathbf{v}, \ldots, \nabla f_m(\mathbf{z}) \cdot \mathbf{v})$. □

Remark 8. This proposition says any differentiable function has the local Lipschitz property. Here δ_0 depends on x_0, and M depends on δ_0 and x_0.

2.5 The total derivative

In the one-dimensional case, a function f with a derivative at c can be approximated near c by a linear polynomial. In fact, if $f'(c)$ exists, let $E_c(h)$ denote the difference

$$E_c(h) = \frac{f(c + h) - f(c)}{h} - f'(c) \quad \text{if } h \neq 0, \tag{33}$$

and let $E_c(0) = 0$. Then we have

$$f(c + h) = f(c) + f'(c)h + hE_c(h), \tag{34}$$

an equation which holds also for $h = 0$. This is called the first-order Taylor formula for approximating $f(c+h) - f(c)$ by $f'(c)h$. The error committed is $hE_c(h)$. From (33) we see that $E_c(h) \to 0$ as $h \to 0$. The error $hE_c(h)$ is said to be of small order than h as $h \to 0$.

We focus attention on two properties of formula (34). First, the quantity $f'(c)h$ is a linear function of h. That is, if we write $T_c(h) = f'(c)h$, then

$$T_c(ah_1 + bh_2) = aT_c(h_1) + bT_c(h_2).$$

Second, the error term $hE_c(h)$ is of smaller order than h as $h \to 0$. The total derivative of a function $\mathbf{f}$ from $\mathbf{R}^n$ to $\mathbf{R}^m$ will now be defined in such a way that it preserves these two properties.

Let $\mathbf{f} : S \to \mathbf{R^m}$ be a function defined on a set S in $\mathbf{R}^n$ with values in $\mathbf{R}^m$. Let $\mathbf{c}$ be an interior point of S, and let $B(\mathbf{c}; r)$ be an n-ball lying in S. Let $\mathbf{v}$ be a point in $\mathbf{R}^n$ with $\|\mathbf{v}\| < r$, so that $\mathbf{c} + \mathbf{v} \in B(\mathbf{c}; r)$.

Definition 14. [1] The function $\mathbf{f}$ is said to be differentiable at $\mathbf{c}$ if there exists a linear function $\mathbf{T_c} : \mathbf{R^n} \to \mathbf{R^m}$ such that

$$\mathbf{f}(\mathbf{c}+\mathbf{v}) = \mathbf{f}(\mathbf{c}) + \mathbf{T_c}(\mathbf{v}) + \|\mathbf{v}\|\mathbf{E_c}(\mathbf{v}), \tag{35}$$

where $\mathbf{E_c}(\mathbf{v}) \to \mathbf{0}$ as $\mathbf{v} \to \mathbf{0}$.

NOTE. Equation (35) is called a first-order Taylor formula. It is to hold for all $\mathbf{v}$ in $\mathbf{R}^n$ with $\|\mathbf{v}\| < r$. **The linear function $\mathbf{T_c}$ is called the total derivative of f at c.** We also write (35) in the form

$$\mathbf{f}(\mathbf{c}+\mathbf{v}) = \mathbf{f}(\mathbf{c}) + \mathbf{T_c}(\mathbf{v}) + o(\|\mathbf{v}\|) \quad \text{as } \mathbf{v} \to \mathbf{0}. \tag{36}$$

Theorem 9. *Assume* $\mathbf{f}$ *is differentiable at* $\mathbf{c}$ *with total derivative* $\mathbf{T_c}$. *Then the directional derivative* $\mathbf{f}'(\mathbf{c}; \mathbf{u})$ *exists for every* $\mathbf{u}$ *in* $\mathbf{R}^n$ *and we have*[2]

$$\mathbf{T_c}(\mathbf{u}) = \mathbf{f}'(\mathbf{c}; \mathbf{u}).$$

Proof. If $\mathbf{v} = \mathbf{0}$ then $\mathbf{f}'(\mathbf{c}; \mathbf{0})$ and $\mathbf{T_c}(\mathbf{0}) = \mathbf{0}$. Therefore we can assume that $\mathbf{v} \neq 0$. Take $\mathbf{v} = h\mathbf{u}$ in Taylor's formula (35), with $h \neq 0$, to get

$$\mathbf{f}(\mathbf{c}+h\mathbf{u}) = \mathbf{f}(\mathbf{c}) + h\mathbf{T_c}(\mathbf{u}) + |h|\|\mathbf{u}\|\mathbf{E_c}(\mathbf{v}).$$

Now divide by h and let $h \to 0$ to obtain $\mathbf{f}'(\mathbf{c}; \mathbf{u}) = \mathbf{T_c}(\mathbf{u})$. □

2.6 The total derivative expressed in terms of partial derivatives

The following theorem shows that the vector $\mathbf{f}'(\mathbf{c})(\mathbf{v})$ is a linear combination of the partial derivatives of $\mathbf{f}$.

Theorem 10. [3] *Let* $\mathbf{f} : S \to \mathbf{R^m}$ *be differentiable at an interior point* $\mathbf{c}$ *of* S, *where* $S \subseteq \mathbf{R}^n$. *If* $\mathbf{v} = v_1\mathbf{u}_1 + \cdots + v_n\mathbf{u}_n$, *where* $\mathbf{u}_1, \cdots, \mathbf{u}_n$ *are the unit coordinate vector in* $\mathbf{R}^n$, *then*

$$\mathbf{f}'(\mathbf{c})(\mathbf{v}) = \sum_{k=1}^{n} v_k D_k \mathbf{f}(\mathbf{c}).$$

In particular, if f *is real-valued* ($m = 1$) *we have*

$$f'(\mathbf{c})(\mathbf{v}) = \nabla f(\mathbf{c}) \cdot \mathbf{v}, \tag{37}$$

the dot product of $\mathbf{v}$ *with the vector* $\nabla f(\mathbf{c}) = (D_1 f(\mathbf{c}), \ldots, D_n f(\mathbf{c}))$.

