MA662 – Multivariable Calculus Notes

Jiahua Chen

January 7, 2019

Subsets of $\mathbb{R}$

Definition 1.1. Let $X \subseteq \mathbb{R}$. Then...

1. $u \in \mathbb{R}$ is called an upper bound of X if $x \leqslant u$, $\forall x \in X$.
2. $l \in \mathbb{R}$ is called a lower bound of X if $x \geqslant l$, $\forall x \in X$.

It is an axiomatic property of $\mathbb{R}$ that each subset of $\mathbb{R}$ bounded above has a least upper bound and, likewise, each subset that is bounded below has a greatest lower bound.

Definition 1.2. Let $X \subseteq \mathbb{R}$ be bounded. Then...

1. $y = sup(X)$ (supremum of X) if y is an upper bound and, y' is another upper bound, then $y' \geqslant y$.
2. $z = inf(X)$ (infinum of X) if z is an lower bound and, z' is another lower bound, then $z' \leqslant z$.

Also if...

1. $sup(X) \in X$, then we call it the maximum of X.
2. $inf(X) \in X$, then we call it the minimum of X.

Example:
$X = (0,1)$ $\quad sup(X) = 1$ $\quad inf(X) = 0$ $\quad$ no max, no min
$X = [0,1]$ $\quad sup(X) = max(X) = 1$ $\quad inf(X) = min(X) = 0$

Proposition 1.3. If $X \subseteq \mathbb{R}$, bounded above, then $y = sup(X)$ iff

(i) y is an upper bound

(ii) $\forall \epsilon > 0, \exists x \in X$ such that $x > y - \epsilon$

Proof: Let $y = sup(X)$.

(i) is true by definition

(ii) Suppose $\exists \epsilon > 0$ such that there is no $x \in X$ with $x > y - \epsilon$.
Then $x \leqslant y - \epsilon \forall x \in X$. But that makes $y - \epsilon < y$ a smaller upper bound of X, which contradicts $y = sup(X)$

Suppose next that (i) and (ii) hold for $y \in \mathbb{R}$. We show that $y = sup(X)$. Clearly, y is an upper bound by (i), so let y' be a smaller upper bound for the sake of contradiction: $X \leqslant y' < y$ for all $x \in X$. Now consider $= y - y'$. Then $y - \epsilon = y - (y - y') = y' \geqslant x \forall x \in X$. This contradicts (ii) because we have found an $\epsilon > 0$ such that $\nexists x \in X$ greater than $y - \epsilon$. □

Proposition 1.4. Let X be bounded below.

$$inf(X) = -sup(-X)$$

where $-X = \{-x \mid x \in X\}$

Proof: Let $y = sup(-X)$. Then $y \geqslant -x \Rightarrow -y \leqslant x$ for all $x \in X$, so $-y$ is a lower bound for X. Now assume for the sake of contradiction that $\exists -y' > -y$, another lower bound of X. Then $-y' \leqslant x \Rightarrow y' \geqslant -x$ for all $x \in X$. But $-y' > -y \Rightarrow y' < y$ so $y \neq sup(-X)$. Hence $\nexists - y'$, another lower bound of X. $\Rightarrow -y = inf(X) \Rightarrow -sup(-X) = inf(X)$ □

Proposition 1.5. If A, B are bounded subsets of $\mathbb{R}$. Then $A \cup B$ is bounded and

$$sup(A \cup B) = max\{sup(A), sup(B)\}$$

January 8, 2019

Subsets $\mathbb{R}^n$ – Point-set topology

Definition 2.1. Let $x \in \mathbb{R}^n$ and $\epsilon > 0$. Then $B_\epsilon(x) = \{y \in \mathbb{R}^n \mid |x - y| < \epsilon\}$. This is called an ϵ-neighborhood of x.

Definition 2.2. Let $X \subseteq \mathbb{R}^n$ and $x \in R^n$. Then x is called

- interior point of X if $\exists \epsilon > 0$ such that $B_\epsilon(x) \subseteq X$.
- boundary point of X if $\exists \epsilon > 0$, $B_\epsilon(x) \cap X \neq \emptyset$ and $B_\epsilon(x) \cap X^c \neq \emptyset$
- exterior point of X if it is an interior point of X^c

Notation: $\mathring{X}$ = interior of X = set of all interior point of X. δX = boundary of X = set of all boundary points of X

Definition 2.3. X is called open if it only consists of interior points. ($X = \mathring{X}$)
X is called closed if its complement is open.
$\Rightarrow$ X is open if it contains none of its boundary points.
$\Rightarrow$ X is closed if it contains all of its boundary points

Exercise 1.5.1, book p.101. For each of the following subsets, state whether it is open or closed (or both or neither), and say why.

a. $\{x \in \mathbb{R} \mid 0 < x \leqslant 1\}$ as a subset of $\mathbb{R}$
   Answer: $\boxed{\text{Neither.}}$ 1 is not an interior point of this set and 0 is not an interior point of the complement of the set.

b. $\left\{\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid \sqrt{x^2 + y^2} < 1\right\}$ as a subset of $\mathbb{R}^2$
   Answer: $\boxed{\text{Open.}}$ The unit circle (which is the boundary) is not contained within the set.

c. the interval $(0, 1]$ as a subset of $\mathbb{R}$
   Answer: $\boxed{\text{Neither.}}$ Similar to a.

