Chapter 1. Topological Properties of Sets in Euclidean Space

2019年5月7日 18:49

1.1 Definition:
For vectors $x, y \in \mathbb{R}^n$ we define the **dot product** of x and y to be

$$x \cdot y = y^T x = \sum_{i=1}^{n} x_i y_i .$$

1.2 Theorem: (Properties of the Dot Product) For all $x, y, z \in \mathbb{R}^n$ and all $t \in \mathbb{R}$, we have

1. (Bilinearity)
   $(x+y)\cdot z = x \cdot z + y \cdot z, (tx)\cdot y = t(x \cdot y)$
   $x \cdot (y+z) = x \cdot y + x \cdot z, x \cdot (ty) = t(x \cdot y)$
2. (Symmetry)
   $x \cdot y = y \cdot x$, and
3. (Positive Definiteness) $x \cdot x \geq 0$ with $x \cdot x = 0$ if and only if $x = 0$.

1.3 Definition: For a vector $x \in \mathbb{R}^n$, we define the **norm** (or **length**) of x to be

$$|x| = \sqrt{x \cdot x} = \sqrt{\sum_{i=1}^{n} x_i^2}$$

We say that x is a **unit vector** when $|x| = 1$.

1.4 Theorem: (Properties of Length) Let $x, y \in \mathbb{R}^n$ and let $t \in \mathbb{R}$. Then

1. (Positive Definiteness) $|x| \geq 0$ with $|x| = 0$ if and only if $x = 0$,
2. (Scaling) $|tx| = |t||x|$,
3. $|x \pm y|^2 = |x|^2 \pm 2(x \cdot y) + |y|^2$.
4. (The Polarization identities) $x \cdot y = \frac{1}{2}\left(|x+y|^2 - |x|^2 - |y|^2\right) = \frac{1}{4}\left(|x+y|^2 - |x-y|^2\right)$,
5. (The Cauchy-Schwarz inequality) $|x \cdot y| \leq |x||y|$ with $|x \cdot y| = |x||y|$ if and only if the set $\{x, y\}$ is linearly dependent, and
6. (The Triangle Inequality) $|x + y| \leq |x| + |y|$

Proof: We leave the proofs of Parts (1), (2) and (3) as an exercise, and we note that (4) follows immediately from (3). To prove part (5), suppose first that $\{x, y\}$ is linearly dependent. Then one of x and y is a multiple of the other, say $y = tx$ with $t \in \mathbb{R}$.

$|x \cdot y| = |x \cdot (tx)| = |t(x \cdot x)| = |t||x|^2 = |x||tx| = |x||y|$.

Suppose next that $\{x, y\}$ is linearly independent. Then for all $t \in \mathbb{R}$ we have $x + ty \neq 0$ and so

$0 \neq |x + ty|^2 = (x + ty)\cdot(x + ty) = |x|^2 + 2t(x \cdot y) + t^2|y|^2$.

Since the quadratic on the right is non-zero for all $t \in \mathbb{R}$, it follows that the discriminant of the quadratic must be negative, that is

$4(x \cdot y)^2 - 4|x|^2|y|^2 < 0$

Thus $(x \cdot y)^2 < |x|^2|y|^2$ and hence $|x \cdot y| < |x||y|$. This proves part (5).

Using part (5) note that

$|x + y|^2 = |x|^2 + 2(x \cdot y) + |y|^2 \leq |x + y|^2 + 2|x \cdot y| + |y|^2 \leq |x|^2 + 2|x||y| + |y|^2 = (|x| + |y|)^2$

And so $|x + y| \leq |x| + |y|$, which proves part (6).

1.5 Definition:
For points $a, b \in \mathbb{R}^n$, we define the **distance** between a and b to be
$\mathrm{dist}(a, b) = |b - a|$.

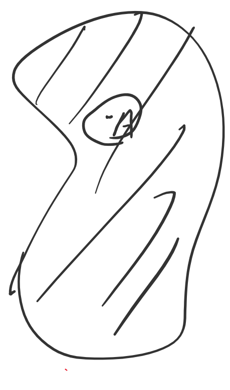
1.6 Theorem: *(Properties of Distance)*
Let $a, b, c \in \mathbb{R}^n$. Then

1. *(Positive Definiteness)* $\mathrm{dist}(a, b) \geq 0$ with $\mathrm{dist}(a, b) = 0$ if and only if $a = b$,
2. *(Symmetry)* $\mathrm{dist}(a, b) = \mathrm{dist}(b, a)$, and
3. *(The Triangle Inequality)* $\mathrm{dist}(a, c) \leq \mathrm{dist}(a, b) + \mathrm{dist}(b, c)$.

Proof: The proof is left as an exercise.

1.7 Definition:
For nonzero vectors $0 \neq u, v \in \mathbb{R}^n$, we define the **angle between** u and v to be $\theta(u, v) = \cos^{-1}\frac{u \cdot v}{|u||v|} \in [0, \pi]$. We say that u and v are **orthogonal** when $u \cdot v = 0$. As an exercise, determine (with proof) some properties of angles.

1. Symmetric
2. Scaling does not change the angle
3. Law of Cosine works
4. Angles add when on the same plane (Difficult)



1.8 Definition:
For $a \in \mathbb{R}^n$ and $0 < r \in \mathbb{R}$, the **sphere**, the **open ball**, the **closed ball**, and the (open) **punctured ball** in $\mathbb{R}^n$ centered at a of radius r are defined to be the sets

$S(a, r) = \{x \in \mathbb{R}^n | \mathrm{dist}(x, a) = r\} = \{x \in \mathbb{R}^n | |a - x| = r\}$,
$B(a, r) = \{x \in \mathbb{R}^n | \mathrm{dist}(x, a) < r\} = \{x \in \mathbb{R}^n | |a - x| < r\}$,
$\overline{B}(a, r) = \{x \in \mathbb{R}^n | \mathrm{dist}(x, a) \leq r\} = \{x \in \mathbb{R}^n | |a - x| \leq r\}$,
$B^*(a, r) = \{x \in \mathbb{R}^n | 0 < \mathrm{dist}(x, a) < r\} = \{x \in \mathbb{R}^n | 0 < |a - x| < r\} = B(a, r) \setminus \{a\}$

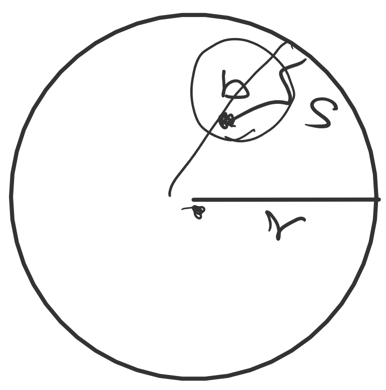
1.9 Definition:
Let $A \subseteq \mathbb{R}^n$. We say that A is **bounded** when $A \subseteq B(a, r)$ for some $a \in \mathbb{R}^n$ and some $0 < r \in \mathbb{R}$. As an exercise, verify that A is bounded if and only if $A \subseteq B(0, r)$ for some $r > 0$.