[1] This is the Definition 12.2 from Apostle's Mathematical Analysis, Second Edition
[2] This is the Theorem 12.3 from Apostle's Mathematical Analysis, Second Edition.
[3] This is the Theorem 12.5 from Apostle's Mathematical Analysis, Second Edition

Proof. We use the linearity of $f'(\mathbf{c})$ to write

$$\begin{aligned}\mathbf{f}'(\mathbf{c})(\mathbf{v}) &= \sum_{k=1}^{n} \mathbf{f}'(\mathbf{c})(v_k \mathbf{u_k}) = \sum_{k=1}^{n} v_k \mathbf{f}'(\mathbf{c})(\mathbf{u_k}) \\ &= \sum_{k=1}^{n} v_k \mathbf{f}'(\mathbf{c}; \mathbf{u_k}) = \sum_{k=1}^{n} v_k D_k \mathbf{f}(\mathbf{c})\end{aligned}$$

□

NOTE. The vector $\nabla f(\mathbf{c})$ in (37) is called the gradient vector of f at $\mathbf{c}$. It is defined at each point where the partials $D_1 f, \ldots, D_n f$ exist. The Taylor formula for real-valued f now takes the form

$$\mathbf{f}(\mathbf{c}+\mathbf{v}) = \mathbf{f}(\mathbf{c}) + \nabla \mathbf{f}(\mathbf{c}) \cdot \mathbf{v} + o(\|\mathbf{v}\|) \quad \text{as } \mathbf{v} \to \mathbf{0}.$$

2.7 The Jacobian matrix

Let $\mathbf{f}$ be a function with values in $\mathbf{R}^m$ which is differentiable at a point $\mathbf{c}$ in $\mathbf{R}^n$, and let $\mathbf{T} = \mathbf{f}'(\mathbf{c})$ be the total derivative of $\mathbf{f}$ at $\mathbf{c}$. To find the matrix of $\mathbf{T}$ we consider its action on the unit coordinate vectors $\mathbf{u_1}, \ldots, \mathbf{u_n}$. By Theorem (9) we have

$$\mathbf{T_c}(\mathbf{u_k}) = \mathbf{f}'(\mathbf{c}; \mathbf{u_k}) = D_k \mathbf{f}(\mathbf{c}).$$

To express this as a linear combination of the unit coordinate vectors $\mathbf{e_1}, \ldots, \mathbf{e_m}$ of $\mathbf{R}^m$ we write $\mathbf{f} = (f_1, \ldots, f_m)$ so that $D_k \mathbf{f} = (D_k f_1, \ldots, D_k f_m)$, and hence

$$\mathbf{T_c}(\mathbf{u_k}) = D_k \mathbf{f}(\mathbf{c}) = \sum_{i=1}^{m} D_k f_i(\mathbf{c} \mathbf{e}_i).$$

Therefore the matrix of $\mathbf{T}$ is $m(\mathbf{T}) = (D_k f_i(\mathbf{c}))$. This is called the Jacobian matrix of $\mathbf{f}$ at $\mathbf{c}$ and is denoted by $\mathbf{Df}(\mathbf{c})$. That is,

$$\mathbf{Df}(\mathbf{c}) = \begin{bmatrix} D_1 f_1(\mathbf{c}) & D_2 f_1(\mathbf{c}) & \cdots & D_k f_1(\mathbf{c}) \\ D_1 f_2(\mathbf{c}) & D_2 f_2(\mathbf{c}) & \cdots & D_k f_2(\mathbf{c}) \\ \vdots & \vdots & & \vdots \\ D_1 f_m(\mathbf{c}) & D_2 f_m(\mathbf{c}) & \cdots & D_k f_m(\mathbf{c}) \end{bmatrix} \tag{38}$$

The kth row of the Jacobian matrix (38) is a vector in $\mathbf{R}^n$ called the gradient vector of f_k, denoted by $\nabla f_k(\mathbf{c})$. That is,

$$\nabla f_k(\mathbf{c}) = (D_1 f_k(\mathbf{c}), \ldots, D_n f_k(\mathbf{c})).$$

In the special case when f is real-valued ($m = 1$), the Jacobian matrix consists of only one row. In this case, $\mathbf{D}f(\mathbf{c}) = \nabla f(\mathbf{c})$[4], and Equation (37) shows that the directional derivative $f'(\mathbf{c}; \mathbf{v})$ is the dot product of the gradient vector $\nabla f(\mathbf{c})$ with the direction $\mathbf{v}$. In terms of values, when f is real-valued, its directional derivative $f'(\mathbf{c}; \mathbf{v})$ is a scalar while its gradient $\nabla f(\mathbf{c})$ is a vector. In addition, a directional derivative depends on both a point and a direction while a gradient is only with respect to a point, i.e., the direction of the tangent line of f at that point.

For a vector-valued function $\mathbf{f} = (f_1, \ldots, f_m)$ we have

$$\mathbf{f}'(\mathbf{c})(\mathbf{v}) = \mathbf{f}'(\mathbf{c}; \mathbf{v}) = \sum_{k=1}^{m} f_k'(\mathbf{c}; \mathbf{v}) \mathbf{e}_k = \sum_{k=1}^{m} \{\nabla f_k(\mathbf{c}) \cdot \mathbf{v}\} \mathbf{e}_k, \tag{39}$$

[4]This subsection is taken from Section 12.8 of Apostle's Mathematical Analysis, Second Edition. Nowadays $\mathbf{D}f(\mathbf{c}) = \nabla f(\mathbf{c})^T$ is more common.

so the vector $\mathbf{f}'(\mathbf{c})(\mathbf{v})$ has components

$$(\nabla f_1(\mathbf{c}) \cdot \mathbf{v}, \ldots, \nabla f_m(\mathbf{c}) \cdot \mathbf{v}).$$

Thus, the components of $\mathbf{f}'(\mathbf{c})(\mathbf{v})$ are obtained by taking the dot product of the successive rows of the Jacobian matrix with the vector $\mathbf{v}$, i.e., $\mathbf{f}'(\mathbf{c})(\mathbf{v}) = \mathbf{Df}(\mathbf{c})\mathbf{v}$.

