MATH 247 LECTURE NOTES

Zhongwei Zhao

My Lecture Notes for MATH 247 2017 Spring

August 2017

Lecture 1, May. 1

Will do multi-variable calculus, extend what you learned from MATH 147, 148 from $\mathbb{R} = \mathbb{R}^1$ to $\mathbb{R}^n$.

The course has 3 parts:

1. Sequence, limits, continuity (also: sets in $\mathbb{R}^n$ - open, closed, compact set)
2. Derivatives in $\mathbb{R}^m$ (requires grasp of linear algebra)
3. Integrals (Connection between derivatives and integrals)

Notation (Vectors).

$$\vec{x} = (x^{(1)}, x^{(2)}, \cdots, x^{(n)}) \in \mathbb{R}^n$$

Notation (Inner Product).

$$\langle \vec{x}\,,\, \vec{y}\,\rangle = x^{(1)}y^{(1)} + x^{(2)}y^{(2)} + \cdots + x^{(n)}y^{(n)},\ \forall \vec{x}, \vec{y} \in \mathbb{R}^n$$

Notation (Norm).

$$||\,\vec{x}\,|| = \sqrt{\langle \vec{x}\,,\, \vec{x}\,\rangle} = \sqrt{\sum (x^{(i)})^2}$$

<u>Some basic inequalities in $\mathbb{R}^n$</u>

1.1 Proposition (Cauchy-Schwarz Inequality)**.**

$$|\ \langle \vec{x}\,,\, \vec{y}\,\rangle\ | \leq ||\,\vec{x}\,|| \cdot ||\,\vec{y}\,||,\ \forall \vec{x}, \vec{y} \in \mathbb{R}^n$$

Proof. If $\vec{y} = \vec{0}$, then we get $0 = 0$.

So we will assume that $\vec{y} \neq \vec{0}$, hence that $||\,\vec{y}\,|| > 0$.

Define $h\colon \mathbb{R} \to \mathbb{R}$ by

$$h(t) = \langle \vec{x} - t\vec{y}\,,\, \vec{x} - t\vec{y}\,\rangle,\ t \in \mathbb{R}$$

Observe that $h(t) \geq 0,\ \forall t \in \mathbb{R}$ by positivity of the inner product.

But on the other hand, use bilinearity and get

$$\begin{aligned} h(t) &= \langle \vec{x}\,,\, \vec{x}\,\rangle - \langle \vec{x}\,,\, t\vec{y}\,\rangle - \langle t\vec{y}\,,\, \vec{x}\,\rangle + \langle -t\vec{y}\,,\, -t\vec{y}\,\rangle \\ &= ||\,\vec{x}\,||^2 - 2t\langle \vec{x}\,,\, \vec{y}\,\rangle + t^2||\,\vec{y}\,||^2 \\ &= at^2 - bt + c \end{aligned}$$

where $a = ||\,\vec{y}\,||^2$, $b = 2\langle \vec{x}\,,\, \vec{y}\,\rangle$, $c = ||\,\vec{x}\,||^2$.

So $h(t)$ is a quadratic function of t, as $||\,\vec{y}\,||^2$ is strictly positive, and $h(t) \geq 0$ for all $t \in \mathbb{R}$. It follows that the discriminant $\Delta = b^2 - 4ac \leq 0$.

Then

$$\langle \vec{x}\,,\, \vec{y}\,\rangle^2 \leq ||\,\vec{x}\,||^2 ||\,\vec{y}\,||^2$$

Hence

$$|\ \langle \vec{x}\,,\, \vec{y}\,\rangle\ | \leq ||\,\vec{x}\,||\,||\,\vec{y}\,||$$

∎

1.2 Definition (Distance)**.** Let $\vec{x} = (x^{(1)}, \cdots, x^{(n)}), \vec{y} = (y^{(1)}, \cdots, y^{(n)}) \in \mathbb{R}^n$. The distance between $\vec{x}$ and $\vec{y}$ is

$$d(\vec{x}, \vec{y}) = ||\,\vec{x} - \vec{y}\,|| = \sqrt{\sum (x^{(i)} - y^{(i)})^2}$$

1.3 Proposition (Triangle Inequality)**.**

Form 1

$$||\,\vec{x} + \vec{y}\,|| \leq ||\,\vec{x}\,|| + ||\,\vec{y}\,||,\ \forall \vec{x}, \vec{y} \in \mathbb{R}^n.$$

Form 2

$$d(\vec{x}, \vec{z}) \leq d(\vec{x}, \vec{y}) + d(\vec{y}, \vec{z}),\ \forall \vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$$

Proof.

$$\begin{aligned}
||\,\vec{x} + \vec{y}\,||^2 =& \langle \vec{x} + \vec{y}\,,\, \vec{x} + \vec{y}\,\rangle \\
=& \langle \vec{x}\,,\, \vec{x}\,\rangle + \langle \vec{x}\,,\, \vec{y}\,\rangle + \langle \vec{y}\,,\, \vec{x}\,\rangle + \langle \vec{y}\,,\, \vec{x}\,\rangle \\
=& ||\,\vec{x}\,||^2 + 2\langle \vec{x}\,,\, \vec{y}\,\rangle + ||\,\vec{y}\,||^2 \\
\leq& ||\,\vec{x}\,||^2 + 2\,|\,\langle \vec{x}\,,\, \vec{y}\,\rangle\,| + ||\,\vec{y}\,||^2 \\
\leq& ||\,\vec{x}\,||^2 + 2||\,\vec{x}\,||||\,\vec{y}\,|| + ||\,\vec{y}\,||^2 \\
=& (||\,\vec{x}\,|| + ||\,\vec{y}\,||)^2.
\end{aligned}$$

So we proved that

$$||\,\vec{x} + \vec{y}\,||^2 \leq (||\,x\,|| + ||\,y\,||)^2.$$

Take the square root on both side and we get form 1.

Form 2: Given $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$. Wrote

$$\begin{aligned}
d(\vec{x}, \vec{z}) =& ||\,\vec{x} - \vec{z}\,|| \\
=& ||\,(\vec{x} - \vec{y}) + (\vec{y} - \vec{z})\,|| \\
\leq& ||\,\vec{x} - \vec{y}\,|| + ||\,\vec{y} - \vec{z}\,|| \\
=& d(\vec{x}, \vec{y}) + d(\vec{y}, \vec{z}).
\end{aligned}$$

□

1.4 Remark. Both propositions extends to "polygon inequalities".

For every $\vec{x}, \vec{y}, \vec{z}, \vec{w} \in \mathbb{R}^n$, we have

$$d(\vec{x}, \vec{w}) \leq d(\vec{x}, \vec{y}) + d(\vec{y}, \vec{z}) + d(\vec{z}, \vec{w})$$

Proof. The proof is by induction on the number of points involved in the polygonal path. □

Notation. For $\vec{x} = (x^{(1)}, \cdots, x^{(n)}) \in \mathbb{R}^n$ denote

$$||\,\vec{x}\,||_1 = \left|\ x^{(1)}\ \right| + \cdots + \left|\ x^{(n)}\ \right|$$

$$||\,\vec{x}\,||_\infty = \max(x^{(1)} + \cdots + x^{(n)})$$

Aside. For every $p \in [1, \infty)$ denote

$$|| \vec{x} ||_p = \left(\left| x^{(1)} \right|^p + \cdots + \left| x^{(n)} \right|^p \right)^{1/p}.$$

Usual norm is $|| \vec{x} ||_2$

1.5 Exercise. Prove that for every $\vec{x} \in \mathbb{R}^n$,

$$|| \vec{x} ||_\infty \leq || \vec{x} || \leq || \vec{x} ||_1 \leq n || \vec{x} ||_\infty$$

Lecture 2, May. 3

Sequences in $\mathbb{R}^n$

We use notation $(\vec{x}_k)_{k=1}^{\infty}$ to mean that we have $\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_k, \cdots$ in $\mathbb{R}^n$ (It is allowed to have $\vec{x}_i = \vec{x}_j$ for $i \neq j$).

2.1 Definition. Let $(\vec{x}_k)_{k=1}^{\infty}$ be a sequence in $\mathbb{R}^n$. Let $\vec{a}$ in $\mathbb{R}^n$. Consider the sequence of real numbers $(||\,\vec{x} - \vec{a}\,||)_{k=1}^{\infty}$. We say that $(\vec{x}_k)_{k=1}^{\infty}$ converges to $\vec{a}$ to mean that $||\,\vec{x} - \vec{a}\,|| \to 0$ as $k \to \infty$.

2.2 Remark.

$$\forall \epsilon > 0,\ \exists k_0 \in \mathbb{N} \text{ such that } \forall k \geq k_0,\ \ ||\,\vec{x}_k - \vec{a}\,|| < \epsilon$$

2.3 Definition (Open Ball of center $\vec{a}$ and radius r)**.** For $\vec{a} \in \mathbb{R}^m$ and $r > 0$ denote

$$\mathrm{B}(\vec{a}, r) = \{\vec{x} \in \mathbb{R}^m \mid ||\,\vec{x} - \vec{a}\,|| < r\}.$$

2.4 Definition (Closed Ball of center $\vec{a}$ and radius r)**.** For $\vec{a} \in \mathbb{R}^m$ and $r > 0$ denote

$$\overline{\mathrm{B}}(\vec{a}, r) = \{\vec{x} \in \mathbb{R}^m \mid ||\,\vec{x} - \vec{a}\,|| \leq r\}.$$

2.5 Remark. Instead of $||\,\vec{x}_k - \vec{a}\,|| < \epsilon$ we can also write $d(\vec{x}_k, \vec{a}) < \epsilon$, or $\vec{x}_k \in \mathrm{B}(\vec{a}, \epsilon)$.

2.6 Definition (Component Sequences)**.** Let $(\vec{x}_k)_{k=1}^{\infty}$ be a sequence in $\mathbb{R}^n$. We write explicitly

$$\vec{x}_1 = (x_1^{(1)}, x_1^{(2)}, \cdots, x_1^{(n)})$$
$$\vec{x}_2 = (x_2^{(1)}, x_2^{(2)}, \cdots, x_2^{(n)})$$
$$\cdots$$
$$\vec{x}_k = (x_k^{(1)}, x_k^{(2)}, \cdots, x_k^{(n)})$$
$$\cdots$$

We get n sequences in $\mathbb{R}^n$: $(x_k^{(1)})_{k=1}^{\infty}, (x_k^{(2)})_{k=1}^{\infty}, \cdots, (x_k^{(n)})_{k=1}^{\infty}$. They are called the *component sequences* of $(\vec{x}_k)_{k=1}^{\infty}$.

(Note: If conversely someone gives you n sequences in $\mathbb{R}$, then you can assemble them into sequence in $\mathbb{R}^n$)

2.7 Proposition. *Let $(\vec{x}_k)_{k=1}^{\infty}$ be a sequence in $\mathbb{R}^n$, $\vec{a} \in \mathbb{R}^n$. Then we have*

$$\vec{x}_k \to \vec{a} \iff \forall i\ \ x_k^{(i)} \to a^{(i)}$$

Proof. ($\Rightarrow$) We have $\vec{x}_k \to \vec{a}$ in $\mathbb{R}^n$, we want to show that $x_k^{(i)} \to a^{(i)}$ in $\mathbb{R}$ for all $i \in [1, n]$.

Fix i. Observe that for every $k \geq 1$,

$$\left|\ x_k^{(i)} - a^{(i)}\ \right| = \left|\ (\vec{x}_k - \vec{a})^{(i)}\ \right| \leq ||\,\vec{x}_k - \vec{a}\,||$$

So for our fixed $i \in [1, n]$ we found that

$$0 \leq \left|\ x_k^{(i)} - a^{(i)}\ \right| \leq ||\,\vec{x}_k - \vec{a}\,|| \to 0$$

Use squeeze to get $\left| \; x_k^{(i)} - a^{(i)} \; \right| \to 0$. Hence $x_k^{(i)} \to a^{(i)}$.