d. $\left\{\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid \sqrt{x^2 + y^2} \leqslant 1\right\}$ as a subset of $\mathbb{R}^2$

e. $\{x \in \mathbb{R} \mid 0 \leqslant x \leqslant 1\}$ as a subset of $\mathbb{R}$.

f. $\{(x, y, z) \in \mathbb{R}^3 \mid \sqrt{x^2 + y^2 + z^2} = 1 \text{ and } x, y, z \neq 0\}$ as a subset of $\mathbb{R}^3$

g. the empty set as a subset of $\mathbb{R}$

Exercise 1.5.2, book p.101. For each of the following subsets, state whether it is open or closed (or both or neither), and say why.

a. (x, y)-plane in $\mathbb{R}^3$

b. $\mathbb{R} \subset \mathbb{C}$

c. the line $x = 5$ in the (x, y)-plane

d. $(0, 1) \subset \mathbb{C}$

e. $\mathbb{R}^n \subset \mathbb{R}^n$

f. the unit sphere in $\mathbb{R}^3$

Exercise 1.5.5. For each of the following subsets of $\mathbb{R}$ and $\mathbb{R}^2$, state whether it is open or closed (or both or neither), and prove it.

a. $\left\{\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid 1 < x^2 + y^2 < 2\right\}$

b. $\left\{\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid xy \neq 0\right\}$

c. $\left\{\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid y = 0\right\}$

d. $\{\mathbb{Q} \subset \mathbb{R}\}$ (the rational numbers)

January 10, 2019

Recall **Prop 1.5**: If A, B are bounded subsets of $\mathbb{R}$. Then $A \cup B$ is bounded and

$$sup(A \cup B) = max\{sup(A), sup(B)\}$$

Proof:

1 Show that $x \leqslant max\{sup(A), sup(B)\}$ for all $x \in A \cup B$
Case 1: $x \in A \Rightarrow x \leqslant sup(A) \leqslant max\{sup(A), sup(B)\}$
Case 2: $x \in B \Rightarrow x \leqslant sup(B) \leqslant max\{sup(A), sup(B)\}$

2 Take $\epsilon > 0$ and consider $max\{sup(a), sup(B)\} - \epsilon$
Case 1: $max\{sup(A), sup(B)\} = supA \Rightarrow \exists x \in A$ such that $x > sup(A) - \epsilon \Rightarrow x \in A \cup B$ such that $x > max\{sup(A), sup(B)\} - \epsilon$
Case 2: $max\{sup(A), sup(B)\} = supB \Rightarrow$ *left to the reader, follows similarly as above.*

□

Also recall...

Exercise 1.5.5. For each of the following subsets of $\mathbb{R}$ and $\mathbb{R}^2$, state whether it is open or closed (or both or neither), and prove it.

a. $\left\{\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid 1 < x^2 + y^2 < 2\right\}$
Answer: $\boxed{\text{Open.}}$
Proof: Let $p \in A$(nnulus). $1 < |p - 0| < \sqrt{2}$. To show: $\exists \epsilon > 0$ s.t. all points in $B_\epsilon(p)$ are between 1 and $\sqrt{2}$ from 0. There is such ϵ, specifically

$$\epsilon = \frac{1}{2} \cdot min(\sqrt{2} - |p|, |p| - 1)$$

Now we show that for $x \in B_\epsilon(p)$, $1 < |x|^2 < 2$:
WLOG: Consider $p \in (1, \sqrt{2})$ on the x-axis. Then the neighborhood of p is:

$$B_\epsilon(p) = \left\{\begin{pmatrix} p + r\sin\theta \\ r\sin\theta \end{pmatrix} \middle| \, r \in [0, \epsilon)\right\}$$

$$\begin{aligned} \left|\begin{pmatrix} p + r\sin\theta \\ r\sin\theta \end{pmatrix}\right|^2 &= p^2 + 2pr\cos\theta + r^2\cos^2\theta + r^2\sin^2\theta \\ &= p^2 + 2pr\cos\theta + r^2 \end{aligned}$$

$$(p - r)^2 = p^2 - 2pr + r^2 \leqslant p^2 + 2pr\cos\theta + r^2 \leqslant p^2 + 2pr + r^2 = (p + r)^2$$

$$\text{Since } r < (\sqrt{2} - p), (p + r)^2 < (p + \sqrt{2} - p)^2 = 2$$

$$\text{Also since } r < (p - 1), (p - r)^2 > (p - (p - 1))^2 = 1$$

$\square$

We could also use the triangle inequality: $|a+b| \leqslant |a| + |b|$:

$$|p+r| \leqslant |p| + |r| < |p| + (\sqrt{2} - |p|) = \sqrt{2}$$

$$|p-r| \geqslant |p| - |r| > |p| - (|p| - 1) = 1$$

b. $\left\{\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid xy \neq 0\right\}$
Answer: $\boxed{\text{Open.}}$
Proof: Consider $B_\epsilon(p)$ with $\epsilon = \frac{1}{2} min\{|x|, |y|\}$. $\square$

c. $\left\{\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid y = 0\right\}$
Answer: $\boxed{\text{Closed.}}$
Proof: Consider the complement, $\left\{\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid y \neq 0\right\}$. Following a similar logic as *b*, consider $\epsilon = \frac{x}{2}$. $\square$

d. $\{\mathbb{Q} \subset \mathbb{R}\}$ (the rational numbers)
Answer: $\boxed{\text{Neither.}}$