NOTE. Equation (39), used in conjunction with the triangle inequality and the Cauchy-Schwarz inequality, gives us

$$\|\mathbf{f}'(\mathbf{c})(\mathbf{v})\| = \left\| \sum_{k=1}^{m} \{\nabla f_k(\mathbf{c}) \cdot \mathbf{v}\} \mathbf{e}_k \right\| \leq \sum_{k=1}^{m} |\nabla f_k(\mathbf{c}) \cdot \mathbf{v}| \leq \|\mathbf{v}\| \sum_{k=1}^{m} \|\nabla f_k(\mathbf{c})\|.$$

Therefore we have

$$\|\mathbf{f}'(\mathbf{c})(\mathbf{v})\| \leq M\|\mathbf{v}\|, \tag{40}$$

where $M = \sum_{k=1}^{m} \|\nabla f_k(\mathbf{c})\|$.

2.8 Mean-Value Theorem

This general Mean-Value theorem for vector-valued functions, including the proof, is exactly the Theorem 12.9 in Tom Apostol's Mathematical Analysis, Second Edition.

In the statement of the theorem we use the notation $L(x, y)$ to denote the line segment joining two points $\mathbf{x}$ and $\mathbf{y}$ in R^n. That is,

$$L(\mathbf{x}, \mathbf{y}) = \{t\mathbf{x} + (1-t)\mathbf{y} : 0 \leq t \leq 1\}.$$

Theorem 11 (Mean-Value Theorem for vector-valued functions)**.** *Let S be an open subset of $\mathbf{R}^n$ and assume that $f : S \to \mathbf{R}^m$ is differentiable at each point of S.* ***Let $\mathbf{x}$ and $\mathbf{y}$ be two points in S such that $L(\mathbf{x}, \mathbf{y}) \subseteq S$. Then for every vector $\mathbf{a}$ in $\mathbf{R}^m$ there is a point $\mathbf{z}$ in $L(\mathbf{x}, \mathbf{y})$ such that***

$$\mathbf{a} \cdot \{f(\mathbf{y}) - f(\mathbf{x})\} = \mathbf{a} \cdot \{f'(\mathbf{z})(\mathbf{y} - \mathbf{x})\}. \tag{41}$$

Proof. Let $\mathbf{u} = \mathbf{y} - \mathbf{x}$. Since S is open and $L(\mathbf{x}, \mathbf{y}) \subseteq S$, there is a $\delta > 0$ such that $\mathbf{x} + t\mathbf{u} \in S$ for all real $t \in (-\delta, 1 + \delta)$. Let $\mathbf{a}$ be a fixed vector in $\mathbf{R}^m$ and let F be the real-valued function defined on $(-\delta, 1 + \delta)$ by the equation

$$F(t) = \mathbf{a} \cdot f(\mathbf{x} + t\mathbf{u})$$

Then F is differentiable on $(-\delta, 1 + \delta)$ and its derivative is given by

$$F'(t) = \mathbf{a} \cdot \{f'(\mathbf{x} + t\mathbf{u})\mathbf{u}\}.$$

By the usual Mean-Value Theorem we have

$$F(1) - F(0) = F'(\theta), \quad \text{where } 0 < \theta < 1$$

Now

$$F'(\theta) = \mathbf{a} \cdot \{f'(\mathbf{x} + \theta\mathbf{u})\mathbf{u}\} = \mathbf{a} \cdot \{f'(\mathbf{z})(\mathbf{y} - \mathbf{x})\}.$$

where $\mathbf{z} = t\mathbf{x} + (1-t)\mathbf{y} \in L(\mathbf{x}, \mathbf{y})$. Plus, $F(1) - F(0) = \mathbf{a} \cdot \{f(\mathbf{y}) - f(\mathbf{x})\}$, so we obtain the claimed result. Of course, the point $\mathbf{z}$ depends on F, and hence on $\mathbf{a}$. $\square$

NOTE. This theorem only holds for any $\mathbf{x}, \mathbf{y}$ whose convex combinations are in S, so if there is a point on the line segment connecting $\mathbf{x}, \mathbf{y}$ which lies outside S, this theorem will not hold for this pair of $\mathbf{x}, \mathbf{y}$. If S is convex, then $L(\mathbf{x}, \mathbf{y}) \subseteq S$ for all $\mathbf{x}, \mathbf{y}$ in S, so (41) naturally holds for all $\mathbf{x}$ and $\mathbf{y}$ in S.

In Pugh's Real Mathematical Analysis, Second Edition, Ex 17(b) of Chapter 5 says that if the derivatives $(\mathbf{DF})_\mathbf{z}$ with $\mathbf{z} \in S$ forms a convex set, the multi-dimensional MVT holds without dot products, i.e.,

$$f(\mathbf{y}) - f(\mathbf{x}) = (\mathbf{Df})_\mathbf{z}(\mathbf{y} - \mathbf{x}). \tag{42}$$

where $\mathbf{z} \in L(\mathbf{x}, \mathbf{y})$.