($\Leftarrow$) First we have

$$\left| \; x_k^{(1)} - a^{(1)} \; \right| + \left| \; x_k^{(2)} - a^{(2)} \; \right| + \cdots + \left| \; x_k^{(n)} - a^{(n)} \; \right| \to 0$$

Hence

$$||\, \vec{x}_k - \vec{a}\, ||_1 \to 0$$

But

$$||\, \vec{x}_k - \vec{a}\, || \leq ||\, \vec{x}_k - \vec{a}\, ||_1, \; \forall k \in \mathbb{N}$$

So we get

$$0 \leq ||\, \vec{x}_k - \vec{a}\, || \leq ||\, \vec{x}_k - \vec{a}\, ||_1 \to 0$$

Use squeeze theorem to get $||\, \vec{x}_k - \vec{a}\, ||$ hence $\vec{x}_k \to \vec{a}$. $\square$

2.8 Remark. Another description of convergence of sequence in $\mathbb{R}^n$ is provided by the Cauchy condition (done in Math 147 for sequence in $\mathbb{R}$). Can upgrade to sequence in $\mathbb{R}^n$ by taking components.

Lecture 3, May. 5

Subsequences and the Bolzano-Weierstrass Theorem

3.1 Definition (Subsequence)**.** Let $(\vec{x}_k)_{k=1}^{\infty}$ be a sequence in $\mathbb{R}^n$. A *subsequece* of $(\vec{x}_k)_{k=1}^{\infty}$ is a sequence of the form $(\vec{x}_{k(p)})_{p=1}^{\infty}$ for some $1 \leq k(1) < k(2) < \cdots < k(p) < \cdots$.

Notation. Instead of $(\vec{x}_{k(p)})_{p=1}^{\infty}$ we may write

$$(\vec{x}_k)_{k \in P}$$

where $P = \{k(1), k(2), \cdots, k(p), \cdots\}$

How to avoid to talk about "sub-subsequences": from the subsequence $(\vec{x}_k)_{k \in P}$ can trim down to $(\vec{x}_k)_{k \in Q}$ where $Q \subseteq P$ is an infinite subset.

3.2 Remark. Let $(\vec{x}_k)_{k=1}^{\infty}$ be a sequence in $\mathbb{R}^n$, $\vec{a} \in \mathbb{R}^n$. If $\vec{x}_k \to \vec{a}$, then for every subsequence $(\vec{x}_{k(p)})_{p=1}^{\infty}$, we have $\vec{x}_{k(p)} \to \vec{a}$ as well.

3.3 Remark (Review from Math 147). Recall that a sequence $(t_k)_{k=1}^{\infty}$ in $\mathbb{R}$ is said to be bounded when $\exists r > 0$ such that $t_k \in [-r, r]$, $\forall k$.

3.4 Definition. A sequence $(\, \vec{x}_k \,)_{k=1}^{\infty}$ in $\mathbb{R}^n$ is said to be bounded when $\exists r \geq 0$ such that $||\, \vec{x}_k \,|| \leq r$, $\forall k \in \mathbb{N}$.

Note. In dimension $n = 1$, this gives back the condition $\mid x_k \mid \leq r$, $\forall k \in \mathbb{N}$, that is $t_k \in [-r, r]$, $\forall k$, which defines bounded sequences in $\mathbb{R}$.

3.5 Proposition. *Let $(\, \vec{x}_k \,)_{k=1}^{\infty}$ be a sequence in $\mathbb{R}^n$, and consider its n component sequences*

$$(x_k^{(1)})_{k=1}^{\infty}, \cdots, (x_k^{(n)})_{k=1}^{\infty}.$$

Then

$$\left(\begin{array}{c} (\, \vec{x}_k \,)_{k=1}^{\infty} \textit{ is bounded} \\ \textit{in } \mathbb{R}^n \end{array} \right) \Leftrightarrow \left(\begin{array}{c} \textit{each of } (x_k^{(1)})_{k=1}^{\infty}, \ldots, (x_k^{(n)})_{k=1}^{\infty} \textit{ is} \\ \textit{bounded in } \mathbb{R} \end{array} \right).$$

Proof. Exercise. □

3.6 Lemma. *Let $(\, \vec{x}_k \,)_{k=1}^{\infty}$ be a sequence $\mathbb{R}^{n+1}$, and for every $k \in \mathbb{N}$ write $\vec{x}_k = (\vec{y}_k, t_k)$ where $\vec{y}_k = (x_k^{(1)}, \cdots, x_k^{(n)}) \in \mathbb{R}^n$, $t_k = x_k^{(k+1)} \in \mathbb{R}$. Suppose that $\vec{y}_k \to \vec{b} \in \mathbb{R}^n$ and $t_k \to t \in \mathbb{R}$. Then $\vec{x}_k \to \vec{a}$ where $\vec{a} = (\vec{b}, t)$.*

Proof. Since $\vec{y}_k \to \vec{b}$, then $x_k^{(i)} \to b^{(i)}$ for all $1 \leq i \leq n$. Since $t_k \to t$, then $x_k^{(n+1)} \to b^{(n+1)}$. Then $x_k^{(i)} \to b^{(i)}$ for all $1 \leq i \leq n+1$. Hence $\vec{x}_k \to \vec{a}$ where $\vec{a} = (\vec{b}, t)$. □

3.7 Theorem (Bolzano-Weierstrass in $\mathbb{R}^n$)**.** *Let $(\, \vec{x}_k \,)_{k=1}^{\infty}$ be a bounded sequence in $\mathbb{R}^n$. Then we can find $1 \leq k(1) < k(2) < \cdots < k(p) < \cdots$ such that the subsequence $(\vec{x}_{k(p)})_{p=1}^{\infty}$ is convergent (to some limit in $\mathbb{R}^n$).*

Proof. By induction on n (dimension of the space).

Base case $n = 1$ was done in Math 147. Here we focus on the induction step $n \Rightarrow (n+1)$.

Assume B-W holds in $\mathbb{R}^n$. Fix a bounded sequence $(\, \vec{x}_k \,)_{k=1}^{\infty}$ in $\mathbb{R}^{n+1}$. Write every $\vec{x}_k = (\vec{y}_k, t_k)$ where $\vec{y}_k = (x_k^{(1)}, \cdots, x_k^{(n)}) \in \mathbb{R}^n$ and $t_k = x_k^{(n+1)} \in \mathbb{R}$.

Claim that $(\,\vec{y}_k\,)_{k=1}^{\infty}$ is a bounded sequence in $\mathbb{R}^n$ and $(t_k)_{k=1}^{\infty}$ is a bounded sequence in $\mathbb{R}$. Since $(\,\vec{x}_k\,)_{k=1}^{\infty}$ is bounded, then $\exists r > 0$ such that $||\,\vec{x}_k\,|| \leq r$, $\forall k \in \mathbb{N}$. But for every $k \in \mathbb{N}$ we have $r^2 \geq ||\,\vec{x}_k\,||^2 = ||\,\vec{y}_k\,||^2 + t_k^2$. Then $||\,\vec{y}_k\,|| \leq r$ and $|\,t_k\,| < r$. It follows that $(\,\vec{y}_k\,)_{k=1}^{\infty}$ is a bounded sequence in $\mathbb{R}^n$ and $(t_k)_{k=1}^{\infty}$ is a bounded sequence in $\mathbb{R}$.

Since B-W holds in $\mathbb{R}^n$, we can find an infinite set $M = \{k(1), k(2), \cdots\} \in \mathbb{N}$ such that the subsequence $(\vec{y}_{k(m)})_{m=1}^{\infty}$ of $(\,\vec{y}_k\,)_{k=1}^{\infty}$ is convergent to some limit $\vec{b} \in \mathbb{R}^n$.

Consider the sequence of real numbers $(t_k)_{k \in M}$. It is a bounded sequence in $\mathbb{R}$. Apply to it the B-W theorem in $\mathbb{R}$. and then we can find an infinite set $P \subseteq M$ such that the subsequence $(t_k)_{k \in P}$ is convergent to a limit $t \in \mathbb{R}$.

Since $\lim_{k \to \infty,\, k \in M} \vec{y}_k \to \vec{b}$ and $P \subseteq M$, we have $\lim_{k \to \infty,\, k \in P} \to \vec{y}_k = \vec{b}$ as well. because $(\vec{y}_k)_{k \in P}$ is a ssubsequence of $(\vec{y}_k)_{k \in M}$.

Now we have $\lim_{k \to \infty,\, k \in P} \vec{y}_k = \vec{b}$ and $\lim_{k \to \infty,\, k \in P} t_k = k$. Then by the previous lemma,

$$\lim_{k \to \infty,\, k \in P} (\vec{y}_k, t_k) = (\vec{b}, t).$$

Then the subsequence $(\vec{x}_k)_{k \in P}$ of $(\,\vec{x}_k\,)_{k=1}^{\infty}$ is convergent to $\vec{a} = (\vec{b}, t) \in \mathbb{R}^{n+1}$. $\square$

Lecture 4, May. 10

Continuity (of a function at a point) and Respect (for sequences)

4.1 Definition. Let $m, n \in \mathbb{N}$, $A \subseteq \mathbb{R}^n$, $f\colon A \to \mathbb{R}^m$, $\vec{a} \in A$. We say that f is <u>continuous at</u> $\vec{a}$ to mean that for all $\epsilon > 0$ there exists a $\delta > 0$ such that $||\, f(\vec{x}) - f(\vec{a}) \,|| < \epsilon$ for all $\vec{x} \in A$ with $||\, \vec{x} - \vec{a} \,|| < \delta$.

Geometric interpretation: given $\epsilon > 0$, we need to find $\delta > 0$ such that

$$f(A \cap \mathrm{B}(\vec{a}\colon \delta)) \subseteq \mathrm{B}(f(\vec{a})\colon \epsilon).$$

4.2 Definition. Let $m, n \in \mathbb{N}$, $A \subseteq \mathbb{R}^n$, $f\colon A \to \mathbb{R}^m$, $\vec{a} \in A$. We say that f <u>respects sequences in A</u> which converges to $\vec{a}$ to mean that whenever $(\, \vec{x}_k \,)_{k=1}^{\infty}$ in A is such that $\vec{x}_k \to \vec{a}$, it follows that $f(\vec{x}_k) \to f(\vec{a})$ in $\mathbb{R}^n$.

4.3 Proposition. *Let $m, n \in \mathbb{N}$, $A \subseteq \mathbb{R}^n$, $f\colon A \to \mathbb{R}^m$, $\vec{a} \in A$. Then f respects sequence that converges to $\vec{a}$ if and only if f is continuous at $\vec{a}$.*

Proof. ($\Rightarrow$) Assume that f respects sequence that converges to $\vec{a}$. Fix $\epsilon > 0$.

Assume, for a contradiction, that f is not continuous at $\vec{a}$. That is for all $\delta > 0$, there exists $\vec{x} \in A$ and $||\, \vec{x} - \vec{a} \,|| < \delta$ with that $||\, f((\vec{x}) - f(\vec{a}) \,|| \geq \epsilon$.

Let $\{\delta_n > 0\}$ be a sequence that converge to 0. For each δ_n there exists $\vec{x}_n \in A$ such that $||\, \vec{x}_n - \vec{a} \,|| < \delta$ with that $||\, f((\vec{x}_n) - f(\vec{a}) \,|| \geq \epsilon$. Since $\lim_{k\to\infty} ||\, \vec{x}_k - \vec{a} \,|| = 0$, we have $\vec{x}_k$ converges to $\vec{a}$. Since f respects sequence that converges to $\vec{a}$, $\{f(\vec{x}_n)\}$ should also converges. However, $||\, f(\vec{x}_n) - f(\vec{a}) \,|| \geq \epsilon$ for all n. It follows by contradiction that f is continuous at $\vec{a}$.

($\Leftarrow$) Assignment (direct application on the definition). □

4.4 Definition. Let $A \subseteq \mathbb{R}^n$, $f\colon A \to \mathbb{R}^m$. For every $\vec{x} \in A$ write explicitly $f(\vec{x}) = (f^{(1)}(\vec{x}), \cdots, f^{(m)}(x)) \in \mathbb{R}^m$. Then $f^{(1)}, \cdots, f^{(m)}\colon A \to \mathbb{R}$ are called the component functions of f.