Exercise 1.5.3. Prove the following statements for open subsets of $\mathbb{R}^n$:

a. **Any union of open sets is open.**
Proof: Let X_i, $i \in I$, be open. Consider $Y = \bigcup_{i \in I} X_i$.
To show: each $y \in Y$ is an interior point of Y.
Let $y \in Y$ belong to arbitrary X_i, for some $i \in I$. As X_i is open, y is also an interior point of X_i. So $\exists \epsilon > 0$ s.t. $B_\epsilon(y) \subset X_i \subseteq Y \Rightarrow y$ is an interior point of Y. $\square$

b. **A finite intersection of open sets is open.**
Proof: Consider $Z = \bigcap_{i=I}^{n} X_i$.
To show: each $z \in Z$ is an interior point of Z. Since $z \in Z$, $z \in X_i$ for $i = 1, \ldots, n$. Since X_i is open, $\exists \epsilon_i > 0 \mid B_{\epsilon_i}(z) \subset X_i$. As there are finitely many i, we choose the smallest $\epsilon = min\{\epsilon_i \mid i = 1, \ldots, n\}$. Then we have

$$B_\epsilon(z) \subset B_{\epsilon_i}(z) \subset X_i \text{ for all } x = 1, \ldots, n$$

Thus $\mathcal{B}_\epsilon(z) \subset Z$, making z an interior point of Z. $\square$

c. **An infinite interesection of open sets is not necessarily open.** *Proof:*

$$\bigcap_{n=1}^{\infty} \left\{ x \,\middle|\, x \in \left(-\frac{1}{n}, \frac{1}{n}\right)\right\} = \{0\}$$

$\square$

Definition 4.1. (Convergent sequence; limit of sequence). A sequence $i \mapsto \boldsymbol{a}_i$ if points in $\mathbb{R}^n$ *converges* to $\boldsymbol{a} \in \mathbb{R}^n$ if

$$\forall \epsilon > 0, \exists M \text{ s.t. } m > M \Rightarrow |\boldsymbol{a}_m - \boldsymbol{a}| < \epsilon$$

We then call $\boldsymbol{a}$ the *limit* of the sequence.

Proposition 4.2. (Convergence in terms of coordinates). A sequence $m \mapsto \boldsymbol{a}_m$ with $\boldsymbol{a}_m \in \mathbb{R}^n$ converges to $\boldsymbol{a}$ if and only if each coordinate converges; i.e., if for all j with $1 \leqslant j \leqslant n$, the jth coordinate of $\boldsymbol{a}_m$ converges to $\boldsymbol{a}_j$.

Proof:

□

Proposition 4.3. (Limit of sequence is unique). If the sequence $i \mapsto \boldsymbol{a}_i$ of points in $\mathbb{R}^n$ converges to $\boldsymbol{a}$ and to $\boldsymbol{b}$, then $\boldsymbol{a} = \boldsymbol{b}$.

Proof: Let the sequence $i \mapsto \boldsymbol{a}_i$ converge to both $\boldsymbol{a}$ and $\boldsymbol{b}$. Then

$$\forall \epsilon > 0, \exists M_a \wedge M_b \text{ s.t. } m > M_a, m > M_b \Rightarrow |\boldsymbol{a} - \boldsymbol{a}_m| < \frac{\epsilon}{2} \wedge |\boldsymbol{a}_m - \boldsymbol{b}| < \frac{\epsilon}{2}$$

$$|\boldsymbol{a} - \boldsymbol{b}| = |(\boldsymbol{a} - \boldsymbol{a}_m) + (\boldsymbol{b}_m - \boldsymbol{b})| \leqslant |\boldsymbol{a} - \boldsymbol{a}_m| + |\boldsymbol{a}_m - \boldsymbol{b}| = \epsilon$$
$$\Rightarrow |\boldsymbol{a} - \boldsymbol{b}| = 0 \Rightarrow \boldsymbol{a} = \boldsymbol{b}$$

□

Theorem 4.4. (The arithmetic of limits of sequences). All arithmetics that apply to limits apply here.

Proposition 4.5. (Sequence in closed set).

1. Let $i \mapsto \boldsymbol{x}_i$ be a sequence in a closed set $C \subset \mathbb{R}^n$ converging to $\boldsymbol{x}_0 \in \mathbb{R}^n$. Then $\boldsymbol{x}_0 \in C$.

2. Conversely, if every convergent sequence in a set $C \in \mathbb{R}^n$ converges to a point in C, then C is closed. a

Definition 5.1. (Limit of a function). Let X be a subset of $\mathbb{R}^n$ and $\mathbf{x}_0$ a point in $\overline{X}$ ($\overline{X} = X \cup \delta X$). A function $\mathbf{f} : X \to \mathbb{R}^m$ has the limit $\mathbf{a}$ at $\mathbf{x}_0$:

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \mathbf{a}$$

if for all $\epsilon > 0$ there exists $\delta > 0$ such that for all $\mathbf{x} \in X$,

$$|\mathbf{x} - \mathbf{x}_0| < \delta \Rightarrow |\mathbf{f}(\mathbf{x}) - a| < \epsilon$$