2.8.1 MVT and Lipschitz condition

If f is real-valued ($m = 1$), we can take $a = 1$ in (41) to obtain

$$f(\mathbf{y}) - f(\mathbf{x}) = \nabla f(\mathbf{z}) \cdot (\mathbf{y} - \mathbf{x}). \tag{43}$$

If $\mathbf{f}$ is vector-valued and if $\mathbf{a}$ is a unit vector in $\mathbf{R}^m$, $\|\mathbf{a}\|= 1$, Eq (41) and the Cauchy-Schwarz give us

$$\|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})\| \leq \|\mathbf{f}'(\mathbf{z})(\mathbf{y} - \mathbf{x})\|.$$

Using (40) we obtain the inequality

$$\|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})\| \leq M\|\mathbf{y} - \mathbf{x}\|,$$

where $M = \sum_{k=1}^{m} \|\nabla f_k(\mathbf{z})\|$. Note that M depends on $\mathbf{z}$ and hence on $\mathbf{x}$ and $\mathbf{y}$. If S is convex and if all the partial derivatives $D_j f_k$ are bounded on S, then there is a constant $A > 0$ such that

$$\|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})\| \leq A\|\mathbf{y} - \mathbf{x}\|, \quad \forall \mathbf{x}, \mathbf{y} \in S.$$

In other words, $\mathbf{f}$ satisfies a Lipschitz condition on S. Charles Pugh gives more knowledge about the connection between Lipschitz condition and the upper bound of derivatives in his book, Real Mathematical Analysis, Second Edition. The following subsection is taken from Theorem 11 of its Chapter 5.

2.8.2 Lipschitz condition and the upper bound of derivatives

Theorem 12. *If $\mathbf{f} : S \to \mathbf{R}^m$ is differentiable on S and the segment $[\mathbf{x}, \mathbf{y}]$ is contained in S then*

$$\|\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})\| \leq A\|\mathbf{y} - \mathbf{x}\|.$$

where $A = \sup\{\|(\mathbf{Df})_{\mathbf{v}}\| \mid \mathbf{v} \in U\}$.

Proof. Fix any unit vector $\mathbf{u} \in \mathbf{R}^m$. The function

$$g(t) = \langle u, \mathbf{f}(\mathbf{x} + \mathbf{t}(\mathbf{y} - \mathbf{x}))\rangle$$

is differentiable and we can calculate its derivative. By the one-dimensional Mean Value Theorem this gives some $\theta \in (0, 1)$ such that $g(1) - g(0) = g'(\theta)$. That is,

$$\begin{aligned}
\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x}) = g'(\theta) &= \langle u, (\mathbf{Df})_{\mathbf{x}+\theta(\mathbf{y}-\mathbf{x})}(\mathbf{y} - \mathbf{x})\rangle \\
&\leq \|\mathbf{Df})_{\mathbf{x}+\theta(\mathbf{y}-\mathbf{x})}(\mathbf{y} - \mathbf{x}\| \quad (\because \text{Cauchy-Schwarz inequality}) \\
&\leq \|\mathbf{Df})_{\mathbf{x}+\theta(\mathbf{y}-\mathbf{x})}\|_2 \|\mathbf{y} - \mathbf{x}\|_2 \\
&\leq A\|\mathbf{y} - \mathbf{x}\|
\end{aligned}$$

where $A = \sup_{\mathbf{v}\in S}\|(\mathbf{Df})_{\mathbf{v}}\|_2$. The second inequality follows from the definition and properties of the induced matrix norm by the norm $\|\cdot\|$. It is sometimes called the operator norm or lub norm (least upper bound norm) associated with the vector norm $\|\cdot\|$, see Definition 5.6.1 and Theorem 5.6.2 in [4]. ☐

The following holds for any $\mathbf{z} = t\mathbf{x} + (1 - t)\mathbf{y} \in L(\mathbf{x}, \mathbf{y})$

$$M = \sum_{k=1}^{m} \|\nabla f_k(\mathbf{z})\| \geq \left(\sum_{k=1}^{m} \|\nabla f_k(\mathbf{z})\|^2\right)^{\frac{1}{2}} = \|(\mathbf{Df})_{\mathbf{z}}\|_F = \sqrt{\sum_{i=1}^{n} \sigma_i^2} \geq \sigma_1 = \|(\mathbf{Df})_{\mathbf{z}}\|_2$$

where $\{\sigma_i\}$ are the singular values of $(\mathbf{Df})_{\mathbf{z}}$ and $\sigma_1 \geq \ldots \geq \sigma_n \geq 0$. Furthermore,

$$A = \sup_{\mathbf{v}\in S}\|(\mathbf{Df})_{\mathbf{v}}\|_2 \leq \sup_{\mathbf{v}\in S}\sum_{k=1}^{m}\|\nabla f_k(\mathbf{v})\| = \sup_{\mathbf{v}\in S} M_{\mathbf{v}}.$$

Hence, we get A as the Lipschitz constant which is stronger than $\sup_{\mathbf{v}\in S} M_{\mathbf{v}}$.

2.9 Intermediate-value theorem for derivatives

Theorem 13 (Intermediate-value theorem for derivatives)**.** *Assume that f is defined on a compact interval $[a,b]$ and that f has a derivative (finite or infinite) at each interior point. Assume also that f has finite one-sided derivatives $f'_+(a)$ and $f'_-(b)$ at the endpoints, with $f'_+(a) \neq f'_-(b)$. Then, if c is a real number between $f'_+(a)$ and $f'_-(b)$, there exists at least one interior point x such that $f'(x) = c$.*

Proof. See Theorem 5.16 on page 112 of Tom Apostle's Mathematical Analysis. □

Theorem 14. *Assume f' exists and is monotonic on an open interval (a,b). Then f' is continuous on (a,b).*

Proof. See Theorem 5.18 on page 112 of Tom Apostle's Mathematical Analysis. □

NOTE. The derivative of a differentiable function need not be continuous. A counterexample is

$$f(x) = \begin{cases} x^2 \sin\frac{1}{x}, & \text{if } x \neq 0 \\ 0, & \text{otherwise.} \end{cases}$$

If $x \neq 0$, $f'(x) = 2x\sin\frac{1}{x} - \cos\frac{1}{x}$. Otherwise, $f'(0) = \lim_{x\to 0} \frac{x^2\sin\frac{1}{x} - 0}{x - 0} = \lim_{x\to 0} x\sin\frac{1}{x} = 0$. However, $\lim_{x\to 0} f'(x)$ does not exist. Hence, $f'(x)$ is discontinuous at 0.