4.5 Proposition. *Let $A \subseteq \mathbb{R}^n$, $f\colon A \to \mathbb{R}^m$, and $f^{(1}, \cdots, f^{(m)}\colon A \to \mathbb{R}$ as above. Let $\vec{a}$ be a point in A. Then*

$$\begin{pmatrix} \textit{f has continuity} \\ \textit{at } \vec{a} \end{pmatrix} \Leftrightarrow \begin{pmatrix} \textit{each of } f^{(1)}, \cdots, f^{(m)} \textit{ has} \\ \textit{continuity at } \vec{a} \end{pmatrix}.$$

Claim.

$$\begin{pmatrix} f \text{ respects sequences} \\ \text{at } \vec{a} \end{pmatrix} \Leftrightarrow \begin{pmatrix} \text{each of } f^{(1)}, \cdots, f^{(m)} \text{ respects} \\ \text{sequences at } \vec{a} \end{pmatrix}.$$

Proof. Take $(\vec{x}_k)_{k=1}^{\infty}$ in A such that $\vec{x}_k \to \vec{a}$. For every $k \in N$, write $f(\vec{x}_k) = (f^{(1)}(\vec{x}_k), \cdots, f^{(m)}(\vec{x}_k))$. Now apply Proposition 2.5, which says

$$f(\vec{x}_k) \to f(\vec{a}) \Leftrightarrow f^{(i)}(\vec{x}_k) \to f^{(i)}(\vec{a})$$

Then the claim follows. □

Claim.

$$\begin{pmatrix} f \text{ has continuity} \\ \text{at } \vec{a} \end{pmatrix} \Leftrightarrow \begin{pmatrix} \text{each of } f^{(1)}, \cdots, f^{(m)} \text{ has} \\ \text{continuity at } \vec{a} \end{pmatrix}.$$

Proof.

$$\begin{pmatrix} f \text{ has continuity} \\ \text{at } \vec{a} \end{pmatrix} \Leftrightarrow \begin{pmatrix} f \text{ respects continuity} \\ \text{at } \vec{a} \end{pmatrix} \Leftrightarrow$$
$$\begin{pmatrix} \text{each of } f^{(1)}, \cdots, f^{(m)} \text{ respects} \\ \text{sequences at } \vec{a} \end{pmatrix} \Leftrightarrow \begin{pmatrix} \text{each of } f^{(1)}, \cdots, f^{(m)} \text{ has} \\ \text{continuity at } \vec{a} \end{pmatrix}.$$

□

4.6 Definition. Let $A \subseteq \mathbb{R}^n$. $f\colon A \to \mathbb{R}^m$. We say that f is continuous on A to mean that f is continuous at every $\vec{a} \in A$.

4.7 Remark. We have two explicit description for "f continuous on A".

1. Given $\vec{a} \in A$ and $\epsilon > 0$, we can find $\delta > 0$ such that
$$||\, f(\vec{x}) - f(\vec{a})\, || < \epsilon$$
for all $\vec{x} \in A$ with $||\, \vec{x} - \vec{a}\, || < \delta$.
2. Whenever $(\, \vec{x}_k\,)_{k=1}^{\infty}$ in A has $\vec{x}_k \to \vec{a}$, it follows that $f(\vec{x}_k) \to f(\vec{a})$ in $\mathbb{R}^m$. (f respects all the convergent sequnce in A).

4.8 Remark (Stronger form of continuity). Let $A \subseteq \mathbb{R}^n$, $f\colon A \to \mathbb{R}^m$. We say that f is uniformly continuous on A to mean that given $\epsilon > 0$, we can find $\delta > 0$ such that
$$||\, f(\vec{x}_1) - f(\vec{x}_2)\, || < \epsilon$$
for all $\vec{x}_1, \vec{x}_2 \in A$ with $||\, \vec{x}_1 - \vec{x}_2\, || < \delta$.

4.9 Remark. Pick a $c \geq 0$. We say that f is c-Lipschitz on A to mean that
$$||\, f(\vec{x}_1) - f(\vec{x}_2)\, || \leq c \cdot ||\, \vec{x}_1 - \vec{x}_2\, ||, \;\; \forall \vec{x}_1, \vec{x}_2 \in A.$$

4.10 Remark. We say that f is Lipschitz on A to mean that $\exists c \geq 0$ such that f is c-Lipschitz.

4.11 Remark. Easy to check
$$f \text{ lipschitz on } A \Rightarrow f \text{ uniformly continuous on } A \Rightarrow f \text{ continuous on } A$$

Lecture 5, May. 15

Interior, closure, boundary for subsets of bR^n

5.1 Definition. Let A be a subset of $\mathbb{R}^n$. A point $\vec{a} \in A$ is said to be an *interior point* of A when there exists $r > 0$ such that $\mathrm{B}(\vec{a}, r) \subseteq A$. The set of all interior points of A is called the *interior* of A denoted by $\mathrm{int}(A)$.

A point $\vec{b} \in \mathbb{R}^n$ is said to be *adherent* to A when it has the property that $\mathrm{B}(\vec{b}, r) \cap A \neq \emptyset, \ \forall r > 0$. The set of all points in $\mathbb{R}^n$ that are adherent to A is called the *closure of* A denoted as $\mathrm{cl}(A)$.

5.2 Remark. For every $A \subseteq \mathbb{R}^n$ we have $\mathrm{int}(A) \subseteq A \subseteq \mathrm{cl}(A)$.

5.3 Definition. The set-difference $\mathrm{cl}(A) \setminus \mathrm{int}(A)$ is caled the *boundary* of A, denoted as $\mathrm{bd}(A)$.

5.4 Example. Say $m = 2$. Let

$$A = \{(s, t) \mid s, t \in \mathbb{R}, \ \ t > 0\} \cup \{(s, 0) \mid s \in \mathbb{R}, \ \ s \geq 0\}.$$

Calculate $\mathrm{int}(A)$, $\mathrm{cl}(A)$, $\mathrm{bd}(A)$.

Solution.

$$\begin{gathered}
\mathrm{int}(A) = \{(s, t) \mid s, t \in \mathbb{R}, \ \ t > 0\} \\
\mathrm{cl}(A) = \{(s, t) \mid s, t \in \mathbb{R}, \ \ t \geq 0\} \\
\mathrm{bd}(A) = \{(s, 0) \mid s \in \mathbb{R}\}
\end{gathered}$$

5.5 Proposition (Duality interior-closure)**.** *For every $A \subseteq \mathbb{R}^n$ we have that*

$$\mathrm{int}(\mathbb{R}^n \setminus A) = \mathbb{R}^n \setminus \mathrm{cl}(A)$$

$$\mathrm{cl}(\mathbb{R}^n \setminus A) = \mathbb{R}^n \setminus \mathrm{int}(A)$$

Proof. We first prove that $\mathrm{int}(\mathbb{R}^n \setminus A) \subseteq \mathbb{R}^n \setminus \mathrm{cl}(A)$. Take $\vec{b} \in \mathrm{int}(\mathbb{R}^n \setminus A)$. Then $\exists r > 0$ such that $\mathrm{B}(\vec{b}, r) \subseteq \mathbb{R}^n \setminus A$. For this $r > 0$, observe that $\mathrm{B}(\vec{b}, r) \cap A = \emptyset$. Then $\vec{b} \notin \mathrm{cl}(A)$ which means that $\vec{b} \in \mathbb{R}^n \setminus \mathrm{cl}(A)$.

Take $\vec{b} \in \mathbb{R}^n \setminus A$. Then $b \notin \mathrm{cl}(A)$. Then $\exists r > 0$ such that $\mathrm{B}(\vec{b}, r) \cap A = \emptyset$. For this r, we have $\mathrm{B}(\vec{b}, r) \subseteq \mathbb{R}^n \setminus A$. Then $\vec{b} \in \mathrm{int}(\mathbb{R}^n \setminus A)$.

Therefore $\mathrm{int}(\mathbb{R}^n \setminus A) = \mathbb{R}^n \setminus \mathrm{cl}(A)$. ◻

5.6 Corollary. *For every $A \subseteq \mathbb{R}^n$, we have* $\mathrm{bd}(A) = \mathrm{cl}(A) \cap \mathrm{cl}(\mathbb{R}^n \setminus A)$.

For every $A \subseteq \mathbb{R}^n$, we have $\mathrm{bd}(A) = \mathrm{bd}(\mathbb{R}^n \setminus A)$.

Proof.

$$\begin{aligned}
\mathrm{bd}(A) &= \mathrm{cl}(A) \setminus \mathrm{int}(A) \\
&= \mathrm{cl}(A) \cap (\mathbb{R}^n \setminus \mathrm{int}(A)) \\
&= \mathrm{cl}(A) \cap \mathrm{cl}(\mathbb{R}^n \setminus A)
\end{aligned}$$

$$\begin{aligned}
\mathrm{bd}(\mathbb{R}^n \setminus A) &= \mathrm{cl}(\mathbb{R}^n \setminus A) \cap \mathrm{cl}(\mathbb{R}^n \setminus (\mathbb{R}^n \setminus A)) \\
&= \mathrm{cl}(\mathbb{R}^n \setminus A) \cap \mathrm{cl}(A) \\
&= \mathrm{cl}(A) \cap \mathrm{cl}(\mathbb{R}^n \setminus A) & = \mathrm{bd}(A)
\end{aligned}$$

$\square$

(Something missing).

Lecture 6, May. 17

Open and closed subsets of $\mathbb{R}^n$

6.1 Definition. $A \subseteq \mathbb{R}^n$ is said to be *open* when it satisfied $A = \text{int}(A)$. That is A is open if and only if every $\vec{a} \in A$ is an interior point of A.

$A \subseteq \mathbb{R}^n$ is said to be *closed* when it satisfied $A = \text{cl}(A)$. That is A is closed if and only of A has no adherent points $\vec{b} \in \mathbb{R}^n \setminus A$.

Most subsets of $\mathbb{R}^n$ are neither open or closed.

In particular, if you have to prove A is closed, it will not suffices to prove A is not open.

6.2 Proposition. *For $A \subseteq \mathbb{R}^n$, we have A is closed if and only if $\mathbb{R}^n \setminus A$ is open.*

Proof. ($\Rightarrow$) A is closed, then $\text{cl}(A) = A$. But then $\text{int}(\mathbb{R}^n \setminus A) = \mathbb{R}^n \setminus \text{cl}(A) = \mathbb{R}^n \setminus A$. So we get $\text{int}(\mathbb{R}^n \setminus A) = \mathbb{R}^n \setminus A$, hence $\mathbb{R}^n \setminus A$ is open.

($\Leftarrow$) $\mathbb{R}^n \setminus A$ is open. then $\text{int}(\mathbb{R}^n \setminus A) = \mathbb{R}^n \setminus A$, then $\mathbb{R}^n \setminus \text{cl}(A) = \mathbb{R}^n \setminus A$. Then $\text{cl}(A) = A$ hence A is closed. $\square$

6.3 Definition. We say that a set $A \subseteq \mathbb{R}^n$ has the "no-escape" property for sequences to mean that whenever $(\,\vec{x}_k\,)_{k=1}^{\infty}$ is a sequence in A that converge to $\vec{b} \in \mathbb{R}^n$, then it follows that $\vec{b} \in A$.

6.4 Proposition. *For $A \subseteq \mathbb{R}^n$ we have A is closed if and only if A has no-escape property.*

Proof. Exercise. $\square$

Lecture 7, May. 19

A new look at interior and closure

7.1 Lemma. *Let $A \subseteq \mathbb{R}^n$. Assume that D is an open set. If $D \subseteq A$, then $D \subseteq \operatorname{int}(A)$.*

Proof. Call $\operatorname{int}(A) = T$. If $\vec{x} \in D$, $\square$

7.2 Lemma. $A \subseteq \mathbb{R}^n$. $\operatorname{int}(A)$ *is an open set.*

7.3 Proposition. *For every $A \subseteq \mathbb{R}^n$. The set* $\operatorname{int}(A)$ *is the largest open set contained in A.*

7.4 Proposition. *For every $A \subseteq \mathbb{R}^n$, the set* $\operatorname{cl}(A)$ *is the smallest closed set which contains A.*

Proof. Let $M = \mathbb{R}^n \setminus A$. We apply Proposition 7.3 to M. It says that

1. $\operatorname{int}(M) \subseteq M$ and $\operatorname{int}(M)$ is open.
2. If $D \subseteq \mathbb{R}^n$ is an open set such that $D \subseteq M$, then it follows that $D \subseteq \operatorname{int}(M)$.