Proposition 5.2. (Convergence by coordinates). Suppose

$$U \subset \mathbb{R}^n, \quad \mathbf{f} = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} : U \to \mathbb{R}^m$$

Theorem 5.3. (Limits of functions). The same rules for traditional limits apply. Additional rules are as follows:

1. Dot product
2. ***

Exercise 1.5.14. State whether the following limits exist, and prove it.

a. $\displaystyle\lim_{(x,y)\to(1,2)} \frac{x^2}{x+y}$

b. $\displaystyle\lim_{(x,y)\to(0,0)} \frac{\sqrt{|x|}y}{x^2+y^2}$

c. ***

Definition 6.1. $X \subseteq \mathbb{R}^n$, define the closure of X: $\overline{X} = X \cup \delta X$

Theorem 6.2. $\overline{X}$ is the smallest closed set that contains X.

Proof: If X is closed, we are done.
Otherwise, assume $\exists Y \subset \mathbb{R}^n$, Y closed, with

$$X \subsetneqq Y \subseteq \overline{X}$$

We show that $Y = \overline{X}$: Assume otherwise for the sake of contradiction that that $\exists z \in \overline{X} - Y \subseteq Y^C$ which is open. Then $\exists \epsilon > 0$ s.t. $B_\epsilon(z) \subseteq Y^C$. Hence $B_\epsilon(z) \subseteq R^n - X$, which contradicts $x \in \overline{X}$. Therefore $\overline{X} - Y = \emptyset$, so $Y = \overline{X}$. □

Definition 6.3. (Continuous function). Let $X \subset \mathbb{R}^n$. A mapping $f : X \to \mathbb{R}^m$ is continuous at $\boldsymbol{x}_0 \in X$ if

$$\lim_{x \to \boldsymbol{x}_0} \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x}_0);$$

$\boldsymbol{f}$ is continuous on X if it is continuous at every point of X. Equivalently, $\boldsymbol{f} : X \to \mathbb{R}^m$ is continuous at $\boldsymbol{x}_0 \in X$ if and only if for every $\epsilon > 0$, there exists $\delta > 0$ such that when $|\boldsymbol{x} - \boldsymbol{x}_0| < \delta$, then $|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}_0)| < \epsilon$.

Theorem 6.4. (Combining continuous mappings). Continuous functions are closed under addition, scalar umltiplication, quotients, and compositions.

Lemma 6.5. Hence polynomials and rational functions (given that the denominator does not vanish) are continous.

Exercise 1.5.21. For the following functions, can you choose a value for f at $\begin{pmatrix}0\\0\end{pmatrix}$ to make the function continuous at the origin?

a. $f\begin{pmatrix}x\\y\end{pmatrix} = \dfrac{1}{x^2+y^2+1}$

Answer: $\boxed{\text{Exists. } f\begin{pmatrix}0\\0\end{pmatrix} = 1.}$

The limit exists at $\begin{pmatrix}0\\0\end{pmatrix}$ by substitution.

b. $f\begin{pmatrix}x\\y\end{pmatrix} = \dfrac{\sqrt{x^2+y^2}}{|x|+|y|^{1/3}}$

Answer: $\boxed{\text{Does not exist.}}$

Proof: Approaching $\begin{pmatrix}0\\0\end{pmatrix}$ from $\begin{pmatrix}x\\0\end{pmatrix}$ gives $\lim\limits_{x\to 0} \dfrac{\sqrt{x^2}}{|x|} = \lim\limits_{x\to 0} \dfrac{|x|}{|x|} = 1$, whilst approaching $\begin{pmatrix}0\\0\end{pmatrix}$ from $\begin{pmatrix}0\\y\end{pmatrix}$ gives $\lim\limits_{y\to 0} \dfrac{\sqrt{y^2}}{|y|^{1/3}} = \lim\limits_{y\to 0} \dfrac{y}{y^{1/3}} = \lim\limits_{y\to 0} y^{2/3} = 0$. $\Rightarrow\Leftarrow$. □

c. $f\begin{pmatrix}x\\y\end{pmatrix} = (x^2+y^2)\ln(x^2+2y^2)$

Answer: $\boxed{f\begin{pmatrix}0\\0\end{pmatrix} = 0.}$

Proof: Consider

$$g\begin{pmatrix}x\\y\end{pmatrix} = (x^2+y^2)\ln(x^2+y^2)$$

$$g\begin{pmatrix}r\\\theta\end{pmatrix} = r^2\ln(r^2) = 2r^2\ln(r)$$

$$\lim_{r\to 0} r^2\ln(r^2) = \lim_{r\to 0}\frac{2\ln(r)}{r^{-2}} = \lim_{r\to 0}\frac{r^{-1}}{-2r^{-3}} = \lim_{r\to 0}\frac{1}{-2}r^2 = 0$$

Now consider bounding $f\begin{pmatrix}0\\0\end{pmatrix}$.

$$g\begin{pmatrix}x\\y\end{pmatrix} \leqslant f\begin{pmatrix}x\\y\end{pmatrix} \leqslant 0 \quad \text{for } \begin{pmatrix}x\\y\end{pmatrix} \text{ sufficiently near } \begin{pmatrix}0\\0\end{pmatrix}$$