2.10 Strongly convex function

We say that a function f: dom$f \to \mathbf{R}$ is α-strongly convex over a convex set $\mathcal{K} \subset \mathrm{dom}f$ with respect to $\|\cdot\|$ if it satisfies the following two equivalent conditions[3]

1. $\forall x, y \in \mathcal{K}$:

$$f(y) \geq f(x) + \langle y - x, \nabla f(x)\rangle + \frac{\alpha}{2}\|y - x\|^2.$$

2. $\forall x, y \in \mathcal{K}, \gamma \in [0,1]$:

$$f(\gamma x + (1-\gamma)y) \leq \gamma f(x) + (1-\gamma)f(y) - \frac{\alpha}{2}\gamma(1-\gamma)\|y - x\|^2.$$

The above definition(part 1) and the first-order optimality imply that for a α-strongly convex function f, if $x^* = \mathrm{argmin}_{x\in\mathcal{K}}$,

$$f(x) - f(x^*) \geq \frac{\alpha}{2}\|x - x^*\|^2. \tag{44}$$

This is exactly the definition of **quadratic functional growth**. Since it is obtained from strong convexity of f, the quadratic functional growth condition is weaker than the condition of strong convexity. Then we have the following derivation,

$$\begin{aligned} \sqrt{\frac{2}{\alpha}(f(x) - f(x^*))} &\geq \|x - x^*\| \\ \sqrt{\frac{2}{\alpha}(f(x) - f(x^*))} \cdot \|\nabla f(x)\|_* &\geq \|x - x^*\| \cdot \|\nabla f(x)\|_* \\ &\geq \langle x - x^*, \nabla f(x)\rangle \\ &\geq f(x) - f(x^*), \end{aligned}$$

where the first inequality follows from (44), the third inequality from Hölder's inequality and the last one from convexity of f. Thus we obtain that at any point $x \in \mathcal{K}$ it holds that

$$\|\nabla f(x)\|_* \geq \sqrt{\frac{\alpha}{2}} \cdot \sqrt{f(x) - f(x^*)}$$

which implies that the magnitude of the gradient of f at point x, $\|\nabla f(x)\|_*$ is at least of the order of the square-root of the approximation error at x, $f(x) - f(x^*)$. **This result only requires f satisfies the quadratic functional growth condition which is weaker than the condition of strong convexity.**

In the general situation of $f : \mathbf{R}^n \to \mathbf{R}^m$. There is a form of the mean value theorem: $a \cdot (f(y) - f(x)) = a \cdot (f'(z)(y - x))$ for every vector $a \in \mathbf{R}^m$ where $\cdot$ denotes dot product. This result is Theorem 12.9 in Tom Apostol's Mathematical Analysis (Second Edition), page No. 355.

Now consider a simpler case $f : \mathbf{R}^n \to \mathbf{R}$. Is the mean value theorem hold in vector space if z on the line segment connecting x and y? Namely, the following equation,

$$f(y) - f(x) = \langle f(z)', (y - x) \rangle$$

where $z = (1 - c)x + cy$ and $0 \le c \le 1$.

2.11 Strongly convex set

[3]

2.12 Sublevel sets

The α-sublevel set of a function f: $\mathbf{R}^n \to \mathbf{R}$ is defined as[2]

$$C_\alpha = \{x \in \mathrm{dom} f \mid f(x) \le \alpha\}.$$

Proposition 10. *Sublevel sets of a convex function are convex, for any value of α.*

2.13 Coercive Functions and Global Min

Theorem 15 (**Continuous functions on closed bounded domain have global extrema**)**.** *A continuous function f on a closed bounded domain D has a global min and max.*

Definition 15. f is coercive if $\lim_{\|x\| \to +\infty} f(x) = +\infty$.

Example:

1.
$$f(x) = x_1^6 + x_2 - x_1 x_2^3$$
Let $x = (0, x_2)$. When $x_2 \to +\infty$, $f(x) \to -\infty$. Hence, $f(x)$ is not coercive.

2.
$$f(x) = x_1^4 + x_2^2 - 4x_1 x_2$$
$x_1^4 + x_2^2$ is dominant, so $f(x)$ is coercive.

□

Theorem 16 (**Continuous and coercive functions have a global min**)**.** *Suppose f is continuous over $\mathbf{R}^n$ and coercive. f must have a global min.*

Proof. Since f is coercive, there exists $r > 0$ such that

$$f(x) > f(0), \quad \forall x \in \{x \mid \|x\| > r\}$$

According to Theorem 15, there is at least one global minimum $x^\star$ in $\bar{B}(0, r)$(closure of B). Thus,

$$f(x) \ge f(x^\star), \quad \forall x \in \bar{B}(0, r).$$

In particular,

$$f(0) \geq f(x^\star)$$

. Hence,

$$f(x) > f(0) \geq f(x^\star), \quad \forall x.$$

This completes the proof.

$\square$

3 Functional analysis

3.1 Important theorems

Theorem 17 (Bolzano-Weierstrass theorem)**.** *Let $(a_n)_{n=0}^{\infty}$ be a bounded sequence (i.e., there exists a real number $M > 0$ such that $|a_n| \leq M$ for all $n \in \mathbf{N}$). Then there is at least one subsequence of $(a_n)_{n=0}^{\infty}$ which converges.*

Theorem 18 (Heine-Borel theorem for the line)**.** *Let X be a subset of $\mathbf{R}$. Then the following two statements are equivalent:*

- *X is closed and bounded.*
- *Given any sequence $(a_n)_{n=0}^{\infty}$ of real numbers which takes values in X (i.e., $a_n \in X$ for all n), there exists a subsequence $(a_{n_j})_{j=0}^{\infty}$ of the original sequence, which converges to some number L in X.*

Remark 9. In the language of metric space topology, this theorem asserts that every subset of the real line which is closed and bounded, is also compact.