$\operatorname{cl}(A) \supseteq A$ known from lecture 5.

$\operatorname{cl}(A) = \operatorname{cl}(\mathbb{R}^n \setminus M) = \mathbb{R}^n \setminus \operatorname{int}(M)$. Since $\operatorname{int}(M)$ is closed, then $\operatorname{cl}(A)$ is closed.

let $F \subseteq \mathbb{R}^n$ be a closed set such that $F \supseteq A$. Put $D = \mathbb{R}^n \setminus F$. Then D is open and $D \subseteq M$. Then $\mathbb{R}^n \setminus D \supseteq \mathbb{R}^n \setminus \operatorname{int}(M) = \operatorname{cl}(A)$. $\square$

Lecture 8, May. 23

Compact subsets of $\mathbb{R}^n$

8.1 Definition. $A \subseteq \mathbb{R}^n$ is said to be bounded when $\exists r \geq 0$ such that $||\,\vec{x}\,|| \leq r,\ \forall \vec{x} \in A$.

8.2 Definition. $A \subseteq \mathbb{R}^n$ is said to be compact when it is closed and bounded.

Note. There are several other equivalent description of compactness, and some of them extends to more general frameworks

8.3 Definition. $A \subseteq \mathbb{R}^n$ is said to be *sequentially compact* when the following happens: For every sequence $(\,\vec{x}_k\,)_{k=1}^{\infty} \in A$ one can find a convergent subsequence $(\vec{x}_{k(p)})_{p=1}^{\infty}$ such that the limit $\vec{a} = \lim_{p\to\infty} \vec{x}_{k(p)}$ still is in A.

8.4 Theorem. *For $A \subseteq \mathbb{R}^n$ have that*

$$A \text{ is compact} \Leftrightarrow A \text{ has sequential compactness}$$

Proof. ($\Rightarrow$) Suppose A is closed and bounded. Let $(\,\vec{x}_k\,)_{k=1}^{\infty}$ be a sequence in A. Then $(\,\vec{x}_k\,)_{k=1}^{\infty}$ is bounded. Then we can extract a convergent subsequence $(\vec{x}_{k(p)})_{p=1}^{\infty}$ (Bolzano-Weierstrass). Denote $\lim_{p\to\infty} \vec{x}_{k(p)} = \vec{a}$. Since A is closed and $(\vec{x}_{k(p)})_{p=1}^{\infty}$ is a sequence in A, it follows that $\vec{a} \in A$ (by the no-escape property of A).

($\Leftarrow$) Suppose that A has sequential compactness.

Assume by contradiction that A is not closed, hence a does not have the no-esape property for sequences. So there exists sequence $(\,\vec{x}_k\,)_{k=1}^{\infty} \in A$ with $\lim_{k\to\infty} \vec{x}_k = \vec{b} \in \mathbb{R}^n \setminus A$. Property sequetial compactness says that we can find $(\vec{x}_{k(p)})_{p=1}^{\infty}$ that converges to a limit $\vec{a} \in A$. But also have that $\vec{x}_k \to \vec{b} \Rightarrow \vec{x}_{k(p)} \to \vec{b}$. (Contradiction.) $\square$

8.5 Proposition. *Let $A \subseteq \mathbb{R}^n$ be compact and let $f\colon A \to \mathbb{R}^n$ be a continuous function. Consider the image-set $M = f(A) = \{\vec{y} \in \mathbb{R}^n \mid \exists \vec{x},\ \ f(\vec{x}) = \vec{y}\}$. Then M is a compact subset of $\mathbb{R}^n$.*

Proof. We will verify that M is sequentially compact. So fix a sequence $(\,\vec{y}_k\,)_{k=1}^{\infty}$ in M. We need to find a convergent subsequence $(\vec{y}_{k(p)})_{p=1}^{\infty}$ with limit still in M.

For every $k \in \mathbb{N}$, we have $\vec{y}_k \in M = f(A)$. Hence $\exists \vec{x}_k \in A$ such that $f(\vec{x}_k) = \vec{y}_k$.

Since A is compact, hence it is sequentially compact. So for the sequence $(\,\vec{x}_k\,)_{k=1}^{\infty}$ in A we can find $1 \leq k(1) \leq k(2) \leq \cdots$ such that $\vec{x}_{k(p)} \to \vec{a} \in A$.

Since f is continuous on A, hence it respects convergence of sequences in A. It follows that $f(\vec{x}_{k(p)}) \to f(\vec{a})$. But $f(\vec{x}_{k(p)}) = \vec{y}_{k(p)}$. Thus $\vec{y}_{k(p)} \to f(\vec{a}) \in M$. $\square$

8.6 Definition. $A \subseteq \mathbb{R}^n$, $f\colon A \to \mathbb{R}$.

1. $\vec{a} \in A$ is said to be a point of global minimum for f on A when $f(\vec{a}) \leq f(\vec{x})$ for all $\vec{x} \in A$
2. $\vec{a} \in A$ is said to be a point of global maximum for f on A when $f(\vec{x}) \leq f(\vec{a})$ for all $\vec{x} \in A$

8.7 Remark. Points of global min/max may or may not exist. When they exist, they may or may not be unique.

8.8 Theorem (E.V.T.)**.** *$A \subseteq \mathbb{R}^n$ compact, $f\colon A \to \mathbb{R}$ continuous. Then f has at least one point of global minimun and at least one point of global maximum on A.*

8.9 Lemma. *Let K be a non-empty compact subset of $\mathbb{R}$. Then $\exists \alpha, \beta \in K$ such that $\alpha \leq t \leq \beta,\ \forall t \in K$.*

Proof. Since K is compact, then K is closed and bounded. Then exists $r > 0$ such that $K \subseteq [-r, r]$. Then K has lower bounds and upper bounds. Then K has a least upper bound $\beta = \sup(K)$ and a greatest lower bound $a = \inf(K)$. For every $k \in \mathbb{N}$, we have $(\beta - 1/k, \beta] \cap K \neq \emptyset$. Hence we can construct a sequence in K that converges to β. Since K is closed, it has the no-escape property. Hence $\beta \in K$. $\square$

Proof of E.V.T.. We want to prove $\exists \vec{a}, \vec{b} \in A$ such that $f(\vec{a}) \leq f(\vec{x}) \leq f(\vec{b}),\ \forall \vec{x} \in A$.

Consider image-set $K = f(A) =\subseteq \mathbb{R}$. Then K is a non-empty compact subset of $\mathbb{R}$. Consider $\alpha = \min(K)$, $\beta = \max(K)$. we have $\alpha, \beta \in K = f(A)$. Hence $\exists \vec{a}, \vec{b} \in A$ such that $f(\vec{a}) = \alpha$ and $f(\vec{b}) = \beta$. But then for every $\vec{x} \in A$ we can write $f(\vec{x}) \in f(A) = K$ then $\alpha \leq f(\vec{x}) \leq \beta$ then $f(\vec{a}) \leq f(\vec{x}) \leq f(\vec{b})$. $\square$

Lecture 9, May. 26

Directional Derivatives

9.1 Remark. $A \subseteq \mathbb{R}^n$, $\vec{a} \in \operatorname{int}(A)$, $\vec{v} \in \mathbb{R}^n$ (a direction). Then $\exists c > 0$ such that $\vec{a} + t\vec{v} \in A, \ \forall t \in (-c, c)$.

9.2 Definition. $A \subseteq \mathbb{R}^n$, $f \colon A \to \mathbb{R}$, $\vec{a} \in \operatorname{int}(A)$. Let $\vec{v}$ be any vector in $\mathbb{R}^n$. Note that exists $c > 0$ such that the quotient

$$\frac{f(\vec{a} + t\vec{v}) - f(\vec{a})}{t}$$

is defined for every $t \in (-c, c) \setminus \{0\}$. If the limit

$$L = \lim_{t \to 0} \frac{f(\vec{a} + t\vec{v}) - f(\vec{a})}{t} \in \mathbb{R}$$

exists, then we say that f has directional derivative in direction $\vec{v}$ at the point $\vec{a}$. The notation for L is $(\partial_{\vec{v}} f)(\vec{a})$.

The limit L exists if and only if whenever $(t_k)_{k=1}^{\infty}$ in $(-c, c) \setminus \{0\}$, we have if $t_k \to 0$ then

$$\lim_{t \to 0} \frac{f(\vec{a} + t_k\vec{v}) - f(\vec{a})}{t} \to L.$$

9.3 Remark. If $\vec{v} = \vec{0}$ then $(\partial_{\vec{v}} f)(\vec{a})$ is sure to exists and is equal to 0.

9.4 Proposition. *$A \subseteq \mathbb{R}^n$, $f \colon A \to \mathbb{R}$, $\vec{a} \in \operatorname{int}(A)$. Let $\vec{v} \neq \vec{0} \in \mathbb{R}^n$. Suppose that $(\partial_{\vec{v}} f)(\vec{a})$ exists. Then for every $\alpha \in \mathbb{R}$, the directional derivative $(\partial_{\alpha\vec{v}} f)(\vec{a})$, and*

$$(\partial_{\alpha\vec{v}} f)(\vec{a}) = \alpha(\partial_{\vec{v}} f)(\vec{a})$$

Proof. If $\alpha = 0$, then the equation becomes $0 = 0$.

Assume $\alpha \neq 0$. Denote $\alpha\vec{v} = \vec{w}$. Then

$$\begin{aligned}
&\lim_{t \to 0} \frac{f(\vec{a} + t\vec{w}) - f(\vec{a})}{t} \\
=&\lim_{t \to 0} \frac{f(\vec{a} + t\alpha\vec{v}) - f(\vec{a})}{t\alpha}\alpha \\
=&\lim_{s \to 0} \frac{f(\vec{a} + s\vec{v}) - f(\vec{a})}{s}\alpha \\
=&(\partial_{\vec{v}} f)(\vec{a}) \cdot \alpha
\end{aligned}$$

$\square$

9.5 Definition. $A \subseteq \mathbb{R}^n$, $f \colon A \to \mathbb{R}$, $\vec{a} \in \operatorname{int}(A)$. Suppose $(\partial_{\vec{v}} f)(\vec{a})$ exists for all $\vec{c} \in \mathbb{R}^n$. We say that "additivity in $\vec{v}$ holds" to mean that

$$\text{(Add)} \qquad (\partial_{\vec{v}_1 + \vec{v}_2} f)(\vec{a}) = (\partial_{\vec{v}_1} f)(\vec{a}) + (\partial_{\vec{v}_2} f)(\vec{a}), \ \forall \vec{v}_1, \vec{v}_2 \in \mathbb{R}^n$$

9.6 Remark. $A \subseteq \mathbb{R}^n$, $f \colon A \to \mathbb{R}$, $\vec{a} \in \operatorname{int}(A)$. Assume $(\partial_{\vec{v}} f)(\vec{a})$ exists for all $\vec{v} \in \mathbb{R}^n$. Then (Add) may or may not hold.

9.7 Definition. $A \subseteq \mathbb{R}^n$, $f\colon A \to \mathbb{R}$, $\vec{a} \in \operatorname{int}(A)$. Consider the special basis $\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_n$ of $\mathbb{R}^n$ where $e_i = \{0, \cdots, 0, 1, 0, \cdots, 0\}$.

When $(\partial_{\vec{e}_i} f)(\vec{a})$ exists, it is called the i-th partial derivative of f at $\vec{a}$ and is denoted as $(\partial_i f)(\vec{a})$.

Suppose all n partials $(\partial_i f)(\vec{a})$ exist for $1 \leq i \leq n$. Then

$$(\nabla f)(\vec{a}) = ((\partial_1 f)(\vec{a}), \cdots, (\partial_n f)(\vec{a})) \in \mathbb{R}^n$$

is called the gradient vector of f at $\vec{a}$.