Bounded also $A \subseteq \overline{B}(0, r)$ for some $r > 0$
$\leftrightarrow$ for some $r > 0$, we have $|a| < r$ for all $a \in A$.

1.10 Definition:
For a set $A \subseteq \mathbb{R}^n$, we say that A is **open** (in $\mathbb{R}^n$) when for every $a \in A$ there exists $r > 0$ such that $B(a, r) \subseteq A$, and we say that A is **closed** (in $\mathbb{R}^n$) when its complement $A^c = \mathbb{R}^n \setminus A$ is open in $\mathbb{R}^n$.

1.11 Exercise:
Show that open intervals in $\mathbb{R}$ are open in $\mathbb{R}$ and closed intervals in $\mathbb{R}$ are closed in $\mathbb{R}$.

1.12 Example:
Show that for $a \in \mathbb{R}^n$ and $0 < r \in \mathbb{R}$, the set $B(a, r)$ is open and the set $\overline{B}(a, r)$ is closed.



Solution:
Let $a \in \mathbb{R}^n$ and let $r > 0$. We claim that $B(a, r)$ is open. We need to show that for all $b \in B(a, r)$ there exists $s > 0$ such that $B(b, s) \subseteq B(a, r)$. Let $b \in B(a, r)$ and note $|b - a| < r$. (Definition of in the ball)
Let $s = r - |b - a| > 0$
Let $x \in B(b, s)$ so $|x - b| < s$
Then $|x - a| = |x - b + b - a|$
$\leq |x - b| + |b - a|$ By the Triangle Inequality
$< s + |b + a|$
$= r - |b - a| + |b - a|$
$= r$

Since $|x - a| < r$, we have $x \in B(a, r)$.

Thus, $B(b,s) \subseteq B(a,r)$.
It follows that $B(a,r)$ is open.

Next we claim that $\bar{B}(a,r)$ is closed.
We need to show that $\bar{B}(a,r)^c$ is open (in $\mathbb{R}^n$).

Let $b \in \bar{B}(a,r)^c$ so $b \in \mathbb{R}^n, b \notin \bar{B}(a,r)$
Since $b \notin \bar{B}(a,r)$, we have $|b-a| > r$.
Let $x \in B(b,s)$ so $x - b| < s$.
Then

$|a-b| = |a-x+x-b|$
$\leq |a-x| + |x-b|$

$|a-x| \geq |a-b| - |x-b|$
$> |a-b| - s$
$= |a-b| - (|b-a| - r) = r$

Since $|x-a| > r$
We have $x \notin \bar{B}(a,r)$
So $x \in \bar{B}(a,r)$.
Thus $B(b,s) \subseteq \bar{B}(a,r)^c$

This proves that $\bar{B}(a,r)^c$ is open
Hence $\bar{B}(a,r)$is closed (in $\mathbb{R}^n$

Definition:
Let $A \subseteq \mathbb{R}^n$
The interior of A (in $\mathbb{R}^n$)
Is the set
~~$A^o = \{a \in A| r > 0$~~

$A^0 = \cup_{U \in S} U$
Where S is the set of all open sets U in $\mathbb{R}^n$ with $U \subseteq A$
And the closure of As the set
$\bar{A} = \cap_{k \in T} K$
Where T is the set of all closed sets K in $\mathbb{R}^n$with $A \subseteq K$

1.13 Theorem
(Basic Properties of Open Sets)
1. $\emptyset$ and $\mathbb{R}^n$ are open in $\mathbb{R}^n$
2. If U_k is an open set for each $k \in K$ (where K is any set) then $\cup_{k \in K} U_k$ is open
3. If $U_1, U_2, \ldots, U_l$ are open sets in $\mathbb{R}^n$, then $\cap_{k=1}^{l} U_k$ is open (where $l \in \mathbb{Z}^+$)

Proof:
1. Vacuously True for $\emptyset$ and $\mathbb{R}^n$ is open because $B(a,r) = \{x \in \mathbb{R}^n | |x-a| < r\} \subseteq \mathbb{R}^n$ for all $a \in \mathbb{R}^n, r > 0$
2. Let U_k be open for each $k \in K$
   Let $U = \cup_{k \in K} U_k$
   Let $a \in U$
   Since $a \in U = \cup_{k \in K} U_k$
   We can choose an index k so
   That $a \in U_k$
   Since U_k is open
   We can choose $r > 0$
   So that $B(a,r) \subseteq U_k$
   Since $B(a,r) \subseteq U_k$
   And $U_k \subseteq U_{j \in K} U_j = U$
   We have $B(a,r) \subseteq U$
   Thus, U is open
3. Let S be a finite set of open sets. If $S = \emptyset$ then we use the convention that $\bigcap S = \mathbb{R}^n$, which is open. Suppose that $S \neq \emptyset$, says $S = \{U_1, U_2, \ldots, U_m\}$ where each U_k is an open set. Let $a \in \bigcap S = \bigcap_{k=1}^{m} U_k$. For each index k, since $a \in U_k$ we can choose $r_k > 0$ so that $B(a, r_k) \subseteq U_k$. Let $r = \min\{r_1, r_2, \ldots, r_m\}$. Then for each index k, it follows that $B(a,r) \subseteq \bigcap_{k=1}^{m} U_k = \bigcap S$. Thus $\bigcap S$ is open, as required.

   Don't really need to consider the vacuous truth. Proof method does not require this.

   If infinitely many sets, countably (inf) could be zero
   Inf could be zero

   But it's false if not finite.

   On class version:
   Let $U_1, U_2, \ldots, U_l$ be open sets in $\mathbb{R}^n$

1.14 Theorem: (Basic Properties of Closed Sets)
1. The sets $\emptyset$ and $\mathbb{R}^n$ are closed in $\mathbb{R}^n$.
2. If S is a set of closed sets then the intersection $\bigcap S = \bigcap_{K \in S} K$ is closed.
3. If S is a finite set of closed sets then the union $\bigcup S = \bigcup_{K \in S} K$ is closed.

Proof: The proof is left as an exercise
Taking Complements

$$\left(\bigcap_{j \in J} K_j\right)^c = \bigcup_{j \in J} K_j^c$$

De Morgan's Law!!!!

1.15 Definition: Let $A \subseteq \mathbb{R}^n$. The **interior** and the **closure** of A (in $\mathbb{R}^n$) are the sets

$$A^0 = \bigcup \{U \subseteq \mathbb{R} \mid U \text{ is open, and } U \subseteq A\},$$

$$\bar{A} = \bigcap \{K \subseteq \mathbb{R}^n \mid K \text{ is closed and } A \subseteq K\}.$$

1.16 Theorem:
Let $A \subseteq \mathbb{R}^n$.
1. The interior of A is the largest open set which is contained in A. In other words, $A^0 \subseteq A$ and A^0 is open, and for every open set U with $U \subseteq A$ we have $U \subseteq A^0$.
2. The closure of A is the smallest closed set which contains A. In other words, $A \subseteq \bar{A}$ and $\bar{A}$ is closed, and for every closed set K with $A \subseteq K$ we have $\bar{A} \subseteq K$.