A more general version of this theorem,

Theorem 19 (Heine-Borel theorem)**.** *Let $(\mathbf{R}^n, d)$ be a Euclidean space with either the Euclidean metric, the taxicab metric, or the sup norm metric. Let $\mathbf{E}$ be a subset of $\mathbf{R}^n$. Then $\mathbf{E}$ is compact if and only if it is closed and bounded.*

3.2 Taylor series expansion

The following 2 Taylor series expansions is taken from a cheatsheet[5](see Page 9).

$$\begin{aligned} \ln(1+x) &= x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4 + \cdots = \sum_{i=1}^{\infty}(-1)^{i+1}\frac{x^i}{i} \\ \ln\frac{1}{1-x} &= x + \frac{1}{2}x^2 + \frac{1}{3}x^3 + \frac{1}{4}x^4 + \cdots = \sum_{i=1}^{\infty}\frac{x^i}{i} \end{aligned} \tag{45}$$

Adding them together, we get

$$\ln\frac{1+x}{1-x} = 2(x + \frac{1}{3}x^3 + \frac{1}{5}x^5 + \cdots) = \tag{46}$$

replacing x with $\frac{1}{x}$, we have

$$\ln\frac{x+1}{x-1} = 2(\frac{1}{x} + \frac{1}{3x^3} + \frac{1}{5x^5} + \cdots) = \tag{47}$$

this is exactly the formula at the bottom line of Page 236 in [5].

[5]This cheatsheet in theoretical computer science, available at `https://www.tug.org/texshowcase/cheat.pdf`

3.3 Closedness

Do linear transformations preserve closedness? The answer is No[6]. For a counterexample, consider $T : \mathbb{R}^2 \to \mathbb{R}$ given by $T(x, y) = x$, and the closed subset $A := \{(x, y) \in \mathbb{R}^2 \mid xy = 1\}$ of $\mathbb{R}^2$. Then $T(A) = \mathbb{R} \setminus \{0\}$ is not closed. Conversely,

Proposition 11. *if $S \in \mathbf{R}^m$ is closed and $A \in \mathbf{R}^{m\times n}$, then $A^{-1}(S) = \{x : Ax \in S\}$, called the preimage of S under A, is closed.*

The following proof is from this discussion[7].

Proof. For a function $f : C \to D$, the preimage of a closed set D, i.e., C, is closed provided that f is continuous. Every linear transformation $A : \mathbf{R}^m \to \mathbf{R}^n$ is continuous, and has a matrix representation. We can conclude that the preimage of any closed set under a linear transformation is closed. □

The definition for polyhedra(the plural of polyhedron) in B & V's Convex Optimization[8] is $\mathcal{P} = \{x|Ax \preceq b, Cx = d\}$. Any polyhedron is closed and convex since halfspaces and hyperplanes are closed, and the fact that the intersection of closed sets is itself closed[9].

3.4 Closed functions

See A.3.3 on Page 639 of B&V's Convex Optimization book[8].

If $f : \mathbf{R}^n \to \mathbf{R}$ is continuous, and $\mathrm{dom}\, f$ is closed, then f is closed. If $f : \mathbf{R}^n \to \mathbf{R}$ is continuous, with $\mathrm{dom}\, f$ open, then f is closed if and only if f converges to ∞ along every sequence converging to a boundary point of $\mathrm{dom}\, f$. In other words, if $\lim_{i\to\infty} x_i = x \in \mathbf{bd}\,\mathrm{dom}\, f$, with $x_i \in \mathrm{dom}\, f$, we have $\lim_{i\to\infty} f(x_i) = \infty$.

3.5 Proper convex (concave) functions

In mathematical analysis, in particular the subfields of convex analysis and optimization, a proper convex function is an extended real-valued convex function with a non-empty domain, that never takes on the value $-\infty$ and also is not identically equal to $+\infty$.[10]

In the language of convex analysis, a convex function is **proper** if its effective domain is nonempty and its epigraph contains no vertical lines. A concave function is proper if its effective domain is nonempty and its hypograph contains no vertical lines. (A **vertical line** in $X \times \mathbf{R}$ is a set of the form $\{x\} \times \mathbf{R}$ for some $x \in X$.) That is, **a convex f is proper if $f(x) < \infty$ for at least one x and $f(x) > -\infty$ for every x.** Every proper convex function is gotten by taking a **finite-valued** convex function defined on some nonempty convex set and extending it to all of X as above. For a proper function, its effective domain is the set of points where it is finite.[11]

3.6 Derivative of a functional

This part is taken from the Appendix D of the famous PRML book. I will expand this part in the future.

We can think of a function $y(x)$ as being an operator that, for any input value x, returns an output value y. In the same way, we can define a functional $F[y]$ to be an operator that takes a function $y(x)$ and returns an output value F.

[6] https://math.stackexchange.com/questions/2762578/preservation-of-closed-sets-under-linear-transformation/2762598
[7] https://math.stackexchange.com/questions/1921018/is-the-preimage-of-a-closed-set-closed/1921020
[8] https://web.stanford.edu/~boyd/cvxbook/
[9] https://math.stackexchange.com/questions/1919425/how-to-prove-that-polyhedron-is-a-closed-set
[10] https://handwiki.org/wiki/Proper_convex_function
[11] https://healy.econ.ohio-state.edu/kcb/Ec181/Lecture13.pdf

We denote the functional derivative of $F[f(x)]$ with respect to $f(x)$ by $\delta F/\delta f(x)$, and define it by the following relation:

$$F[y(x)+\epsilon\eta(x)] = F[y(x)] + \epsilon \int \frac{\delta F}{\delta f(x)}\eta(x)\,\mathrm{d}\,x + O(\epsilon^2).$$