And the squeeze theorem gives that $\lim_{(x,y)\to(0,0)} f\begin{pmatrix}x\\y\end{pmatrix} = 0.$ □

d. $f\begin{pmatrix}x\\y\end{pmatrix} = (x^2+y^2)\ln|x+y|$

Answer: $\boxed{\text{Limit does not exist.}}$

Proof: Consider approaching $f\begin{pmatrix}x\\y\end{pmatrix}$ from $y=-x$. We then have

$$\lim_{(x,y)\to(0,0)} f\begin{pmatrix}x\\y\end{pmatrix} = \lim_{y\to 0} 2y^2\cdot\ln|0| = \infty!$$

□

Exercise 1.5.16b. Either show that the limit exists at 0 and find it, or show that it does not exist:

$$f\begin{pmatrix}x\\y\end{pmatrix} = \frac{\sin(x+y)}{\sqrt{x^2+y^2}}$$

Answer: $\boxed{\text{Does not exist.}}$

Proof: Consider approaching $\begin{pmatrix}0\\0\end{pmatrix}$ from $\begin{pmatrix}x\\0\end{pmatrix}$. We then have

$$\lim_{(x,y)\to(0,0)} f\begin{pmatrix}x\\y\end{pmatrix} = \lim_{x\to 0}\frac{\sin(x)}{|x|}$$

$$\lim_{x\to 0^+}\frac{\sin(x)}{|x|} = +1 \quad \text{but} \quad \lim_{x\to 0^-}\frac{\sin(x)}{|x|} = -1 \neq +1$$

January 19, 2019

Recall from previously, we were trying to solve:

$$g\begin{pmatrix}x\\y\end{pmatrix} = (|x|+|y|)\cdot\ln\left(x^2+y^4\right) < 0 \text{ near } \begin{pmatrix}0\\0\end{pmatrix}$$

The solution then follows that

$$(|x|+|y|)\cdot\ln\left(x^4+y^4\right) < g\begin{pmatrix}x\\y\end{pmatrix} < 0$$

We can use lp-norms to estimate one of the values in the above function.

$$\left\|\begin{pmatrix}x\\y\end{pmatrix}\right\|_p = \left(|x|^p+|y|^p\right)^{\frac{1}{p}}, p \geqslant 1 \quad lp\text{-norms}$$

Definition 7.1. (Bounded set). A subset $X \in \mathbb{R}^n$ is bounded if it is contained in a ball in $\mathbb{R}^n$ centered at the origin:

$$X \subset B_R(0) \quad \text{for some } R < \infty.$$

Definition 7.2. (Compact set). A nonempty subset $C \subset \mathbb{R}^n$ is compact if it is closed and bounded.

Theorem 7.3. (Convergent subsequence in a compact set). If a compact set $C \subset \mathbb{R}^n$ contains a sequence $i \mapsto \boldsymbol{x}_i$, then that sequence has a convergent subsequence $j \mapsto \boldsymbol{x}_{i(j)}$ whose limit is in C.

January 22, 2019

Definition 8.1. (Supremum). ***

Definition 8.2. (Infimum). ***

Definition 8.3. (Minimum value; minimum). ***

Theorem 8.4. (Existence of minima and maxima). Let $C \subset \mathbb{R}^n$ be a compact subset, and let $f : C \to \mathbb{R}$ be a continuous function. Then there exists a point $\boldsymbol{a} \in C$ such that $f(\boldsymbol{a}) \geqslant f(\boldsymbol{x})$ for all $\boldsymbol{x} \in C$, and a point $\boldsymbol{b} \in C$ such that $f(\boldsymbol{b}) \leqslant f(\boldsymbol{x})$ for all $\boldsymbol{x} \in C$.

Proof: Detailed in textbook. □

Theorem 8.5. (Mean value theorem). If $f : [a,b] \to \mathbb{R}$ is continuous, and f is differentiable on (a,b), then there exists $c \in (a,b)$ such that

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

Theorem 8.6. (Fundamental theorem of algebra). Let

$$p(z) = z^k + a_{k-1} z^{k-1} + \cdots + a_0$$

***

Derivatives

Idea: Replace a complicated nonlinear equation by a linear one with the understanding that the results only hold approximately in a small neighborhood around a point $p \in \mathbb{R}^n$ but that the error vanishes faster than the distance to p.

$$\lim_{h \to 0} \frac{f(x+h) - f(x) - f'(x) \cdot h}{h} = 0$$

For $f : \mathbb{R}^n \to \mathbb{R}^m$, we are looking for a function $\boldsymbol{Df}(\boldsymbol{x}_0) \in L(\mathbb{R}^n, \mathbb{R}^m)$ such that

$$\lim_{\vec{h} \to 0} \frac{\left\{f(\boldsymbol{x} + \vec{h}) - f(\boldsymbol{x})\right\} - \left\{[\boldsymbol{Df}(\boldsymbol{x}_0)]\vec{h}\right\}}{|\vec{h}|} = 0$$

As a linear transformation $\mathbb{R}^n \to \mathbb{R}^m$, $\boldsymbol{Df}(\boldsymbol{x}_0)$ has a matrix which is called the Javobian of f at $\boldsymbol{x}_0$: $[\boldsymbol{Df}(\boldsymbol{x}_0)] = [\boldsymbol{Jf}(\boldsymbol{x}_0)]$

Test on Thursday: Topology, sets, limits, supremum, infimum, continuity,

Derivatives (continued)

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

Goal: local linearization of f with error going to zero sufficiently fast.

$$\lim_{\vec{h}\to 0} \frac{1}{|\vec{h}|}(f(\boldsymbol{x}+\vec{h}) - f(\vec{\boldsymbol{x}})) - [\boldsymbol{Df}(\boldsymbol{x})]\vec{h} = \vec{0} \qquad [\boldsymbol{Df}(\boldsymbol{x})] \in L(\mathbb{R}^n, \mathbb{R}^m)$$

If we know that $[\boldsymbol{Df}(\boldsymbol{x})]$ exists, we can calculate its matrix $[\boldsymbol{Jf}(\boldsymbol{x})]$ (Javobian matrix) by evaluating $[\boldsymbol{Df}(\boldsymbol{x})]$ on the standard basis vectors.