Largest / smallest exist.
Ordering
Partial ordering

So does not exist???
No!!!!! See left

For open ball $B(a,r)$, the closure is $\bar{B}(a,r)$.
For closed ball $\bar{B}(a,r)$, the interior is $B(a,r)$.

Proof:
Note that A^0 is open by Part (2) of Theorem 8.10, because A^0 is equal to the union of a set of open sets. Also note that $A^0 \subseteq A$ because A^0 is equal to the union of a set of subsets of A. Finally note that for any open set U with $U \subseteq A$ we have $U \in S$ so that $U \subseteq \cup S = A^0$. This completes the proof of Part (1), and the proof of Part (2) is similar.

The proof is really immediate! Like a Corollary.

1.17 Corollary: Let $A \subseteq \mathbb{R}^n$
1. $\left(A^0\right)^0 = A^0$ and $\bar{\bar{A}} = \bar{A}$
2. A is open if and only if $A = A^0$
3. A is closed if and only if $A = \bar{A}$.

Proof: The proof is left as an exercise.

1.18 Definition:
Let $A \subseteq \mathbb{R}^n$. An **interior point** of A is a point $a \in A$ such that for some $r > 0$ we have $B(a,r) \subseteq A$. (Surrouded)

A **limit point** of A is a point $a \in \mathbb{R}^n$ such that for every $r > 0$ we have $B^*(a,r) \cap A \neq \emptyset$. The set of all limit points of A is denoted by A'. The **boundary** of A, is the set of all boundary points of A.

What's the difference? Intuitive Example?
For disk, they are the same?

A **boundary point** of A when for all $r > 0$
$B(a,r) \cap A \neq \emptyset$, and $B(a,r) \cap A^c \neq \emptyset$

And the set of boundary points of A is denoted by ∂A and ∂A is called the boundary of A.

1.19 Theorem: (Properties of Interior, Limit and Boundary Points)
Let $A \subseteq \mathbb{R}^n$.
1. A^0 is equal to the set of all interior points of A.
2. A is closed if and only if $A' \subseteq A$.
3. $\bar{A} = A \cup A'$.
4. $\partial A = \bar{A} \setminus A^0$. or equivalently $\bar{A} = A^0 \cup \partial A$ and $A^0 \cap \partial A = \emptyset$

Proof: Two side approaches and contrapositive.
We leave the proof of Parts (1) and (4) as exercises. To prove Part (2) note that when $a \notin A$ we have $B(a,r) \cap A = B^*(a,r) \cap A$ and so

A is closed $\Leftrightarrow A^c$ is open
$\Leftrightarrow \forall a \in A^c\ \exists r > 0\ B(a,r) \subseteq A^c$
$\Leftrightarrow \forall a \in \mathbb{R}^n (a \notin A \Rightarrow \exists r > 0\ B(a,r) \subseteq A^c)$
$\Leftrightarrow \forall a \in \mathbb{R}^n (a \notin A \Rightarrow \exists r > 0\ B(a,r) \cap A = \emptyset)$
$\Leftrightarrow \forall a \in \mathbb{R}^n (a \notin A \Rightarrow \exists r > 0\ B^*(a,r) \cap A = \emptyset)$
$\Leftrightarrow \forall a \in \mathbb{R}^n (\forall r > 0\ B^*(a,r) \cap A \neq \emptyset \Rightarrow a \in A)$
$\Leftrightarrow \forall a \in \mathbb{R}^n (a \in A' \Rightarrow a \in A)$
$\Leftrightarrow A' \subseteq A$.

$A' \subseteq A \Leftrightarrow \forall a \in \mathbb{R}^n (a \in A' \Rightarrow a \in A)$
$\Leftrightarrow \forall a \in \mathbb{R}^n (\forall r > 0\ B^*(a,r) \cap A \neq \emptyset \Rightarrow a \in A)$

To prove Part (3) we shall prove that $A \cup A'$ is the smallest closed set which contains A. It is clear that $A \cup A'$ contains A. We claim that $A \cup A'$ is closed, that is $(A \cup A')^c$ is open. Let $a \in (A \cup A')^c$, that is let $a \in \mathbb{R}^n$ with $a \notin A$ and $a \notin A'$. Since $a \notin A'$ we can choose $r > 0$ so that $B(a,r) \cap A = \emptyset$. We claim that because $B(a,r) \cap A = \emptyset$ it follows that $B(a,r) \cap A' = \emptyset$.

(Point b arbitrarity close to a)
Suppose, for a contradiction, that $B(a,r) \cap A' \neq \emptyset$. Choose $b \in B(a,r) \cap A'$. Since $b \in B(a,r)$ and $B(a,r)$ is open, we can choose $s > 0$ so that $B(b,s) \subseteq B(a,r)$. Since $b \in A'$ (b is a limit point) it follows that $B(b,s) \cap A \neq \emptyset$. Choose $x \in B(b,s) \cap A$. Then we have $x \in B(b,s) \subseteq B(a,r)$ and $x \in A$ and so $x \in B(a,r) \cap A$, which contradicts the fact that $B(a,r) \cap A = \emptyset$.

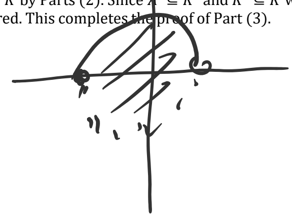
Thus $B(a,r) \cap A' = \emptyset$, as claimed. Since $B(a,r) \cap A = \emptyset$ and $B(a,r) \cap A' = \emptyset$ it follows that $B(a,r) \cap (A \cup A') = \emptyset$ hence $B(a,r) \subseteq (A \cup A')^c$. Thus proves that $(A \cup A')^c$ is open, and hence $A \cup A'$ is closed.

It remains to show that for every closed set K with $A \subseteq K$ we have $A \cup A' \subseteq K$. Let K be a closed set in $\mathbb{R}^n$ with $A \subseteq K$. Note that since $A \subseteq K$ it follows that $A' \subseteq K'$ because if $a \in A'$ then for all $r > 0$ we have $B(a,r) \cap A \neq \emptyset$ hence $B(a,r) \cap K \neq \emptyset$ and so $a \in K'$. Since K is closed we have $K' \subseteq K$ by Parts (2). Since $A' \subseteq K'$ and $K' \subseteq K$ we have $A' \subseteq K$. Since $A \subseteq K$ and $A' \subseteq K$ we have $A \cup A' \subseteq K$, as required. This completes the proof of Part (3).

Limit points visualize?

Example!!