4 Subdifferential

The definition of the subdifferential of $f(x)$ is

$$\partial f(x) = \{\mathbf{v}\in\mathbf{R}^n |\ f(y) \geq f(x) + \mathbf{v}^T(y-x),\ \forall x, y \in \mathrm{dom}(f)\}$$

For $f(x) = \|x\|_2,\ x\in\mathbf{R}^n$, when $x \neq 0$, $f(x)$ is differentiable and its subdifferential is the set of its gradients, i.e., $\nabla f(x) = x/\|x\|_2$. But it is not differentiable at $x = 0$, by definition,

$$\begin{aligned} &\partial f(0) = \{\mathbf{v}\in\mathbf{R}^n | \|y\|_2 \geq \mathbf{v}^T y,\ \forall y \in \mathrm{dom}(f)\} \\ \Longrightarrow &\partial f(0) = \{\mathbf{v}\in\mathbf{R}^n | \|\mathbf{v}\|_2 \leq 1\} \end{aligned}$$

We can use this result to derive the solution to group lasso.

Proposition 12. *The proximal operator with* $\|\cdot\|$ *has a closed form solution:*

$$\underset{\mathbf{x}}{\operatorname{argmin}} \frac{1}{2}\|\mathbf{x}-\mathbf{y}\|^2 + \lambda\|\mathbf{x}\| = \max(\|\mathbf{y}\|-\lambda, 0)\frac{\mathbf{y}}{\|\mathbf{y}\|}, \tag{48}$$

where $\lambda > 0$ *and* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

Proof. Recall the definition of the subdifferential of a function $f(\mathbf{x})$ with regard to $\mathbf{x}$:

$$\partial f(\mathbf{x}) = \{\mathbf{v}\in\mathbf{R}^n |\ f(\mathbf{y}) \geq f(\mathbf{x}) + \mathbf{v}^{\mathrm{T}}(\mathbf{y}-\mathbf{x}),\ \forall \mathbf{x}, \mathbf{y} \in \mathrm{dom}(f)\}$$

Given $f(\mathbf{x}) = \|\mathbf{x}\|_2$, its gradient at $\mathbf{x}\neq\mathbf{0}$ is $\nabla f(\mathbf{x}) = \mathbf{x}/\|\mathbf{x}\|_2$. Its subdifferential at $\mathbf{x} = \mathbf{0}$ is

$$\partial f(\mathbf{0}) = \{\mathbf{v}\in\mathbf{R}^n | \|\mathbf{y}\|_2 \geq \mathbf{v}^T\mathbf{y},\ \forall \mathbf{y}\in\mathrm{dom}(f)\} \Rightarrow \partial f(\mathbf{0}) = \{\mathbf{v}\in\mathbf{R}^n | \|\mathbf{v}\|_2 \leq 1\}.$$

Taking derivatives of $\ell_\lambda(\mathbf{x})$ w.r.t $\mathbf{x}$, according to the first-order optimality condition, we get

$$\partial\ell_\lambda(\mathbf{x}) = \mathbf{x} - \mathbf{y} + \lambda\partial\|\mathbf{x}\|_2 \ni \mathbf{0} \tag{49}$$

We need to consider two cases: $\mathbf{x}\neq\mathbf{0}$ and $\mathbf{x} = \mathbf{0}$.

- Case 1: When $\mathbf{x}\neq\mathbf{0}$, $\partial\|\mathbf{x}\|_2 = \mathbf{x}/\|\mathbf{x}\|_2$. Thus,

$$\begin{aligned} \mathbf{x} - \mathbf{y} + \lambda\frac{\mathbf{x}}{\|\mathbf{x}\|_2} = \mathbf{0} &\Longleftrightarrow (1 + \frac{\lambda}{\|\mathbf{x}\|_2})\mathbf{x} = \mathbf{y}\ (\mathbf{x} \text{ and } \mathbf{y} \text{ share the same direction}) \\ &\Longleftrightarrow (1+\lambda/\|\mathbf{x}\|_2)\|\mathbf{x}\|_2 = \|\mathbf{y}\|_2 \Longleftrightarrow \|\mathbf{x}\|_2 + \lambda = \|\mathbf{y}\|_2 \\ &\Longleftrightarrow \|\mathbf{x}\|_2 = \|\mathbf{y}\|_2 - \lambda\ (\text{this is the amplitude of } \mathbf{x}) \\ &\Longleftrightarrow \mathbf{x} = (\|\mathbf{y}\|_2-\lambda)\frac{\mathbf{y}}{\|\mathbf{y}\|_2} = (1 - \frac{\lambda}{\|\mathbf{y}\|_2})\mathbf{y}\ (\text{with } \|\mathbf{y}\|_2 > \lambda), \end{aligned}$$

  where the second last "$\Longleftrightarrow$" indicates $\|\mathbf{y}\|_2 > \lambda$ since $\mathbf{x}\neq\mathbf{0}$. Note that the "$\Longleftarrow$" of the second "$\Longleftrightarrow$" follows from the last "$\Longleftrightarrow$", i.e., $\mathbf{x} = (\|\mathbf{y}\|_2-\lambda)\frac{\mathbf{y}}{\|\mathbf{y}\|_2}$.

- Case 2: When $\mathbf{x} = \mathbf{0}$, after substituting it into (49), we have

$$\mathbf{x} = \mathbf{0} \Longleftrightarrow -\mathbf{y} + \lambda\mathbf{v} \ni \mathbf{0} \Longleftrightarrow \mathbf{y}\in\lambda\mathbf{v} \Longleftrightarrow \|\mathbf{y}\|_2 \leq \lambda,$$

  where $\mathbf{v}\in\partial f(\mathbf{0})$. The last "$\Longleftrightarrow$" follows from $\|\mathbf{v}\|_2 \leq 1$.