9.8 Proposition. $A \subseteq \mathbb{R}^n$, $f\colon A \to \mathbb{R}$, $\vec{a} \in \operatorname{int}(A)$. *Suppose* $(\partial_{\vec{v}} f)(\vec{a})$ *exists for all* $\vec{v} \in bR^n$, *and that we have (Add) property. Then for every* $\vec{v} \in \mathbb{R}^n$, *we have*

$$(\partial_{\vec{v}} f)(\vec{a}) = \langle \vec{v}\,,\, (\nabla f)(\vec{a})\,\rangle$$

Proof. Fix $\vec{v} = (v^{(1)}, \cdots, v^{(n)}) \in \mathbb{R}^n$, and write $\vec{v} = v^{(1)}\vec{e}_1 + \cdots + v^{(n)}\vec{e}_n$. Then

$$\begin{aligned}(\partial_{\vec{v}} f)(\vec{a}) &= (\partial_{v^{(1)}\vec{e}_1 + \cdots + v^{(n)}\vec{e}_n} f)(\vec{a}) \\ &= (\partial_{v^{(1)}\vec{e}_1} f)(\vec{a}) + \cdots + (\partial_{v^{(n)}\vec{e}_n} f)(\vec{a}) \\ &= v^{(1)}(\partial_1 f)(\vec{a}) + \cdots + v^{(1)}(\partial_n f)(\vec{a}) \\ &= \langle \vec{v}\,,\, (\nabla f)(\vec{a})\,\rangle\end{aligned}$$

$\square$

9.9 Remark. $(\partial_i f)(\vec{a})$ are computed in practical examples by using Calculus 1.

Rule of thumb: if $\vec{x} = (x^{(1)}, \cdots, x^{(n)}) \in A$, treat all $x^{(j)}$ with $j \neq i$ as constants $x^{(j)} = a^{(j)}$ and do 1-dimensional derivative with respect to $x^{(i)}$ at $a^{(i)}$.

9.10 Example.

$$f((x,y,z)) = x \cdot \sin y \cdot z^2 - 3 \arctan z$$

The rule of thumb says

$$(\partial_1 f)(x,y,z) = \sin y \cdot z^2$$

$$(\partial_2 f)(x,y,z) = (xz^2)\cos y$$

$$(\partial_3 f)(x,y,z) = 2x \cdot \sin y \cdot z - \frac{3}{1+z^2}$$

Lecture 10, May. 31

Two basic applications of directional derivatives

10.1 Definition. $A \subseteq \mathbb{R}^n$, $f\colon A \to \mathbb{R}$. We say that $\vec{a} \in A$ is a point of local minimum for f to mean that $\exists r > 0$ such that $f(\vec{a}) \leq f(\vec{x})$ for all $\vec{x} \in A \cap \mathrm{B}(\vec{a}, r)$.

We say a point $\vec{b} \in A$ is a local maximum to mean that $\exists r > 0$ such that $f(\vec{b}) \geq f(\vec{x})$ for all $\vec{x} \in A \cap \mathrm{B}(\vec{a}, r)$.

A point of A which is either a point of local min or a point of local max for f is called a point of local extremum for f.

10.2 Proposition. *$A \in \mathbb{R}^n$, $f\colon A \to \mathbb{R}$, $\vec{a} \in \mathrm{int}(A)$, $\vec{v} \in \mathbb{R}^n$ and suppose that $(\partial_{\vec{v}} f)(\vec{a})$ exists. If $\vec{a}$ is a point of local extremum for f, then $(\partial_{\vec{v}} f)(\vec{a}) = 0$.*

Proof. Assume that $\vec{a}$ is a local minimum. Let $r > 0$ be such that $\mathrm{B}(\vec{a}, r) \subseteq A$ and $f(\vec{a}) \leq f(\vec{x})$ for all $\vec{x} \in \mathrm{B}(\vec{a}, r)$. Let $c = \frac{r}{1 + ||\,\vec{v}\,||}$. Then have $\vec{a} + t\vec{v} \in \mathrm{B}(\vec{a}, r)$, $\forall t \in (-c, c)$. Define $h\colon (-c, c) \to \mathbb{R}$

$$h(t) = f(\vec{a} + t\vec{v})$$

For every $t \in I$, we have $h(t) = f(\vec{a} + t\vec{v}) \geq f(\vec{a}) = h(0)$. Hence 0 is a point of minimum for h on I.

Then we have

$$\lim_{t \to 0} \frac{h(t) - h(0)}{t - 0} = \lim_{t \to 0} \frac{f(\vec{a} + t\vec{v}) - f(\vec{a})}{t} = (\partial_{\vec{v}} f)(\vec{a})$$

Since $(\partial_{\vec{v}} f)(\vec{a})$ exists, then the left hand side must exist as well. Then $h'(0) = (\partial_{\vec{v}} f)(\vec{a})$. Since $h'(0)$ exists and 0 is a point of minimum, then $h'(0) = 0$. So $(\partial_{\vec{v}} f)(\vec{a}) = 0$, as required. $\square$

10.3 Corollary (Gradient-test for local extremum)**.** *$A \subseteq \mathbb{R}^n$, $f\colon A \to \mathbb{R}$, $\vec{a} \in \mathrm{int}(A)$, and suppose that f has partial derivatives at $\vec{a}$, that is $(\partial_{\vec{v}} f)(\vec{a})$ exists for all $1 \leq i \leq n$. If $\vec{a}$ is point of local extremum for f, then $(\nabla f)(\vec{a}) = \vec{0}$. When $(\nabla f)(\vec{a}) = \vec{0}$, one says that $\vec{a}$ is stationary point for f.*

Notation. For $\vec{x}, \vec{y} \in \mathbb{R}^n$ denote $\mathrm{Co}(\vec{x}, \vec{y}) = \{(1 - t)\vec{x} + t\vec{y} \mid t \in [0, 1]\}$. $\mathrm{Co}(\vec{x}, \vec{y})$ is called the line segment in $\mathbb{R}^n$ with endpoints $\vec{x}$ and $\vec{y}$.

10.4 Example. Let A be an open convex set in $\mathbb{R}^n$. $f\colon A \to \mathbb{R}$. $(\partial_{\vec{v}} f)(\vec{a})$ exists for all $\vec{\in} A, \vec{v} \in \mathbb{R}^n$. Then $(\partial_{\vec{v}} f)(\vec{a}) = 0$ for all $\vec{\in} A, \vec{v} \in \mathbb{R}^n$

Lecture 11, June. 5

C^1-functions and their linear approximation

11.1 Definition. $A \subseteq \mathbb{R}^n$, $f\colon A \to \mathbb{R}$. We say that f is a C^1-function to mean that it has the following properties

- f is continuous on A
- f has partial derivatives $(\partial_i f)(\vec{a})$ exists for all $\vec{a} \in A$ and $i \in \{1, \cdots, n\}$
- The new functions $\partial_i f\colon A \to \mathbb{R}$ are continuous on A.

The collection of all C^1-functions from A to $\mathbb{R}$ is denoted as $C^1(A, \mathbb{R})$.

One also uses notation

$$C^0(A, \mathbb{R}) = \{f\colon A \to \mathbb{R} \mid f \text{ is continuous on } A\}$$

11.2 Remark.

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

11.3 Theorem. $A \subseteq \mathbb{R}^n$, $f \in C^1(A, \mathbb{R})$. *Then for every* $\vec{a} \in A$ *we have*

$$(\mathrm{L-Approx}) \qquad \lim_{\vec{x} \to \vec{a}} \frac{|\ f(\vec{x}) - f(\vec{a}) - \langle \vec{x} - \vec{a}\,,\, (\nabla f)(\vec{a})\,\rangle\ |}{||\,\vec{x} - \vec{a}\,||} = 0$$

11.4 Corollary. $A \subseteq \mathbb{R}^n$ *open,* $f \in C^1(A, \mathbb{R})$, $\vec{a} \in A$. *Then for every* $\vec{v} \in \mathbb{R}^n$, *the directional derivative* $(\partial_{\vec{v}} f)(\vec{a})$ *exists, and have* $(\partial_{\vec{v}} f)(\vec{a}) = \langle \vec{v}\,,\, (\nabla f)(\vec{a})\,\rangle$.

Proof. In (L-Approx) we pick $\vec{x}$ of the form $\vec{a} + t\vec{v}$. Then $\vec{x} \to \vec{a}$ be comes $t \to 0$.

Then multiply the limit by $||\,\vec{v}\,||$

$$\lim_{t \to 0} \frac{|\ f(\vec{a} + t\vec{v}) - f(\vec{a}) - \langle (\vec{a} + t\vec{v}) - \vec{a}\,,\, (\nabla f)(\vec{a})\,\rangle\ |}{||\,\vec{a} + t\vec{v} - \vec{a}\,||} \cdot ||\,\vec{v}\,|| = 0||\,\vec{v}\,|| = 0$$

$$\lim_{t \to 0} \left| \frac{f(\vec{a} + t\vec{v}) - f(\vec{a}) - t\langle \vec{v}\,,\, (\nabla f)(\vec{a})\,\rangle}{t} \right| = 0$$

$$\lim_{t \to 0} \left| \frac{f(\vec{a} + t\vec{v}) - f(\vec{a})}{a} - \langle \vec{v}\,,\, (\nabla f)(\vec{a})\,\rangle \right| = 0$$

$$\lim_{t \to 0} \frac{f(\vec{a} + t\vec{v}) - f(\vec{a})}{a} = \langle \vec{v}\,,\, (\nabla f)(\vec{a})\,\rangle$$

□

11.5 Corollary. $A \subseteq \mathbb{R}^n$ *open,* $f \in C^1(A, \mathbb{R})$, $\vec{a} \in A$. *The directional derivatives at* $\vec{a}$ *have (Add) property.*

11.6 Lemma. $A \subseteq \mathbb{R}^n$ *open,* $f \in C^1(A, \mathbb{R})$, $\vec{a} \in \mathbb{R}$. *Pick* $r > 0$ *such that* $\mathrm{B}(\vec{a}, r) \subseteq A$. *Then for every* $\vec{x} \in \mathrm{B}(\vec{a}, r)$ *we can find* $\vec{b}_1, \cdots \vec{b}_n \in \mathrm{B}(\vec{a}, r)$ *such that* $f(\vec{x}) - f(\vec{a}) = \langle \vec{x} - \vec{a}\,,\, \vec{w}\,\rangle$ *with* $\vec{w} = ((\partial_1 f)(\vec{b}_1), \cdots, (\partial_n f)(\vec{b}_n))$.

Proof. Fix $\vec{x} \in \mathrm{B}(\vec{a}, r)$. Consider vectors $\vec{x}_0, \vec{x}_1, \cdots, \vec{x}_n$, defined as follows:

$$\vec{x}_0 = \vec{a} = (a^{(1)}, \cdots, a^{(n)})$$
$$\vec{x}_1 = \vec{a} = (x^{(1)}, \cdots, a^{(n)})$$
$$\vec{x}_2 = \vec{a} = (x^{(1)}, x^{(2)}, \cdots, a^{(n)})$$
$$\cdots$$
$$\vec{x}_n = \vec{a} = (x^{(1)}, \cdots, x^{(n)}) = \vec{x}$$

Note that for every $1 \leq i \leq n$ we have $||\, \vec{x}_i - \vec{a}\, || \leq ||\, \vec{x} - \vec{a}\, || < r$. Hence $\vec{x}_0, \vec{x}_1, \cdots, \vec{x}_n \in \mathrm{B}(\vec{a}, r) \subseteq A$.

Claim for every $1 \leq i \leq n$ there exists $\vec{b}_i \in \mathrm{Co}(\vec{x}_{i-1}, \vec{x}_i)$ such that

$$f(\vec{x}_i) - f(\vec{x}_i) - f(\vec{x}_{i-1}) = (x^{(i)} - a^{(i)})(\partial_i f)(\vec{b}_i)$$

Verification of the claim.

$$\begin{aligned} &\vec{x}_i - \vec{x}_{i-1} \\ =&(x^{(i)} - a^{(i)}) \cdot \vec{e}_i \\ =&\alpha \vec{e}_i \end{aligned}$$

Apply MVT in direction $\vec{e}_i$ with endpoints $\vec{x}_{i-1}$ and $\vec{x}_i$, then $\exists \vec{b}_i \in \mathrm{Co}(\vec{x}_{i-1}, \vec{x}_i)$ such that $f(\vec{x}_i) - f(\vec{x}_{i-1}) = (x^{(i)} - a^{(i)})(\partial_i f)(\vec{b}_i)$. Done with claim.