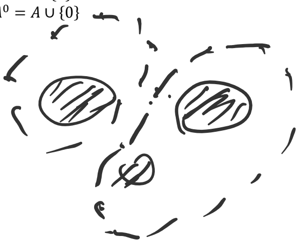
If $A = B(0,1) \cup \{(\cos t, \sin t) |\ 0 \leq t < \pi\} \subseteq \mathbb{R}^2$



$A^0 = B(0,1)$
$\bar{A} = \bar{B}(0,1)$
$\partial A = \bar{A} \setminus A^0$
$= S(0,1)$
$A' = \bar{A}$

Example!!!
When $A = \left\{\frac{1}{n} \middle|\ n \in \mathbb{Z}^+\right\} \subseteq R'$

$A^0 = \emptyset$
$A' = \{0\}$
$\bar{A} = A \cup A' = A \cup \{0\}$
$\partial A = \bar{A} \setminus A^0 = A \cup \{0\}$



Separate using open sets

1.20 Definition:
Let $A \subseteq \mathbb{R}^n$. For sets $U, V \subseteq \mathbb{R}^n$, we say that U and V **separate** A when
$U \cap A \neq \emptyset, V \cap A \neq \emptyset, U \cap V = \emptyset$ and $A \subseteq U \cup V$.
We say that A is **connected** when there do not exist open sets U and V in $\mathbb{R}^n$ which separate A. We say that A is **disconnected** when it is not connected, that is when there do exist open sets U and V in $\mathbb{R}^n$ which separate A.

1.21 Theorem
The connected sets in $\mathbb{R}$ are the intervals, that is the sets of one of the forms
$(a,b), [a,b), (a,b], [a,b], (a,\infty), [a,\infty), (-\infty,b), (-\infty,b], (-\infty,\infty)$
For some $a, b \in \mathbb{R}$ with $a \leq b$. We include the case that $a = b$ in order to include the degenerate intervals $\emptyset =$

(a, a) and $\{a\} = [a, a]$.

Sketch proof:

Supremum arguments to prove the intermediate value property

Munkres Topology

Hoffman

Analysis elementary

Davidson

Real Analy. and Application

Donsig

Springer

Proof:
We use the fact that the intervals in $\mathbb{R}$ are the sets with the intermediate value property (a set $A \subseteq \mathbb{R}$ has **the intermediate value property** when for all $a, b \in A$ and all $x \in \mathbb{R}$, if $a < x < b$ then $x \in A$). Let $A \subseteq \mathbb{R}$. Suppose that A is not an interval. Then A does not have the intermediate value property so we can choose $a, b \in A$ and $u \in \mathbb{R}$ with $a < x < b$. And $x \notin A$. Then $U = (-\infty, x)$ and $V = (x, \infty)$ separate A and so A is disconnected.

Suppose, conversely, that A is disconnnected. Choose open sets U and V which separate A. Choose $a \in U$ and $b \in V$. Note that $a \neq b$ since $U \cap V = \emptyset$. Suppose that

Choose $a \in U \cap A, b \in V \cap A$. Note that they are not equal.

Note that $a \neq b$. Since $U \cap V = \emptyset$.
Say $a < b$. (The case $b < a$ is similar)

Let $S = U \cap [a, b]$
Note that $S \neq \emptyset$ ($a \in S$) and S is bounded above (by b)

S has a supremum in $\mathbb{R}$.

(Least upper bound or completeness)

Let $x = \sup S$

Note that since $S \subseteq [a, b]$ we have $x = \sup S \in [a, b]$

Since $A \in U$ and U is open we can choose $0 < r < b - a$. So that $B(a, r) \subseteq U$. That is $(a - r, a + r) \subseteq U$.

Hence $[a, a + r) \subseteq U \cap [a, b] = S$

And hence $x = \sup S \geq a + r > a$

Since $b \in V$ and V is open, we can choose $0 < r < b - a$
So that $(b - r, b + r) \subseteq V$ and since $U \cap V = \emptyset$, we have $S = U \cap [a, b] \subseteq [a, b - r]$

Hence $x = \sup S \leq b - r < b$

Thus we have $a < x < b$.

If we had $x \in U$
We could choose $r > 0$ with $r \leq \min(x - a, b - x)$ (the radius is small enough) so that $(x - r, x + r) \subseteq U \cap [a, b]$.
but then $x = \sup S \geq x + r > x$ which is impossible

Similar argument shows $x \notin V$

If we had $x \in V$. We could choose $0 < r < \min(x - a, b - x)$. So that $(x - r, x + r) \subseteq V \cap [a, b]$

Core of the proof:
Contradiction
Supremum! Two supremum

But then (since $U \cap V = \emptyset$)

$S = U \cap [a, b] \subseteq [a, x - r] \cup [x + r, b]$

Restriction on the possible value of supremum

If we had $S \subseteq [a, x - r]$ we would have $x = \sup S \leq x - r < r$
And otherwise we would have $x = \sup S \geq x + r > x$

In either case, we have a contradiction, so $x \notin V$.

Since $x \notin U$ and $x \notin V$
And $A \subseteq U \cup V$

It follows that $x \notin A$.

Thus, we have $a, b \in A$ and $x \notin A$ with $a < x < b$,
So A does not satisfy the IVP, so A is not an interval.

If we had $x = a$

Note that $u \neq a$ since we can choose $\delta > 0$ such that $[a, a + \delta) \subseteq U \cap [a, b]$ and we have $u = \sup(U \cap [a, b]) \geq a + \delta$. Note that $u \neq b$ since we can choose $\delta > 0$ such that $(b - \delta, b] \subseteq V \cap [a, b]$ and then we have $u = \sup(U \cap [a, b]) \leq b - \delta$ since $U \cap V = \emptyset$.

which contradicts the fact the $u = \sup(U \cap [a, b])$ because $U \cap V = \emptyset$.

And note that $x \in [a, b]$.

(Show that $a < x < b$ and $x \notin U$ and $x \notin V$ so $x \notin A \subseteq U \cup V$)

The ball here is neither open nor closed
Considered open in P

(Show that $a < x < b$ and $x \notin U$ and $x \notin V$ so $x \notin A \subseteq U \cup V$)

Open and Closed sets in $P \subseteq \mathbb{R}^n$

Definition: When $P \subseteq \mathbb{R}^n, a \in P$ and $0 < r \in \mathbb{R}$,
The open ball in P centered at a radius of r is the set
$B_P(a,r) = \{x \in P \mid |x-a| < r\} = B(a,r) \cap P$

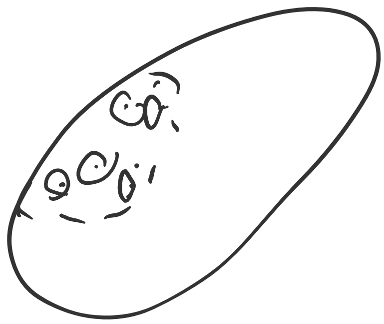
Similarly, we define $\bar{B}_P(a,r), B_P^*(a,r)$, and $S_P(a,r)$

When $A \subseteq P \subseteq \mathbb{R}^n$, we say that A is open in P when for all $a \in A$, there exists $r > 0$ such that $B_P(a,r) \subseteq A$

And we say that A is closed in P when A^c is open in P. (where $A^c = P \setminus A$)

Thm: (Basic Properties of Open Sets in $P \subseteq \mathbb{R}^n$)
1. $\emptyset$ and P are open in P
2. If $E_k \subseteq P$ is open in P for each $k \in K$, then
   $\bigcup_{k\in K} E_k$ is open in P
3. If $E_1, \ldots, E_\ell \subseteq P$ are open in P then
   $\bigcap_{k=\ell}^{\ell} E_k$ is open in P

Proof. Exercise.