Combining the above two cases, we get the solution to the group lasso problem

$$\mathbf{x} = \begin{cases} (1 - \frac{\lambda}{\|\mathbf{y}\|_2})\mathbf{y}, & \text{if } \|\mathbf{y}\|_2 > \lambda \\ \mathbf{0}, & \text{if } \|\mathbf{y}\|_2 \leq \lambda. \end{cases}$$

For notational simplicity,

$$\mathbf{x} = (1 - \frac{\lambda}{\|\mathbf{y}\|_2})_+ \mathbf{y} = \max\{1 - \frac{\lambda}{\|\mathbf{y}\|_2}, 0\}\mathbf{y}$$

□

Remark 10. We can also show the sufficiency of $\|\mathbf{y}\| \leq \lambda$ for $\mathbf{x} = \mathbf{0}$ to be a solution by contradiction. Specifically, given $\|\mathbf{y}\| \leq \lambda$, we suppose $\mathbf{x} \neq \mathbf{0}$, then $\partial f(\mathbf{x}) = \mathbf{x}/\|\mathbf{x}\|_2$ holds. According to the first-order optimality condition,

$$\mathbf{x} - \mathbf{y} + \lambda \mathbf{x}/\|\mathbf{x}\|_2 \ni \mathbf{0} \Longrightarrow \mathbf{x}(1 + \lambda/\|\mathbf{x}\|_2) \ni \mathbf{y} \Longrightarrow \|\mathbf{x}\|_2 + \lambda = \|\mathbf{y}\|_2 \leq \lambda \Longrightarrow \|\mathbf{x}\|_2 \leq 0$$

which contradicts $\mathbf{x} \neq \mathbf{0}$. Hence, $\mathbf{x} = \mathbf{0}$ must hold.

Proposition 13. *The subdifferential of the dual norm of a given norm, e.g. $\|x\|$, can be expressed as*

$$\partial \|x\|_\star = \underset{\|z\| \leq 1}{\operatorname{argmax}} \ x^T z$$

where $\|x\|_\star$ is the dual norm of $\|x\|$.[12]

Proof. Suppose that the given norm is $\|x\|_p$ with $p \geq 1$, then its dual norm is

$$\|x\|_\star = \|x\|_q = \max_{\|z\|_p \leq 1} \ x^T z$$

where $\frac{1}{p} + \frac{1}{q} = 1$. This is a known result. We also know the following famous Hölder inequality.

$$x^T z \leq \|x\|_p \|z\|_q.$$

Thus, we have $x^T z \leq \|x\|_\star \|z\|$. We first show $\partial \|x\|_\star \subset \operatorname{argmax}_{\|z\| \leq 1} \ x^T z$. If $g \in \partial \|x\|_\star$, we substitute $y = 0$ and $y = 2x$ into the definition of $\partial \|x\|_\star$, i.e. $\partial \|x\|_\star = \{g \in \mathbf{R}^n \mid \|y\|_\star \geq \|x\|_\star + g^T(y - x)\}$, respectively. We get $g^T x \geq \|x\|_\star$ and $g^T x \leq \|x\|_\star$, which gives $g^T x = \|x\|_\star$. Substituting this into the definition of $\partial \|x\|_\star$, we get

$$\begin{aligned} \|y\|_\star &\geq \|x\|_\star + g^T y - g^T x\} \\ &= \|x\|_\star + g^T y - \|x\|_\star \\ &= g^T y \end{aligned}$$

By Hölder inequality, we have $\|y\|_\star \|g\| \geq g^T y$ for all y. To ensure $\|y\|_\star \geq g^T y$ for all y, we need to let $\|y\|_\star \|g\| \leq \|y\|_\star$ which gives $\|g\| \leq 1$. Now $\|g\| \leq 1$ and $g^T x = \|x\|_\star$ are obtained, which demonstrates $g \in \operatorname{argmax}_{\|z\| \leq 1} \ x^T z$ by definition, i.e. $\partial \|x\|_\star \subset \operatorname{argmax}_{\|z\| \leq 1} \ x^T z$.

Now we show $\operatorname{argmax}_{\|z\| \leq 1} \ x^T z \subset \partial \|x\|_\star$. Suppose $z_+ \in \operatorname{argmax}_{\|z\| \leq 1} \ x^T z$, then $\|x\|_\star = x^T z_+$ and $\|z_+\| \leq 1$. We have

$$\begin{aligned} \|x\|_\star + z_+^T(y - x) &= \|x\|_\star + z_+^T y - x^T z_+ \\ &= \|x\|_\star + z_+^T y - \|x\|_\star \\ &= z_+^T y \\ &\leq \|z_+^T\| \|y\|_\star \text{ (use Hölder inequality)} \\ &\leq \|y\|_\star \ (\because \|z_+\| \leq 1) \end{aligned}$$

[12] https://math.stackexchange.com/questions/1950226/how-can-i-prove-this-theorem-regarding-the-sub-gradient-of-p-norm-using-holder-i

which gives $\|y\|_\star \geq \|x\|_\star + z_+^T(y-x)$, i.e., $z_+ \in \partial\|x\|_\star$. This shows $\text{argmax}_{\|z\|\leq 1}\ x^Tz \subset \partial\|x\|_\star$. Hence, $\partial\|x\|_\star = \text{argmax}_{\|z\|\leq 1}\ x^Tz$. $\square$

5 Inequalities

References

[1] Tom M. Apostol. *Mathematical Analysis, Second Edition.* 1974.

[2] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization.* 2004.

[3] Dan Garber and Elad Hazan. Faster rates for the frank-wolfe method over strongly-convex sets. In *Proceedings of the 32th International Conference on Machine Learning*, pages 541–549, 2015.

[4] Roger A. Horn and Charles R. Johnson. *Matrix Analysis, Second Edition.* 2013.

[5] Xianda Zhang. *Matrix Analysis and Applications (Chinese version).* Tsinghua University Press, 2nd edition, 2013.