Then

$$\begin{aligned} f(\vec{x}) - f(\vec{a}) &= f(\vec{x}_m) - f(\vec{x}_0) \\ &= f(\vec{x}_m) - f(\vec{x}_{m-1}) + \cdots + f(\vec{x}_1) - f(\vec{x}_0) \\ &= \sum_{i=1}^{m} f(\vec{x}_i) - f(\vec{x}_{i-1}) \\ &= \sum_{i=1}^{m} (x^{(i)} - a^{(i)})(\partial_i f)(\vec{b}_i) \\ &= \langle \vec{x} - \vec{a}\, ,\, \vec{w}\, \rangle \end{aligned}$$

where $\vec{w} = ((\partial_1 f)(\vec{b}_1), \cdots, (\partial_n f)(\vec{b}_n))$. □

Proof of Theorem 11.3. Given $\epsilon > 0$, we want to find $\delta > 0$ such that $\mathrm{B}(\vec{a}, \delta) \subseteq A$ and such that

$$\text{(Want)} \qquad \frac{|\ f(\vec{x}) - f(\vec{a}) - \langle \vec{x} - \vec{a}\, ,\, (\nabla f)(\vec{a})\, \rangle\ |}{||\, \vec{x} - \vec{a}\, ||} < \epsilon$$

for all $\vec{x} \in \mathrm{B}(\vec{a}, \delta) \setminus \{\vec{a}\}$.

Fix $r_0 > 0$ such that $\mathrm{B}(\vec{a}, r_0) \subseteq A$. For every $1 \leq i \leq n$, we know that $\partial_i f$ is continuous at $\vec{a}$ hence $\exists 0 \leq r_i \leq r_0$ such that for all $\vec{y} \in \mathrm{B}(\vec{a}, r_i)$ we have

$$|\ (\partial_i f)(\vec{y}) - (\partial_i f)(\vec{a})\ | < \frac{\epsilon}{n}$$

Put $\delta = \min(r_1, \cdots, r_n)$. Claim δ is good for (Want).

Verification of claim. Pick $\vec{x} \in \mathrm{B}(\vec{a}, \delta) \setminus \{a\}$ for which we prove that

$$(\text{Want}') \qquad |\, f(\vec{x}) - f(\vec{a}\,| - \langle \vec{x} - \vec{a}\,,\, (\nabla f)(\vec{a})\,\rangle < \epsilon ||\, \vec{x} - \vec{a}\,||.$$

Lemma 11.6 gives us points $\vec{b}_1, \cdots, \vec{b}_n \in \mathrm{B}(\vec{a}, \delta)$ such that

$$f(\vec{x} - \vec{a}) = \langle \vec{x} - \vec{a}\,,\, \vec{w}\,\rangle$$

where $\vec{w} = ((\partial_1 f)(\vec{b}_1), \cdots, (\partial_n f)(\vec{b}_n))$.

Then

$$\begin{aligned}
&|\, f(\vec{x}) - f(\vec{a}\,| - \langle \vec{x} - \vec{a}\,,\, (\nabla f)(\vec{a})\,\rangle \\
=&|\, \langle \vec{x} - \vec{a}\,,\, \vec{w}\,\rangle - \langle \vec{x} - \vec{a}\,,\, (\nabla f)(\vec{a})\,\rangle\,| \\
=&|\, \langle \vec{x} - \vec{a}\,,\, \vec{w} - (\nabla f)(\vec{a})\,\rangle\,| \\
\leq&||\, \vec{x} - \vec{a}\,|| \cdot ||\, \vec{w} - (\nabla f)(\vec{a})\,|| \\
\leq&||\, \vec{x} - \vec{a}\,|| \cdot ||\, \vec{w} - (\nabla f)(\vec{a})\,||_1 \\
=&||\, \vec{x} - \vec{a}\,|| \cdot \sum_{i=1}^{m} \Big|\, (\partial_i f)(\vec{b}_i) - (\partial_i f)(\vec{a})\,\Big| \\
<&||\, \vec{x} - \vec{a}\,|| \cdot \sum_{i=1}^{m} \frac{\epsilon}{n} \\
=&\epsilon \cdot ||\, \vec{x} - \vec{a}\,||
\end{aligned}$$

□

Lecture 12, June. 9

Geometric interpretation of (L-Approx)

Theorem (Theorem 11.3)**.** *$A \subseteq \mathbb{R}^n$, $f \in C^1(A,\mathbb{R})$. Then for every $\vec{a} \in A$ we have*

$$(\mathrm{L-Approx}) \qquad \lim_{\vec{x}\to\vec{a}} \frac{|\ f(\vec{x}) - f(\vec{a}) - \langle \vec{x}-\vec{a}\,,\,(\nabla f)(\vec{a})\,\rangle\ |}{||\,\vec{x}-\vec{a}\,||} = 0$$

In particular, for $\vec{x}$ close to $\vec{a}$ we have

$$f(\vec{x}) \approx f(\vec{a}) + \langle \vec{x}-\vec{a}\,,\,(\nabla f)(\vec{a})\,\rangle$$

Review case $n = 1$

$$f(x) \approx f(a) + (x-a)f'(a)$$

Linear approximation says that for x close to a we can approximate Q by Q' where $Q = (x, f(x))$ and $Q' = (x, f(a) + (x-a)f'(a))$.

12.1 Definition. $A \subseteq \mathbb{R}^n$, $f\colon A \to \mathbb{R}$. The graph of f is the set of points

$$\Gamma = \{(\vec{x}, t) \in \mathbb{R}^{n+1} \mid \vec{x} \in A,\ t = f(\vec{x})\}$$

12.2 Definition. A hyperplane going through a point $\vec{p} \in \mathbb{R}^{n+1}$ is a set of the form

$$H = \{\vec{p} + \sum_{i=1}^{n} \alpha_i \vec{y}_i \mid \alpha_1, \cdots, \alpha_n \in \mathbb{R}\}$$

where $\vec{y}_1, \cdots \vec{y}_n \in \mathbb{R}^{n+1}$ are linearly independent.

A vector $\vec{z} \in \mathbb{R}^{n+1}$, $\vec{z} \neq \vec{0}$, such that $\vec{z} \perp \vec{y}_i$, $\forall 1 \leq i \leq n$ is said to be a normal vector to the hyperplane H.

12.3 Remark. Consider hyperplane $H \subseteq \mathbb{R}^{n+1}$ as in Definition 12.2. If $\vec{z}_1, \vec{z}_2$ are normal vectors to H, then $\exists \alpha \in \mathbb{R} \setminus \{0\}$ such that $\vec{z}_2 = \alpha \vec{z}_1$

12.4 Exercise. Given $A \subseteq \mathbb{R}^n$ open, $f \in C^1(A,\mathbb{R})$, $\vec{a} \in A$. Let Γ be the graph of f. What is the equation of the hyperplane tangent to Γ at the point $\vec{p} = (\vec{a}, f(\vec{a})) \in \Gamma$?

Solution. Pick $\delta > 0$ such that $\mathrm{B}(\vec{a}, \delta) \subseteq A$. Pick $\vec{x} \in A$ of the form $\vec{x} = \vec{a} + \vec{v}$ with $||\,\vec{v}\,|| < \delta$. Look at the points $\vec{q} = (\vec{x}, f(\vec{x})) \in \Gamma$ and

$$\vec{q}' = (\vec{x}, f(\vec{a}) + \langle \vec{x}-\vec{a}\,,\,(\nabla f)(\vec{a})\,\rangle)$$

(L-Approx) says $\vec{q}'$ is the linear approximation of $\vec{q}$.

Our exercise becomes: how do we define H so that $\vec{q}'$ is sure to be in H.

$$\begin{aligned}
\vec{q}' - \vec{p} &= (\vec{x}, f(\vec{a}) + \langle \vec{x}-\vec{a}\,,\,(\nabla f)(\vec{a})\,\rangle) - (\vec{a}, f(\vec{a})) \\
&= (\vec{x}-\vec{a}, \langle \vec{x}-\vec{a}\,,\,(\nabla f)(\vec{a})\,\rangle) \\
&= (\vec{v}, \langle \vec{v}\,,\,(\nabla f)(\vec{a})\,\rangle) \\
&= (v^{(1)}, \cdots, v^{(n), v^{(1)}(\partial_1 f)(\vec{a}) + \cdots + v^{(n)}(\partial_n f)(\vec{a})}) \\
&= v^{(1)} \cdot (1, 0, \cdots, 0, v^{(1)}(\partial_1 f)(\vec{a})) + \cdots + v^{(n)} \cdot (0, 0, \cdots, 1, v^{(n)}(\partial_n f)(\vec{a})) \\
&= v^{(1)}\vec{y}_1 + \cdots + v^{(n)}\vec{y}_n
\end{aligned}$$

where $\vec{y}_i = (0, \cdots, \underbrace{0, 1, 0}_{\text{position} i}, \cdots, 0, (\partial_i f)(\vec{a}))$.

12.5 Definition. Given $A \subseteq \mathbb{R}^n$ open, $f \in C^1(A, \mathbb{R})$, $\vec{a} \in A$. Let Γ be the graph of f. Let $\vec{a} \in A$ and $\vec{p} = (\vec{a}, f(\vec{a})) \in \Gamma$. The hyperplane

$$H = \{\vec{p} + \sum_{i=1}^{n} \alpha_i \vec{y}_i \mid \alpha_1, \cdots, \alpha_n \in \mathbb{R}\}$$

with $\vec{y}_i = (0, \cdots, \underbrace{0, 1, 0}_{\text{position} i}, \cdots, 0, (\partial_i f)(\vec{a}))$ for $1 \leq i \leq n$ is called the tangent hyperplane to Γ at the point $\vec{p}$.

12.6 Remark. What is the normal vector to the tangent hyperplane of Definition 12.5?

We need $\vec{z} \in \mathbb{R}^{n+1}$ such that $\vec{z} \perp \vec{y}_i$ for all $1 \leq i \leq n$. Observe that

$$\vec{z} = (-(\partial_1 f)(\vec{a}), \cdots, -(\partial_n f)(\vec{a}), 1)$$

will do.

Lecture 13, June. 12

The Chain Rule

13.1 Remark. Let $f, g\colon \mathbb{R} \to \mathbb{R}$ differentiable. Let $u\colon \mathbb{R} \to \mathbb{R}$ be

$$u = f \circ g = f(g(t))$$

Then u is differentiable with $u'(t) = f'(g(t)) \cdot g'(t)$.

13.2 Theorem. *$A \subseteq \mathbb{R}^n$ open, $f \in C^1(A, \mathbb{R})$. Let $I \subseteq \mathbb{R}$ be an open interval and let $\gamma\colon I \to \mathbb{R}^n$ be a differentiable path such that $\gamma(t) \in A$ for all $t \in I$. Define $u\colon I \to \mathbb{R}$ by $u(t) = f(\gamma(t))$ Then u is differentiable with*

$$u'(t) = \langle (\nabla f)(\gamma(t)) \,,\, \gamma'(t) \rangle$$

13.3 Example. Let $A = (0, \infty)^2$. $f\colon A \to \mathbb{R}$ be $f((s,t)) = s^t$. Then $f \in C^1(A, \mathbb{R})$.

Let $I = (0, \infty) \subseteq \mathbb{R}$ and let $\gamma\colon I \to \mathbb{R}^n$ be $\gamma(t) = (t,t)$. Then $\gamma'(t) = (1,1)$

Then $u(t) = t^t$.