Theorem: (Characterization of Open and Closed Sets in $P \subseteq \mathbb{R}^n$)

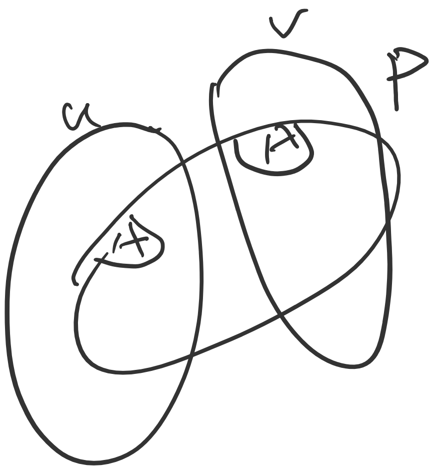
Let $A \subseteq P \subseteq \mathbb{R}^n$
1. A is open in P if and only if $A = U \cap P$ for some set $U \subseteq \mathbb{R}^n$ which is open in $\mathbb{R}^n$.
2. A is closed in P if and only if $A = K \cap P$ for some set $K \subseteq \mathbb{R}^n$ which is closed in $\mathbb{R}^n$.

Proof:
1. Suppose A is open in P. For each $a \in A$. Choose $r_a > 0$ such that $B_P(a, r_a) \subseteq A$, that is $B(a, r_a) \cap P \subseteq A$.
   Let $U = \bigcup_{a\in A} B(a, r_a)$
   Note that U is open in $\mathbb{R}^n$, because it is a union of open sets in $\mathbb{R}^n$.
   Note that $A \subseteq U$ (since for each $a \in A, a \in B(a, r_A) \subseteq U$) and $A \subseteq P$ so $A \subseteq U \cap P$
   Also, note that $U \cap P = \left(\bigcup_{a\in A} B(a,r_a)\right) \cap P = \bigcup_{a\in A}\left(B(a,r_a) \cap P\right) = \bigcup_{a\in A} B_P(a, r_a) \subseteq A$
   Since $B_P(a, r_a) \subseteq A$ for all $a \in A$.

   Thus we have $A = U \cap P$.

   Suppose conversely, that $A = U \cap P$, where $U \subseteq \mathbb{R}^n$ is an open set in $\mathbb{R}^n$
   Let $a \in A$. Since $a \in A$ and $a = U \cap P \subseteq U$. We have $a \in U$.
   Since $a \in U$ and U is open in $\mathbb{R}^n$. We can choose $r > 0$ so that $B(a,r) \subseteq U$.
   Then we have $B_P(a,r) = B(a,r) \cap P \subseteq U \cap P = A$
   That A is open in P.

   End of Part 1.
   Part 2 left as an exercise.



1.32 Theorem: Let $A \subseteq P \subseteq \mathbb{R}^n$. Define A to be **connected in** P when there do not exists set $E, F \subseteq P$ which are open in P and which separate A. Define A to be **compact in** P when for every set S of open sets in P such that $A \subseteq \bigcup S$ there exists a finite subset $T \subseteq S$ such that $A \subset \bigcup T$. Then

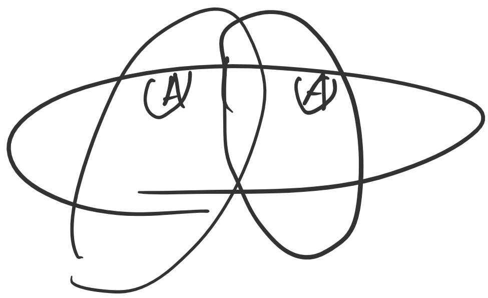
A is connected in $P \leftrightarrow A$ is connected in R^n

Proof:

For the purpose of this theorem alone, we define A to be

Suppose that A is disconnected in $\mathbb{R}^n$. Let U and V be open sets in $\mathbb{R}^n$ which separate A.
Let $E = U \cap P$ and $F = V \cap P$.
Note that E and F are open in $\mathbb{R}^n$.
Verify that E and F separate A

Suppose conversely that A is disconnected in P. Choose open sets $E, F \subseteq P$ which are open in P and separate A.



Radius 2 ball shunk so no intersect

Choose open sets U and V in $\mathbb{R}^n$
Such that $E = U \cap P$ and $F = V \cap P$.

We have $U \cap A \supseteq E \cap A \neq \emptyset, V \cap A \supseteq F \cap A \neq \emptyset, A \subseteq E \cup F \subseteq U \cup V$, but we might have $U \cap V \neq \emptyset$

For each $a \in U$, choose $r_a > 0$ so that $B_P(a, 2r_a) \subseteq U$ (we can do this since U is open in $\mathbb{R}^n$)

Then let $U_o = \bigcup_{a\in E} B(a, r_a)$

Note that U_o is open in $\mathbb{R}^n$ (since it is a union of open sets)
And $E \subseteq U_o$ (since each $a \in E$ lies in $B(a, r_a) \subseteq U$)

And $U_0 \cap P = \left(\bigcup_{a\in E} B(a,r)\right) \cap P = \bigcup_{a\in E}\left(B(a,r_a) \cap P\right) = \bigcup_{a\in E} B_P(a,r)$

And $E \subseteq P$ so that $E \subseteq U_0 \cap P$
And $U_0 \subseteq U$ (since each $B(a,r) \subseteq U$)

So $U_0 \cap P \subseteq U \cap P = E$

Thus $E = U_0 \cap P$

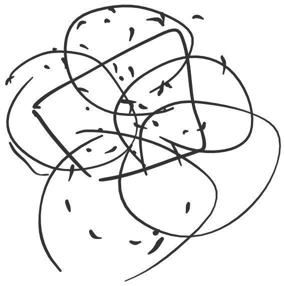
Similarly, for each $b \in V$ choose $s_b > 0$ so that $B(b, 2s_b) \subseteq V$
Then let $V_0 = \bigcup_{b\in F} B(b, s_b)$
And then we have $V \cap P = F$

We have $E \cap A \subseteq U_0 \cap A \neq \emptyset$
$\emptyset \neq F \cap A \subseteq V_0 \cap A$
And $A \subseteq E \cup F \subseteq U_0 \cup V_0$

We claim that $U_0 \cap V_0 = \emptyset$

Suppose, for a contradiction, that $U_0 \cap V_0 \neq \emptyset$
Choose $c \in U_0 \cap V_0$
Since $c \in U_0 = \bigcup_{a \in E} B(a, r)$
We can choose $a \in E$
Such that $c \in B(a, r_a)$
Likewise, we can choose $b \in F$
So that $c \in B(b, s_b)$

Say $r_a \leq S_b$
Then $|a-b| \leq |a-c| + |c-b| < r_a + s_b < 2s_b$
So we have $a \in B(b, 2s_b) \subseteq V$
But then $a \in P$ and $a \in V$ so $a \in V \cap P = F$
And $a \in E$
Which contradicts the fact that $E \cap F = \emptyset$



Corollary
For $A \subseteq \mathbb{R}^n$
A is connected if and only if the only subsets of A. Which are both open and closed in A. Are the sets $\emptyset$ and A

Both open and closed

Compactness
1.22 Definition:
Let $A \subseteq \mathbb{R}^n$. An **open cover** of A is a set S of open sets in $\mathbb{R}^n$ such that $A \subseteq \bigcup S$. A **subcover** of an open cover S of A is a subset $T \subseteq S$ such that $A \subseteq \bigcup T$. We say that A is **compact** when every open cover of A has a finite subcover.