We know that

$$\begin{aligned}
0 &= \lim_{|h|\to 0} \frac{1}{|h\vec{e_i}|}(f(\boldsymbol{x}+h\vec{e_i}) - f(\boldsymbol{x}) - [\boldsymbol{Df}(\boldsymbol{x})](h\vec{e_i})) \\
&= \lim_{|h|\to 0} \frac{1}{|h|}(f(\boldsymbol{x}+h\vec{e_i}) - f(\boldsymbol{x}) - h[\boldsymbol{Df}(\boldsymbol{x})](\vec{e_i})) \\
&= \lim_{h\to 0} \frac{f(\boldsymbol{x}+h\vec{e_i}) - f(\boldsymbol{x})}{h} - [\boldsymbol{Df}(\boldsymbol{x})](\vec{e_i})
\end{aligned}$$

$$\Rightarrow [\boldsymbol{Df}(\boldsymbol{x})]\vec{e_i} = \lim_{h\to 0} \frac{f(\boldsymbol{x}+h\vec{e_i}) - f(\boldsymbol{x})}{h} \qquad (\#)$$

The right-hand side of # is called the partial derivative of f at x: in components, it looks as the follows:

$$\lim_{h\to 0} \frac{f\begin{pmatrix} x_1 & \\ x_2 & \\ \vdots & \\ x_i & +h \\ \vdots & \\ x_n & \end{pmatrix} - f\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{pmatrix}}{h}$$

Therefore, we can calculate it by considerinf x_i the only variable, and holding all other components constant.

There are a variety of notations for this derivative:

- $D_i f(x)$
- $D_x f(\boldsymbol{x}), D_y f(\boldsymbol{x}), D_z f(\boldsymbol{x})$
- $\frac{\delta f}{\delta x}, \frac{\delta f}{\delta x_2} \ldots$

- $f_x, f_y \dots$

Example:

$$\begin{aligned}
f\begin{pmatrix} x \\ y \end{pmatrix} &= \sin\left(x^2 + y^3\right) \\
D_x f\begin{pmatrix} x \\ y \end{pmatrix} &= \cos\left(x^2 + y^3\right) \cdot 2x \\
D_y f\begin{pmatrix} x \\ y \end{pmatrix} &= \cos\left(x^2 + y^3\right) \cdot 3y^2 \\
\left[Df\begin{pmatrix} x \\ y \end{pmatrix}\right] &= \cos\left(x^2 + y^3\right)\begin{bmatrix} 2x & 3y^2 \end{bmatrix}
\end{aligned}$$

Warning: The Jacobian matrix is only the matrix of the derivative if the function is actually differentiable!

(**Preview:** We will see shortly that f is differentiable if all its partials exist and are continuous.)

This Df gives us the rate of change in the axes, if we want to find the directional rate of change in any direction, we have to use a direction derivative.

Definition 9.1. The directional derivative of f at x in direction gives the rate of change of f as we step into direction $\vec{v}$. It is defined as

$$\lim_{h \to 0} \frac{f\left(x + h\vec{v}\right) - f(x)}{h}$$

We will see shortly that this evaluates to $[Df(x)]\vec{v}$.

Let's brush up on some simple derivatives first:

a. $f(x) = \sin^3\left(x^2 + \cos x\right) = \sin$

b. ***

Exercise 1.7.4. Using the definition, check whether the following functions are differentiable at 0.

a. $f(x) = |x|^{3/2}$

b. $f(x) = \begin{cases} x \cdot \ln|x| & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$

   Answer: $\boxed{\text{Does not exist.}}$

c. $f(x) = \begin{cases} x / \ln|x| & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$

   Answer: $\boxed{\text{Exists.}}$

Proposition 10.1. (***)

Proof: Recall that the expression

$$\boldsymbol{r}(\vec{\boldsymbol{h}}) = (\boldsymbol{f}(\boldsymbol{a}+\vec{\boldsymbol{h}}) - \boldsymbol{f}(\boldsymbol{a})) - [\boldsymbol{D}\boldsymbol{f}(\boldsymbol{a})]\vec{\boldsymbol{h}}$$

*** $\square$

Rules for calculting derivatives

(A lot of them are surprisingly similar)

1. If $\boldsymbol{f} : U \to \mathbb{R}^m$ is a constant function, then $\boldsymbol{f}$ is differentiable, and its derivative is $[0]$.
2. If $\boldsymbol{f} : \mathbb{R}^n \to \mathbb{R}^m$ is linear, then it is differentiable everywhere, and its derivative at all points $\boldsymbol{a}$ is $\boldsymbol{f}$, *i.e.*, ***
3. (differentiable just take singular derivative)
4. sum
5. product
6. quotient rule
7. composition