Theorem 13.2 says that

$$\begin{aligned}(t^t)' &= u'(t) \\ &= \langle (\nabla f)(\gamma(t)) \,,\, \gamma'(t) \rangle \\ &= \langle (\nabla f)((t,t)) \,,\, (1,1) \rangle \\ &= \langle (t \cdot t^{t-1}, t^t \ln t) \,,\, (1,1) \rangle \\ &= t^t + t^t \ln t\end{aligned}$$

Proof of Theorem 13.2. Fix $t_0 \in I$ for which we will prove that the Chain Rule holds. So we need

$$\lim_{t \to t_0} \frac{u(t) - u(t_0)}{t - t_0} = \langle (\nabla f)(\gamma(t_0)) \,,\, \gamma'(t_0) \rangle$$

We will do this limit by sequence. Let $(t_k)_{k=1}^{\infty}$ in I such that $t_k \to t_0$. Will show that

$$\lim_{k \to \infty} \frac{u(t_k) - u(t_0)}{t_k - t_0} = \langle (\nabla f)(\gamma(t_0)) \,,\, \gamma'(t_0) \rangle$$

Denote $\gamma(t_0) = \vec{a} \in A$, $\gamma(t_k) = \vec{t}_k \in A, \forall k \in \mathbb{N}$. Then $(\vec{x}_k)_{k=1}^{\infty}$ is a sequence in A.

Claim 1. We have $\vec{x}_k \to \vec{a}$, and moreover that

$$\lim_{k \to \infty} \frac{1}{t_k - t_0} (\vec{x}_k - \vec{a}) = \gamma'(t_0)$$

Verif of Claim 1. For every $k \in \mathbb{N}$ we have

$$\vec{x}_k = \gamma(t_k) = (\gamma^{(1)}(t_k), \cdots, \gamma^{(n)}(t_k))$$

where $\gamma^{(1)}, \cdots \gamma^{(n)}\colon I \to \mathbb{R}$ are differentiable, hence continuous.

When $k \to \infty$, get $\gamma^{(i)}(t_k) \to \gamma^{(i)}(t_0)$. So $\vec{x}_k \to (\gamma^{(1)}(t_0), \cdots, \gamma^{(n)}(t_0))$. Hence $\vec{x}_k \to \vec{a}$ as needed.

Moreover,

$$\frac{1}{t_k - t_0}(\vec{x}_k - \vec{a}) = (\frac{\gamma^{(1)}(t_k) - \gamma^{(1)}(t_0)}{t_k - t_0}, \cdots, \frac{\gamma^{(n)}(t_k) - \gamma^{(n)}(t_0)}{t_k - t_0}) \to ((\gamma^{(1)})'(t_0), \cdots, (\gamma^{(n)})'(t_0)) = \gamma'(t_0)$$

Claim 2. Pick $r > 0$ such that $\mathrm{B}(\vec{a}, r) \subseteq A$, and pick $k_0 \in \mathbb{N}$ such that $\vec{x}_k \in \mathrm{B}(\vec{a}, r)$ for all $k \geq k_0$. Then for every $k \geq k_0$ have $\mathrm{Co}(\vec{a}, \vec{x}_k) \subseteq A$, and we can find $\vec{b}_k \in \mathrm{Co}(\vec{a}, \vec{x}_k)$ such that

$$\frac{u(t_k) - u(t_0)}{t_k - t_0} = \langle (\nabla f)(\vec{b}_k)\, , \, \frac{1}{t_k - t_0}(\vec{x}_k - \vec{a})\, \rangle$$

Verif of Claim 2. Application of MVT.

Claim 3. Let $(\vec{b}_k)_{k=0}^{\infty}$ be as in Claim 2. Then $\vec{v}_k \to \vec{a}$, and therefore $(\nabla f)(\vec{b}_k) \to (\nabla f)(\vec{a})$.

Verif of Claim 3. For every $k \geq k_0$ we have $\vec{b}_k \in \mathrm{Co}(\vec{a}, \vec{x}_k)$, have $||\, \vec{b}_k - \vec{a}\, || \leq ||\, \vec{x}_k - \vec{a}\, ||$. By squeeze theorem $\vec{b}_k \to \vec{a}$. Then for every $1 \leq i \leq n$ get $(\partial_i f)(\vec{b}_k) \to (\partial_i f)(\vec{a})$ because $\partial_i f$ is continuous on A. Then $(\nabla f)(\vec{b}_k) \to (\nabla f)(\vec{a})$.

Claim 4. We have

$$\lim_{k \to \infty} \frac{u(t_k) - u(t_0)}{t_k - t_0} = \langle (\nabla f)(\gamma(t_0))\, , \, \gamma'(t_0)\, \rangle$$

Verif of Claim 4. Have $(\nabla f)(\vec{b}_k) \to (\nabla f)(\vec{a})$

$$\lim_{k \to \infty} \frac{1}{t_k - t_0}(\vec{x}_k - \vec{a}) = \gamma'(t_0)$$

So

$$\langle (\nabla f)(\vec{b}_k)\, , \, \frac{1}{t_k - t_0}(\vec{x}_k - \vec{a})\, \rangle \to \langle (\nabla f)(\vec{a})\, , \, \gamma'(t)\, \rangle$$

Since

$$\frac{u(t_k) - u(t_0)}{t_k - t_0} = \langle (\nabla f)(\vec{b}_k)\, , \, \frac{1}{t_k - t_0}(\vec{x}_k - \vec{a})\, \rangle$$

,

then

$$\lim_{k \to \infty} \frac{u(t_k) - u(t_0)}{t_k - t_0} = \langle (\nabla f)(\gamma(t_0))\, , \, \gamma'(t_0)\, \rangle$$

□

Lecture 14, June. 19

Jacobian matrix, and the Chain Rule in the general case

14.1 Definition. $A \subseteq \mathbb{R}^n$ open, $f\colon A \to \mathbb{R}^m$. For every $\vec{x} \in A$, write $f(\vec{x}) = (f^{(1)}(x), \cdots, f^{(m)}(x)) \in \mathbb{R}^m$, and look at the component functions $f^{(i)}\colon A \to \mathbb{R}$, $1 \le i \le m$. If $f^{(i)} \in C^1(A, \mathbb{R})$. then we say that $f \in C^1(A, \mathbb{R}^m)$.

14.2 Definition. $A \subseteq \mathbb{R}^n$, $f \in C^1(A, \mathbb{R}^m)$ with component functions $f^{(1)}(x), \cdots, f^{(m)}(x) \in C^1(A, \mathbb{R})$. Then for every $\vec{a} \in A$, the matrix

$$(\mathrm{J}f)(\vec{a}) = \begin{bmatrix} (\partial_1 f^{(1)})(\vec{a}) & \cdots & (\partial_n f^{(1)})(\vec{a}) \\ \vdots & & \\ (\partial_1 f^{(m)})(\vec{a}) & \cdots & (\partial_n f^{(m)})(\vec{a}) \end{bmatrix} = \begin{bmatrix} (\nabla f^{(1)})(\vec{a}) \\ \vdots \\ (\nabla f^{(m)})(\vec{a}) \end{bmatrix}$$

is called the Jacobian matrix of f at $\vec{a}$.

14.3 Remark. When $m = 1$, $f \in C^1(A, \mathbb{R})$, so $(\mathrm{J}f)(\vec{a}) \in M_{1\times m}(\mathbb{R})$. $(\mathrm{J}f)(\vec{a})$ is $(\nabla f)(\vec{a})$, treated as a row matrix.

When $n = 1$. Take $A = I =$ open interval in $\mathbb{R}$. and take $f\colon I \to \mathbb{R}^m$ with $f^{(1)}(x), \cdots, f^{(m)}(x)\colon I \to \mathbb{R}$. Have

$$f \in C^1(A, \mathbb{R}^m) \iff f^i \in C^1(I, \mathbb{R}),\ \forall 1 \le i \le m$$

$$(\mathrm{J}f)(a) = \begin{bmatrix} (f^{(1)})'(a) \\ \vdots \\ (f^{(m)})'(a) \end{bmatrix}$$

Hence $(\mathrm{J}f)(a)$ is the velocity vector of f at a, treated as a column matrix.

14.4 Theorem (Chain Rule)**.** *$A \subseteq \mathbb{R}^n$ open, $f \in C^1(A, \mathbb{R}^n)$. Suppose we also have $B \subseteq \mathbb{R}^p$ open, and $g \in C^1(B, \mathbb{R}^n)$, where $g(\vec{b}) \in A$ for all $\vec{b} \in B$.*

Define the composed function $u = f \circ g$.

Then

1. *$u \in C^1(B, \mathbb{R}^m)$*
2. *Fix $\vec{b} \in B$, denote $g(\vec{b}) = \vec{a} \in A$ and consider the Jacobian matrices:*

$$(\mathrm{J}f)(\vec{a}) \in M_{m\times n}(\mathbb{R})$$

$$(\mathrm{J}g)(\vec{b}) \in M_{n\times p}(\mathbb{R})$$

$$(\mathrm{J}u)(\vec{b}) \in M_{m\times p}(\mathbb{R})$$

Then we have

$$(CR) \qquad (\mathrm{J}u)(\vec{b}) = (\mathrm{J}f)(\vec{a}) \cdot (\mathrm{J}g)(\vec{b})$$

14.5 Remark. Can write (CR) as

$$(\mathrm{J}(f\circ g))(\vec{b}) = (\mathrm{J}f)(g(\vec{b}))\cdot(\mathrm{J}g)(\vec{b})$$

This is analogous to chain rule in calc 1

$$(f\circ g)'(b) = f'(g(b))\cdot g'(b)$$

with appropriate dictionary

$$\text{Derivatives} \iff \text{Jacobian matrices}$$

$$\text{multiplication of numbers} \iff \text{multiplication of matrices}$$

Consider the special case $1\to n\to 1$. So $B\subseteq\mathbb{R}, g\colon B\to\mathbb{R}^n$ is a C^1-path and have

$$(\mathrm{J}g)(b) = \begin{bmatrix} (g^{(1)})'(b) \\ \vdots \\ (g^{(m)})'(b) \end{bmatrix}$$

On the other hand, since $m=1$, have $f\in C^1(A,\mathbb{R})$ and $(\mathrm{J}f)(\vec{a}) = [(\partial_1 f)(\vec{a}),\cdots,(\partial_n f)(\vec{a})]$.

Put $u = f\circ g$. Denote $g(b) = a$.

Theorem 13.2 says

$$u'(b) = \langle(\nabla f)(\vec{a})\,,\,g'(b)\,\rangle$$

Theorem 14.4 says

$$(\mathrm{J}u)(b) = (\mathrm{J}f)(\vec{a})\cdot(\mathrm{J}g)(b)$$

14.6 Example ("Polar Coordinates in $\mathbb{R}^2$")**.** $p=2$, $n=2$, $B=\{(r,\theta)\mid 0<r<1,\, 0<\theta<2\pi\}\subseteq\mathbb{R}^2$, $A=\{\vec{x}\in\mathbb{R}^2\mid ||\,\vec{x}\,||<1\}\subseteq\mathbb{R}^2$. Let $g\colon B\to A$, $g((r,\theta)) = (r\cdot\cos\theta, r\cdot\sin\theta)$.

The Jacobian matrix for g

$$(\mathrm{J}g)((r,\theta)) = \begin{bmatrix} \cos\theta & -r\cdot\sin\theta \\ \sin\theta & r\cdot\cos\theta \end{bmatrix}$$

Now take $f\in C^1(A,\mathbb{R})$, say $f((x,y)) = (x^2+y^2)^{3/2}$. Consider the composed function $u=f\circ g$, $u\colon B\to\mathbb{R}$.

(CR) says that for $\vec{b}\in B$ we have

$$(\mathrm{J}u)(\vec{b}) = (\mathrm{J}f)(g(\vec{b}))\cdot(\mathrm{J}g)(\vec{b})$$

$$\begin{aligned}
&[(\partial_1 u)(r,\theta),(\partial_2 u)(r,\theta)] \\
&= [(\partial_1 f)(r\cdot\cos\theta, r\cdot\sin\theta),(\partial_2 f)(r\cdot\cos\theta, r\cdot\sin\theta)]\cdot\begin{bmatrix} \cos\theta & -r\cdot\sin\theta \\ \sin\theta & r\cdot\cos\theta \end{bmatrix} \\
&= [3r^2\cdot\cos\theta, 3r^2\cdot\sin\theta]\cdot\begin{bmatrix} \cos\theta & -r\cdot\sin\theta \\ \sin\theta & r\cdot\cos\theta \end{bmatrix} \\
&= [3r^2, 0]
\end{aligned}$$

14.7 Remark. $\vec{b}\in B$ is denoted as (r,θ), so instead of $\partial_1 g$, $\partial_2 g$, people use $\dfrac{\partial g}{\partial r}$, $\dfrac{\partial g}{\partial\theta}$

Lecture 15, May. 23

C^2-functions and the 2nd derivative test

15.1 Definition. $A \subseteq \mathbb{R}^n$ open, $f \in C^1(A, \mathbb{R})$. Consider the partial derivative $\partial_i f \colon A \to \mathbb{R}$, $1 \leq i \leq n$. If $\partial_i f \in C^1(A, \mathbb{R})$, $\forall\, 1 \leq i \leq n$, we say $f \in C^2(A, \mathbb{R})$.