1.24 Theorem:
(The Nested Interval Theorem) Let $I_0, I_1, I_2, \ldots$ be nonempty, closed bounded intervals in $\mathbb{R}$. Suppose that $I_0 \supseteq I_1 \supset I_2 \supset \cdots$, Then $\bigcap_{k=0}^{\infty} I_k \neq \emptyset$.

Proof: For each $k \geq 1$, let $I_k = [a_k, b_k]$ with $a_k < b_k$. For each k, since $I_{k+1} \subseteq I_k$ we have $a_k \leq a_{k+1} < b_{k+1} \leq b_{k+1}$. Since $a_k \leq a_{k+1}$ for all k, the sequence (a_k) is increasing. Since $a_k < b_k \leq b_{k-1} \leq \cdots \leq b_1$ for all k, the sequence (a_k) is bounded above by b_1. Since (a_k) is increasing and bounded above, it converges. Let $a = \sup\{a_k\} = \lim_{k\to\infty} a_k$. Similarly, (b_k) is decreasing and bounded below by a_1, and so it converges. Let $b = \inf\{b_k\} = \lim_{k\to\infty} b_k$. Fix $m \geq 1$. For all $k \geq m$ we have $a_m < b_m \leq b_{m+1} \leq \cdots \leq b_k$. Since $a_k \leq b_k$ for all k, by the Comparison Theorem we have $a \leq b$, and so the interval $[a, b]$ is not empty. Since (a_k) is increasing with $a_k \to a$, it follows (we leave the proof as an exercise) that $a_k \leq a$ for all $k \geq 1$. Similarly, we have $b_k \geq b$ for all $k \geq 1$ and so $[a, b] \subseteq [a_k, b_k] = I_k$. Thus $[a, b] \subseteq \bigcap_{k=1}^{\infty} I_k$, and so $\bigcap_{k=1}^{\infty} I_k \neq \emptyset$.

Cartesian product

1.25 Definition:
A **closed rectangle** in $\mathbb{R}^n$ is a set of the form
$$R = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$$
$$= \{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \,|\, a_j \leq x_j \leq b_j \text{ for all } j\}.$$

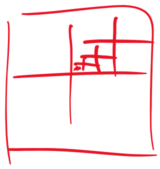
1.26 Theorem:
(Nested Rectangles) Let $R_1, R_2, R_3, \ldots$ be closed rectangles in $\mathbb{R}^n$ with $R_1 \supseteq R_2 \supseteq R_3 \supseteq \cdots$. Then
$$\bigcap_{k=1}^{\infty} R_k \neq \emptyset.$$

Divide each interval into two equal subintervals

Proof: Let $R_k = [a_{k,1}, b_{k,1}] \times [a_{k,2}, b_{k,2}] \times \cdots \times [a_{k,n}, b_{k,n}]$. Since $R_1 \supseteq R_2 \supseteq \cdots$, it follows that for each index j with $1 \leq j \leq n$ we have $[a_{1,j}, b_{1,j}] \supseteq [a_{2,j}, b_{2,j}] \supseteq \cdots$. By the Nested Interval Theorem, for each index j we can choose $u_j \in \bigcap_{k=1}^{\infty} [a_{k,j}, b_{k,j}]$. Then for $u = (u_1, u_2, \ldots, u_n)$ we have $u \in \bigcap_{k=1}^{\infty} R_k$.

Core: Construct a rectangle to a point. And open cover
Must contain, thus a finite subcover, contradiction.

Zoom in to a point



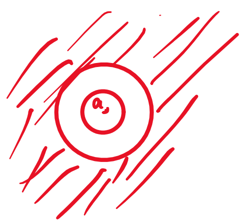
1.27 Theorem: (Compactness of Rectangles) Every closed rectangle in $\mathbb{R}^n$ is compact.
Proof: Let $R = I_1 \times I_2 \times \cdots \times I_n$ where $I_j = [a_j, b_j]$ with $a_j \leq b_j$. Let d be the diameter of R, that is $d = \text{diam}(R) = \left(\sum_{j=1}^{n} (b_j - a_j)^2\right)^{\frac{1}{2}} = \frac{d}{2^{k-1}}$. Let S be an open cover of R. Suppose, for a contradiction, that S does not have a finite subset which covers R. Let $a_{1,j} = a_j, b_{1,j} = b_j, I_{1,j} = I_j = [a_{1,j}, b_{1,j}]$ and $R_1 = R = I_{1,1} \times \cdots \times I_{1,n}$. Recursively, we construct rectangles $R = R_1 \supseteq R_2 \supseteq R_3 \supseteq \cdots$, with $R_k = I_{k,1} \times \cdots \times I_{k,n}$ where $I_{k,j} = [a_{k,j}, b_{k,j}]$, and $d_k = \text{diam}(R_k) = \left(\sum_{j=1}^{n} (b_{k,j} - a_{k,j})^2\right)^{\frac{1}{2}} = \frac{d}{2^{k-1}}$, such that the open cover S does not have a finite subset which covers any of the rectangles R_k. We do this recursive construction as follows. Having constructed one of the rectangles R_k, we partition each of the intervals $I_{k,j} = [a_{k,j}, b_{k,j}]$ into the two equal-sized subintervals $\left[a_{k,j}, \frac{a_{k,j}+b_{k,j}}{2}\right]$ and $\left[\frac{a_{k,j}+b_{k,h}}{2}, b_{k,j}\right]$, and we thereby partition the rectangle R_k into 2^n equal-sized sub-rectangles. We choose R_{k+1} to be equal one of these 2^n sub-rectangles with the property that the open cover S does not have a finite subset which covers R_{k+1} (if each of the 2^n sub-rectangles could be covered by a finite subset of S then the union of these 2^n finite subsets would be a finite subset of S which covers R_k).