Theorem 10.2. (Chain rule). Let $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$ be open sets, let $\boldsymbol{g} : U \to V$ and $\boldsymbol{f} : V \to \mathbb{R}^p$ be mappings, and let $\boldsymbol{a}$ be a point of U. If $\boldsymbol{g}$ is differentiable at $\boldsymbol{a}$ and $\boldsymbol{f}$ is differentiable at $\boldsymbol{g}(\boldsymbol{a})$, then the composition $\boldsymbol{f} \circ \boldsymbol{g}$ is differentiable at $\boldsymbol{a}$, and its derivative is given by

$$\boldsymbol{D}[(\boldsymbol{f} \circ \boldsymbol{g})(\boldsymbol{a})] = [\boldsymbol{D}\boldsymbol{f}(\boldsymbol{g}(\boldsymbol{a}))] \circ [\boldsymbol{D}\boldsymbol{g}(\boldsymbol{a})]$$

Example: (The derivative of a composition). Define $\boldsymbol{g} : \mathbb{R} \to \mathbb{R}^3$ and $f : \mathbb{R}^3 \to \mathbb{R}$

$$f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x^2 + y^2 + z^2; \qquad \boldsymbol{g}(t) = \begin{pmatrix} t \\ t^2 \\ t^3 \end{pmatrix}$$

Theorem 11.1. (Mean value theorem for functions of several variables). Let $U \subset \mathbb{R}^n$ be open, let $f : U \to \mathbb{R}$ be differentiable, and let the segment $[\boldsymbol{a}, \boldsymbol{b}]$ joining $\boldsymbol{a}$ to $\boldsymbol{b}$ be comtained in U. Then there exists $\boldsymbol{c}_0 \in [\boldsymbol{a}, \boldsymbol{b}]$ such that

$$f(\boldsymbol{b} - (\boldsymbol{a}) = [\boldsymbol{D}f(\boldsymbol{c}_0)](\boldsymbol{b} \vec{-} \boldsymbol{a})$$

Definition 11.2. (C^p function). A <u>C^p *function*</u> on $U \subset \mathbb{R}^n$ is a function that is p times continuously differentiable: all of its partial derivatives up to order p exist and are continuous on U.

Exercise 1.9.1. Show that

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{cases} \dfrac{x^4 + y^4}{x^2 + y^2} & \text{if } \begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ 0 & \text{if } \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{cases}$$

***

Exercise 1.9.2. Show that for

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{cases} \dfrac{3x^2 y - y^3}{x^2 + y^2} & \text{if } \begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ 0 & \text{if } \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{cases}$$

all directional derivaties exist, but that f is not differentiable at the origin.

Recall

Exercise 1.9.2a. Show that for

$$f\begin{pmatrix}x\\y\end{pmatrix}=\begin{cases}\dfrac{3x^2y-y^3}{x^2+y^2} & \text{if } \begin{pmatrix}x\\y\end{pmatrix}\neq\begin{pmatrix}0\\0\end{pmatrix}\\ 0 & \text{if } \begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}\end{cases}$$

all directional derivaties exist, but that f is not differentiable at the origin. *Proof:*

$$\begin{aligned}D_xf\begin{pmatrix}0\\0\end{pmatrix}&=\lim_{h\to 0}\frac{f\begin{pmatrix}0+h\\0\end{pmatrix}-0}{h}\\&=\lim_{h\to 0}\frac{3h^2\cdot 0-0^3}{h(h^2+0^2)}=0\\D_yf\begin{pmatrix}0\\0\end{pmatrix}&=\lim_{h\to 0}\frac{0-h^3}{h(0^2+h^2)}=1\end{aligned}$$

$$\boldsymbol{Jf}\begin{pmatrix}0\\0\end{pmatrix}=\begin{bmatrix}0 & -1\end{bmatrix}$$

$$\vec{v}=\begin{bmatrix}v_1\\v_2\end{bmatrix}\qquad\lim_{h\to 0}\frac{f(0+h\vec{v})-f(0)}{h}=\lim_{h\to 0}\frac{3h^2v_1^2hv_2-h^3v_2^2}{h(h^2v_1+h^2v_2^2)}=\frac{3v_1^2v_2-v_2^3}{v_1^2+v_2^2}$$

$$\Rightarrow \text{ all directional derivatives exist}$$

If f is differentiable, the directional derivative is $\left[\boldsymbol{Jf}\begin{pmatrix}0\\0\end{pmatrix}\right]\vec{v}=-v_2$, so setting $\vec{v}=\begin{bmatrix}1\\1\end{bmatrix}$ gives a contradiction, hence $f\begin{pmatrix}x\\y\end{pmatrix}$ is not differentiable at the origin. □

Exercise 1.9.2b. Show that

$$g\begin{pmatrix}x\\y\end{pmatrix}=\begin{cases}\dfrac{x^2y}{x^4+y^2} & \text{if } \begin{pmatrix}x\\y\end{pmatrix}\neq\begin{pmatrix}0\\0\end{pmatrix}\\ 0 & \text{if } \begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}\end{cases}$$

has directional derivatives at every point but is not continuous.