So in order to have $f \in C^2(A, \mathbb{R})$, we need the second partial derivatives $\partial_j(\partial_i f)$ to exist and be continuous on A $(1 \leq i, j \leq n)$

Convention: omit brackets, just write $\partial_j \partial_i f$. Also if $i = j$, then write $\partial_i^2 f$ instead of $\partial_i \partial_j f$.

15.2 Theorem. *$A \subseteq \mathbb{R}^n$ open, $f \in C^2(A, \mathbb{R})$. Then for every $1 \leq i, j \leq n$ we have $\partial_j \partial_j f = \partial_j \partial_i f$.*

15.3 Definition. $A \subseteq \mathbb{R}^n$ open, $f \in C^2(A, \mathbb{R})$, $\vec{a} \in A$. The $n \times n$ matrix

$$[\mathrm{H}f](\vec{a}) = \begin{bmatrix} (\partial_1^2 f)(\vec{a}) & (\partial_1 \partial_2 f)(\vec{a}) & \cdots & (\partial_1 \partial_n f)(\vec{a}) \\ \cdots & & & \\ (\partial_n^2 f)(\vec{a}) & (\partial_n \partial_2 f)(\vec{a}) & \cdots & (\partial_n^2 f)(\vec{a}) \end{bmatrix}$$

is called the Hessian matrix of f at $\vec{a}$. Theorem 15.2 says that $\mathrm{H} = \mathrm{H}^T$

15.4 Remark. Let $\lambda_1, \cdots, \lambda_n$ be the eigenvalues of H. Important points

1. Can diagonalize: $\mathrm{H} = VDV^T$ where D is a diagonal matrix and V is an orthogonal matrix.
2. If $\lambda_1, \cdots, \lambda_n > 0$, then we say that H is positive definite. Equivalent description for this property is $\langle \mathrm{H} \cdot \vec{v}\,, \vec{v}\,\rangle > 0$ for all $\vec{v} \in \mathbb{R}^n$, $\vec{v} \neq \vec{0}$. Likewise for negative definite.

15.5 Remark. $I \subseteq \mathbb{R}$ open interval, $f \colon I \to \mathbb{R}$ twice differentiable. Suppose that $f'(a) = 0$, $f''(a) \neq 0$, then $f''(a) > 0$ then a is local minimum for f and if $f''(a) < 0$ then a is local maximum for f.

15.6 Theorem. *$A \subseteq \mathbb{R}^n$ open, $f \in C^2(A, \mathbb{R})$, $\vec{a} \in A$. Consider the gradient vector $(\nabla f)(\vec{a}) \in \mathbb{R}^n$ and the Hessian matrix $[\mathrm{H}f](\vec{a}) \in M_{n \times n}(\mathbb{R})$ with eigenvector $\lambda_1, \cdots, \lambda_n \in \mathbb{R}$. Suppose that $(\nabla f)(\vec{a}) = \vec{0}$. ($\vec{a}$ is stationary point for f), and $\lambda_1, \cdots, \lambda_n \neq 0$. Then*

1. *$\lambda_1, \cdots, \lambda_n > 0$ ($[\mathrm{H}f](\vec{a})$ is positive def) then $\vec{a}$ is local min for f*
2. *$\lambda_1, \cdots, \lambda_n < 0$ ($[\mathrm{H}f](\vec{a})$ is negative def) then $\vec{a}$ is local max for f*
3. *If $\exists i, j$ such that $\lambda_i > 0$, $\lambda_j < 0$ then $\vec{a}$ is not a local extremum for f.*

15.7 Example. $n = 2$, $f \in C^2(\mathbb{R}^2, \mathbb{R})$. $f((s,t)) = s^3 + t^3 - 3st$, $s, t \in \mathbb{R}$.

15.8 Remark. $A \subseteq \mathbb{R}^2$ open, $f \in C^2(A, \mathbb{R})$, $\vec{a} \in A$ with $(\nabla f)(\vec{a}) = \vec{0}$. Denote $[\mathrm{H}f](\vec{a}) = \begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix}$ and λ_1, λ_2 be the eigenvector of $\begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix}$.

Linear algebra facts

$$\lambda_1 + \lambda_2 = \alpha + \gamma$$

$$\lambda_1 \cdot \lambda_2 = \alpha \cdot \gamma - \beta_2$$

2nd derivative test gives this

1. if $\alpha\beta < \beta^2$, then $\vec{a}$ is saddle point for f.
2. if $\alpha\beta > \beta^2$ and $\alpha > 0$, then $\vec{a}$ is local min for f
3. if $\alpha\beta > \beta^2$ and $\alpha < 0$, then $\vec{a}$ is local min for f

Lecture 16, May. 26

Basic integrability I: divisions and lower/upper sums

16.1 Definition. Let A be a non-empty set in $\mathcal{M}_n$. By a division of A we mean a set $\Delta = \{A_1, \cdots, A_r\}$ where $A_1, \cdots, A_r$ are non-empty sets in $\mathcal{M}_n$, such that $A_1 \cup \cdots \cup A_r = A$ and $A_i \cap A_j = \emptyset$ for $i \neq j$.

16.2 Definition. $A \neq \emptyset$ in $\mathcal{M}_n$. Let $\Delta = \{A_1, \cdots, A_r\}$ and $\Gamma = \{B_1, \cdots, B_q\}$ be two divisions of A. We say that Γ refines Δ written as $\Gamma \prec \Delta$ to mean that for every $1 \leq j \leq q$ there exists $1 \leq i \leq r$ such that $B_j \subseteq A_i$.

16.3 Remark. Condition from Definition 16.2 forces every A_i of Δ to be a disjoint union of B_j's from Γ. That is, we can rewrite $\Delta = \{A_1, \cdots, A_r\}$,

$$\Gamma = \{B_{1,1}, \cdots, B_{1,q_1}, \cdots B_{r,1}, \cdots, B_{r,q_r}\}$$

where $q = q_1 + \cdots + q_r$.

16.4 Proposition. *$A \neq \emptyset$ in $\mathcal{M}_n$. Let Δ' and Δ'' be two divisions of A. Then exists division Γ of A such that $\Gamma \prec \Delta'$ and $\Gamma \prec \Delta''$.*

Proof. Write $\Delta' = \{A_1', \cdots, A_r'\}$, $\Delta'' = \{A_1'', \cdots, A_s''\}$.

Put $\Gamma = \{A_i' \cap A_j'' \mid 1 \leq i \leq r, 1 \leq j \leq s, A_i' \cup A_j'' \neq \emptyset\}$ $\square$

Now start looking at functions defined on sets $A \in \mathcal{M}_n$. Need to review: oscillation.

Let $f \colon A \to \mathbb{R}$. Suppose f is bounded. For $B \subseteq A$, we denote

$$\sup_B(f) = \sup\{f(\vec{x}) \mid \vec{x} \in B\}$$

$$\inf_B(f) = \inf\{f(\vec{x}) \mid \vec{x} \in B\}$$

Then we put

$$\operatorname*{osc}_B(f) = \sup\{\mid f(\vec{x}) - f(\vec{y}) \mid \mid \vec{x}, \vec{y} \in B\}$$

16.5 Definition. $f \colon A \to \mathbb{R}$ bounded function where $A \in \mathcal{M}_n$. Let $\Delta = \{A_1, \cdots, A_r\}$ be a division of A. Then

$$U(f, \Delta) = \sum_{i=1}^{r} \mathrm{Vol}(A_i) \cdot \sup_{A_i}(f)$$

is called the upper (Darboux) sum for f and Δ and

$$L(f, \Delta) = \sum_{i=1}^{r} \mathrm{Vol}(A_i) \cdot \inf_{A_i}(f)$$

is called the lower (Darboux) sum for f and Δ.

16.6 Remark. A, f, Δ as in Definition 16.5, have

$$U(f,\Delta) - L(f,\Delta) = \sum_{i=1}^{r} \mathrm{Vol}(A_i) \cdot \underset{A_i}{\mathrm{osc}}(f)$$

In particular, it is obvious that $U(f,\Delta) \geq L(f,\Delta)$.

16.7 Lemma. *$A \in \mathcal{M}_n$, $f\colon A \to \mathbb{R}$ bounded. Let Δ, Γ be divisions of A such that $\Gamma \prec \Delta$. Then we have $U(f,\Gamma) \leq U(F,\Delta)$ and $L(f,\Gamma) \geq L(f,\Delta)$.*

Proof. Will prove the inequality for upper sums. Let $\Delta = \{A_1, \cdots, A_r\}$. $\Gamma = \{B_{1,1}, \cdots B_{1,q_1}, \cdots, B_{r,1}, \cdots B_{r,q_r}\}$ where $B_{i,1} \cup \cdots \cup B_{i,q_i} = A_i$, $1 \leq i$
$leqr$.

Then

$$U(f,\Gamma) = \sum_{i=1}^{r} \left(\sum_{j=1}^{q_i} \mathrm{Vol}(B_{i,j}) \cdot \sup_{b_{i,j}}(f) \right) \leq \sum_{i=1}^{r} \left(\sum_{j=1}^{q_i} \mathrm{Vol}(B_{i,j}) \right) \cdot \sup_{A_i}(f) = U(f,\Delta).$$

$\square$

16.8 Proposition. *$A \in \mathcal{M}_n$, $f\colon A \to \mathbb{R}$. Let Δ', Δ'' be two divisions of A. Then $U(f,\Delta') \geq L(f,\Delta'')$.*

Lecture 17, June. 30

Basic integrability II: the integral.

17.1 Definition. $A \in \mathcal{M}_n$. $f\colon A \to \mathbb{R}$ bounded. The set of real numbers

$$T = \{U(f, \Delta) \mid \Delta \text{ division of } A\}$$

is bounded from below, so has an inf.

The number $\inf(T) \in \mathbb{R}$ is called the upper integral of f on Am denoted as $\overline{\int}_A f$ or $\overline{\int}_A f(\vec{x})\mathrm{d}\vec{x}$.

The set of real numbers

$$S = \{L(f, \Delta) \mid \Delta \text{ division of } A\}$$

is bounded from above, so has an sup.

The number $\sup(S) \in \mathbb{R}$ is called the lower integral of f on Am denoted as $\underline{\int}_A f$ or $\underline{\int}_A f(\vec{x})\mathrm{d}\vec{x}$.

One has $\underline{\int}_A f \leq \overline{\int}_A f$

17.2 Definition. If $\underline{\int}_A f = \overline{\int}_A f$, then we say that f is integrable on A, and define its integral to be $\underline{\int}_A f = \int_A f = \overline{\int}_A f$

17.3 Theorem. *$A \in \mathcal{M}_n$, $f\colon A \to \mathbb{R}$ bounded. Then TFAE*

1. *f is integrable on A.*
2. *for every $\epsilon > 0$ there exists a division Δ of A such that $U(f, \Delta) - L(f, \Delta) < \epsilon$.*
3. *There exists a sequence $(\Delta_k)_{k=1}^{\infty}$ of divisions of A such that $U(f, \Delta_k) - L(f, \Delta_k) \to 0$.*

Proof. Will prove (1) $\to$ (2). Others are left as exercises.

Denote $\int_A f = I$. So have $\underline{\int}_A f = I = \overline{\int}_A f$. Given $\epsilon > 0$, we need to find a division Δ of A such that $U(f, \Delta) - L(f, \Delta) < \epsilon$.

The idea is to find Δ' such that $I \leq U(f, \Delta') < I + \epsilon/2$. Find Δ'' such that $I - \epsilon/2 < L(f, \Delta'') \leq I$. Then let $\Delta \prec \Delta'$ and $\Delta \prec \Delta''$. Then we find such Δ. $\square$