Choose a r > 0
$n \to \infty$, will contain since open

By the Nested Rectangles Theorem, we can choose an element $u \in \bigcap_{k=1}^{\infty} R_k$. Since $u \in R$ and S covers R we can choose an open set $U \in S$ such that $u \in U$. Since U is open, we can choose $r > 0$ such that $B(u, r) \subseteq U$. Since $d_k \to 0$ we can choose k so that $d_k < r$. Since $u \in R_k$ and $\text{diam}(R_k) = d_k < r$ we have $R_k \subseteq B(u, r) \subseteq U$. Thus S does have a finite subset, namely $\{U\}$, which covers R_k, giving the desired contradiction.

1.28 Theorem:
Let $A \subseteq K \subseteq \mathbb{R}^n$. If A is closed and K is compact then A is compact.

Proof: Suppose that A is closed in $\mathbb{R}^n$ and that K is compact. Let S be an open cover of A. Let $A^c = \mathbb{R}^n \setminus A$. Since $A \subseteq \bigcup S$ we have $\bigcup S \cup \{A^c\} = \mathbb{R}^n$ and so $S \cup \{A^c\}$ is an open cover of K. Since K is compact, we can choose a finite subset $T \subseteq S \cup \{A^c\}$ with $K \subseteq \bigcup T$. Since $A \subseteq K \subseteq \bigcup T$ we also have $A \subseteq \bigcup (T \setminus \{A^c\})$. Thus, the open cover S of A does have a finite subcover, namely $T \setminus \{A^c\}$, and so A is compact, as required.



1.29 Theorem: (The Heine-Borel Theorem) Let $A \subseteq \mathbb{R}^n$. Then A is compact if and only if A is closed and bounded.
Preliminary theorem: Theorem. (Closed Subsets of Compact sets are compact)
Proof: Suppose that A is compact. Suppose, for a contradiction, that A is not bounded. For each $k \in \mathbb{Z}^+$ let $U_k = B(0, k)$ and let $S = \{U_k | k \in \mathbb{Z}^+\}$. Then $\bigcup S = \mathbb{R}^n$ so S is an open cover of A. Let T be any finite subset of S. If $T = \emptyset$ then $\bigcup T = \emptyset$ and $A \not\subseteq \bigcup T$. This shows that the open cover S has no finite subcover, T, which contradicts the fact that A is compact.

Next suppose, for a contradiction, that A is not closed. By Part (1) of Theorem 8.16, it follows that $A' \not\subseteq A$. Choose $a \in A'$ with $a \notin A$. For each $k \in \mathbb{Z}^+$ let U_k be the open set $U_k = \bar{B}\left(a, \frac{1}{k}\right)^c = \left\{x \in \mathbb{R}^n \middle| |x-a| > \frac{1}{k}\right\}$ and let $S = \{U_k | k \in \mathbb{Z}^+\}$. Note that $\bigcup S = \mathbb{R}^n \setminus \{a\}$ so S is an open cover of A. Let T be any finite subset of S. If $T = \emptyset$ then $\bigcup T = \emptyset$ so $A \not\subseteq \bigcup T$ (since A is not closed so $A \neq \emptyset$). Suppose that $T \neq \emptyset$, say $T = \{U_{k_1}, U_{k_2}, \ldots, U_{k_m}\}$ with $k_1 < k_2 < \cdots <$

k_m. Since $U_{k_1} \subseteq U_{k_2} \subseteq \cdots \subseteq U_{k_m}$ we have $\cup T = \cup_{i=1}^{m} U_{k_i} = U_{k_m} = \bar{B}\left(a, \frac{1}{k_m}\right)^c$. Since a is a limit point of A we have $B\left(a, \frac{1}{k_m}\right) \neq \emptyset$ hence $\bar{B}\left(a, \frac{1}{k_m}\right) \cap A \neq \emptyset$ and so $A \not\subseteq \bar{B}\left(a, \frac{1}{k_m}\right)^c$, hence $A \not\subseteq \cup T$. This shows that the open cover S has no finite subcover T, which again contradicts the fact that A is compact.

Suppose, conversely, that A is closed and bounded. Since A is bounded we can choose $r > 0$ so that $A \subseteq B(0, r)$. Let R be the closed rectangle $R = \{x \in \mathbb{R}^n \mid |x_k| \leq r \text{ for all } k\}$.
Note that $B(0, r) \subseteq R$ since when $x = (x_1, \dots, x_n) \in B(0, r)$, for each index k we have

$$|x_k| = (x_k^2)^{\frac{1}{2}} \leq \left(\sum_{i=1}^{n} x_i^2\right)^{\frac{1}{2}} = |x| < r.$$

Since A is closed and $A \subseteq R$ and R is compact, it follows that A is compact, by the above theorem.

Eg. Let $A = \left\{\frac{1}{n} \middle| n \in Z^+\right\}$
Show A is not compact

Sketch solution:

For $n \in \mathbb{Z}^+$, let $U_n = B\left(\frac{1}{n}, \frac{1}{2n(n+1)}\right) = \left(\frac{1}{n} - \left(\frac{1}{2n(n+1)}\right), \left(\frac{1}{n} + \frac{1}{2n(n+1)}\right)\right)$
Verify that for each $n \in \mathbb{Z}^+$
$\frac{1}{n} \in U_k \leftrightarrow k = n$

And hence $A \subseteq \cup_{n=1}^{\infty} U_n$
But A is not contained in the union of finitely many of the sets U_r.

Covers the tail?
Limit point

Eg. Let $B = \bar{A} = \left\{\frac{1}{n} \middle| n \in \mathbb{Z}^+\right\} \cup \{0\}$
Show that B is compact

Etch solution:

Let S be an open cover of B. Choose $U \in S$ with $0 \in U$
Choose $r > 0$ so that $B(a, r) \subseteq U$

Note that $B(0, r)$ contains all the points in B except for
$\frac{1}{1}, \frac{1}{2}, \dots, \frac{1}{n}$

Where $\frac{1}{n+1} < r$

For $1 \leq k \leq n$, choose
$U_k \in S$ so that $\frac{1}{k} \in U_k$

Then $T = \{U_0, U_1, \dots, U_n\}$ is a finite subcover of S.

Theorem. (Closed Subsets of Compact sets are compact)

Let $A \subseteq P \subseteq R^n$. If P is compact and A is closed in $\mathbb{R}^n$, then A is compact.

Proof. Suppose P is compact and A is closed (in $\mathbb{R}^n$)
Let S be any open cover of A, since A is closed, A^c is open, where $A^c = \mathbb{R}^n \setminus A$
Since S covers A, $S \cup \{A^c\}$ covers $\mathbb{R}^n$.
Hence $S \cup \{A^c\}$ covers P
Since P is compact, we can choose a finite subset T of S
Such that $T \cup \{A^c\}$ covers P.
It follows that T covers A (if $a \in A$ then $a \in \cup(T \cup \{A^c\}) = \cup T \cup A^c$
So either $a \in \cup T$ or $a \in A^c$ but $a \in A$ so $a \notin A^c$ hence $a \in \cup T$)
Thus, A is compact

Thm

Chapter 2. Introduction to Vector Valued Functions

2019年5月22日 2:26

2.1 Definition: Let $D \subseteq \mathbb{R}^n$. We say that f is a **function** or **map** from D to $\mathbb{R}^m$, and we write $f: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$, when for every $x \in D$ there is a unique point $y = f(x) \in \mathbb{R}^m$. The set D is called the **domain** of the function f.