Proof: ***

$$\vec{v}=\begin{bmatrix}v_1\\v_2\end{bmatrix}\qquad\lim_{h\to 0}\frac{g(0+h\vec{v})-g(0)}{h}=\lim_{h\to 0}\frac{h^2v_1^2\cdot hv_2}{h(h^4v_1^4+h^2v_2^2)}=\lim_{h\to 0}\frac{v_1^2v_2}{h^2v_1^4+v_2^2}=\frac{v_1^2}{v_2}$$

□

Exercise 1.9.2c. Show that

$$h\begin{pmatrix}x\\y\end{pmatrix}=\begin{cases}\dfrac{x^2y}{x^6+y^2} & \text{if } \begin{pmatrix}x\\y\end{pmatrix}\neq\begin{pmatrix}0\\0\end{pmatrix}\\ 0 & \text{if } \begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}\end{cases}$$

***.

Exercise 1.8.11. Show that if $f\begin{pmatrix}x\\y\end{pmatrix}=\varphi\left(\dfrac{x+y}{x-y}\right)$ for some differentiable function $\varphi:\mathbb{R}\mapsto\mathbb{R}$, then

$$xD_xf+yD_yf=0$$

Proof:

$$xD_xf=xD_x\varphi\left(\frac{x+y}{x-y}\right)=x\cdot\varphi'\left(\frac{x+y}{x-y}\right)\cdot\frac{(x-y)-(x+y)}{(x-y)^2}=\varphi\left(\frac{x+y}{x-y}\right)\frac{-2xy}{(x-y)^2}$$

$$yD_yf=yD_y\varphi\left(\frac{x+y}{x-y}\right)=x\cdot\varphi'\left(\frac{x+y}{x-y}\right)\cdot\frac{(x-y)-(-1)(x+y)}{(x-y)^2}=\varphi\left(\frac{x+y}{x-y}\right)\frac{2xy}{(x-y)^2}$$

□

Just for the future, we might want to switch to a new coordinate system. We'll have to use the chain rule in these cases. For example:

$$x=r\cos\theta\quad y=r\sin\theta$$

$$\boldsymbol{D}_\theta f\begin{pmatrix}X(r,\theta)\\Y(r,\theta)\end{pmatrix}$$

$$\boldsymbol{D}_r f\begin{pmatrix}X(r,\theta)\\Y(r,\theta)\end{pmatrix}$$

$$\boldsymbol{D}f\begin{pmatrix}X(r,\theta)\\Y(r,\theta)\end{pmatrix}$$

Consider f as the "outside" function, and $h:\begin{pmatrix}r\\\theta\end{pmatrix}\mapsto\begin{pmatrix}x\\y\end{pmatrix}$. Recall

$$\boldsymbol{D}[(\boldsymbol{f}\circ\boldsymbol{g})(\boldsymbol{a})]=[\boldsymbol{D}\boldsymbol{f}(\boldsymbol{g}(\boldsymbol{a}))]\circ[\boldsymbol{D}\boldsymbol{g}(\boldsymbol{a})]$$

We then get

$$\boldsymbol{D}h=\begin{bmatrix}\cos\theta & -r\sin\theta\\ \sin\theta & r\cos\theta\end{bmatrix}$$

$$\boldsymbol{D}f=\left[\boldsymbol{D}_xf(r,\theta)\quad\boldsymbol{D}_yf(r,\theta)\right]$$

$$[\boldsymbol{D}f][\boldsymbol{D}h]=\left[\boldsymbol{D}_xf(r,\theta)\quad\boldsymbol{D}_yf(r,\theta)\right]\begin{bmatrix}\cos\theta & -r\sin\theta\\ \sin\theta & r\cos\theta\end{bmatrix}$$

Exercise 1.8.10 (p.144).

Definition 13.1. (Newton's method). Let $\vec{\boldsymbol{f}}$ be a differentiable map from U to $\mathbb{R}^n$, where U is an open subset of $\mathbb{R}^n$. Newton's method consists of starting with some guesses $\boldsymbol{a}_0$ for a sollution of $\vec{\boldsymbol{f}}(\boldsymbol{x}) = \vec{0}$. Then linearize the equation at $\boldsymbol{a}_0$: replace the increment to the function, $\vec{\boldsymbol{f}}(\boldsymbol{x}) - \vec{\boldsymbol{f}}(\boldsymbol{a}_0)$, by a linear function of the increment, $[\boldsymbol{D}\vec{\boldsymbol{f}}(\boldsymbol{a}_0)](\boldsymbol{x} - \boldsymbol{a}_0)$. Now solve the corresponding linear equation:

$$\vec{f}(\boldsymbol{a}_0) + [\boldsymbol{D}\vec{\boldsymbol{f}}(\boldsymbol{a}_0)](\boldsymbol{x} - \boldsymbol{a}_0) = \vec{0}$$

***

$$\begin{aligned} a_0 &= \text{initial guess} \\ a_{n+1} &= a_n - [\boldsymbol{D}\vec{\boldsymbol{f}}(a_n)]^{-1}\boldsymbol{f}(a_n) \\ x_{n+1} &= x_n - \frac{\boldsymbol{f}(x)}{\boldsymbol{f}'(x_n)} \end{aligned}$$

Example:

$$f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos x + y - 1.1 \\ \text{x} + \cos(x+y) - 0.9 \end{pmatrix}$$