The **graph** of the function f is the set

$$\text{Graph}(f) = \{(x, f(x)) | x \in D\} \subseteq \mathbb{R}^{n+m}$$

We say the graph of f is defined **explicitly** by the equation $y = f(x)$.

Geometric Objects

The **null set** of f is the set (kernel)

$$\text{Null}(f) = f^{-1}(0) = \{x \in D | f(x) = 0\} \subseteq \mathbb{R}^n.$$

More generally, given $k \in \mathbb{R}^m$, the **level set** $f^{-1}(k)$, also called the **inverse image** of k under f, is the set

$$f^{-1}(k) = \{x \in D | f(x) = k\} \subseteq \mathbb{R}^n.$$

More generally still, given a subset $B \subseteq \mathbb{R}^n$, the **inverse image** of B under f is the set

$$f^{-1}(B) = \{x \in D | f(x) \in B\} \subseteq \mathbb{R}^n.$$

We say the level set $f^{-1}(k)$ is defined **implicitly** by the equation $f(x) = k$.

The **range** of f, also called the **image** of f, is the set

$$\text{Image}(f) = \text{Range}(f) = f(D) = \{f(x) | x \in D\} \subseteq \mathbb{R}^m.$$

More generally, given a set $A \subseteq D$, the **image** of A under f is the set

$$f(A) = \{f(x) | x \in A\} \subseteq \mathbb{R}^m.$$

We say the range of f is defined **parametrically** by the equation $y = f(x)$, and for $x = (x_1, x_2, \dots, x_n) \in D$, the variables $x_1, x_2, \dots, x_n$ are called the **parameters.**

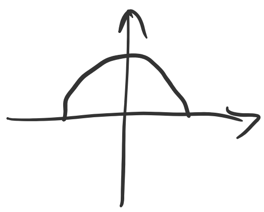
2.2 Note: The graph, the level sets and the range of a function $f: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ are geometric objects such as points, curves, surfaces, or higher dimensional analogues of these. In accordance with the above definitions, a curve in $\mathbb{R}^2$ or in $\mathbb{R}^3$, or a surface in $\mathbb{R}^3$, can be defined explicitly, implicitly, or parametrically.

A curve in $\mathbb{R}^2$ can be defined explicitly as the graph of a function $f: D \subseteq \mathbb{R} \to \mathbb{R}$, implicitly as the null set (or a level set) of a function $g: D \subseteq \mathbb{R}^2 \to \mathbb{R}$, or parametrically as the range of a function $\alpha: D \subseteq \mathbb{R} \to \mathbb{R}^3$.

A curve in $\mathbb{R}^3$ can be defined explicitly as the graph of a function $f: D \subseteq \mathbb{R} \to \mathbb{R}^2$, implicitly as the null set (or a level set) of a function $g: D \subseteq \mathbb{R}^3 \to \mathbb{R}^2$, or parametrically as the range of a map $\alpha: D \subseteq \mathbb{R} \to \mathbb{R}^3$.

A surface in $\mathbb{R}^3$ can be defined explicitly as the graph of a function $f: \mathbb{R}^2 \to \mathbb{R}$, implicitly as the null set (or as a level set) of a function $g: D \subseteq \mathbb{R}^3 \to \mathbb{R}$, or parametrically as the range of a function $\sigma: D \subseteq \mathbb{R}^2 \to \mathbb{R}^3$.

2.3 Example:
Consider the unit circle $x^2 + y^2 = 1$ in $\mathbb{R}^2$. For $f: [-1,1] \subseteq \mathbb{R} \to \mathbb{R}$ given by $f(x) = \sqrt{1 - x^2}$, the graph of f, that is the curve $y = f(x)$, is equal to the top half of the unit circle. For $g: \mathbb{R}^2 \to \mathbb{R}$ given by $g(x, y) = x^2 + y^2 - 1$, the null set of g, that is the curve $x^2 + y^2 = 1$, is equal to the entire circle. For $\alpha: \mathbb{R} \to \mathbb{R}^2$ given by $\alpha(t) = (\cos t, \sin t)$, the range of α, that is the curve $(x, y) = \alpha(t)$, is equal to the entire circle.



2.4 Example:
Consider the ellipse which is the intersection of the cylinder $x^2 + y^2 = 1$ with the plane $z = x + y$ in $\mathbb{R}^3$. The ellipse is given implicitly by the two equations $x^2 + y^2 = 1$ and $z = x + y$ which can be written in vector form as the single equation $(x^2 + y^2 - 1, z - x - y) = (0,0)$, and so it is the null set of the function $g: \mathbb{R}^3 \to \mathbb{R}^2$ given by $g(x, y, z) = (x^2 + y^2 - 1, z - x - y)$. To obtain a parametric description of the ellipse, note that to get $x^2 + y^2 = 1$ we can take $x = \cos t$ and $y = \sin t$, and then to get $z = x + y$ we can take $z = \cos t + \sin t$, and so the ellipse is given parametrically by $(x, y, z) = (\cos t, \sin t, \cos t + \sin t)$. In other words, the ellipse is the range of the function $\alpha: \mathbb{R} \to \mathbb{R}^3$ given by $\alpha(t) = (\cos t, \sin t, \cos t + \sin t)$. To obtain an explicit description for half of the ellipse, note that the top half of the circle $x^2 + y^2 = 1$ is given by $y = \sqrt{1 - x^2}$ and then to get $z = x + y$ we need $z = x + \sqrt{1 - x^2}$, and so the right half of the ellipse (when the y-axis points to the right) is given explicitly by $(y, z) = \left(\sqrt{1 - x^2}, x + \sqrt{1 - x^2}\right)$. In other words, the right half of the ellipse is the graph of the function $g: [-1,1] \subseteq \mathbb{R} \to \mathbb{R}^2$ given by $g(x) = \left(\sqrt{1 - x^2}, x + \sqrt{1 - x^2}\right)$.

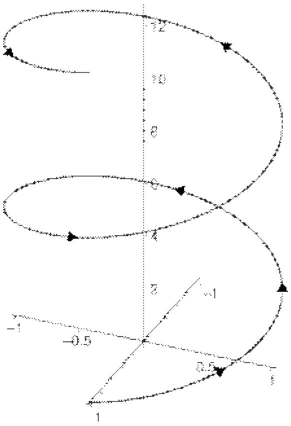
Parametric
$(x, y) = g(t) = (x(t), y(t)) = (\cos t, \sin t)$
Or by $x = x(t) = \cos t$, and $y = y(t) = \sin t$.

The entire circle can be given implicitly by $x^2 + y^2 = 1$

Null set of which $h: \mathbb{R}^2 \to \mathbb{R}$ is given by $h(x, y) = x^2 + y^2 - 1$

2.8 Exercise: The **helix** in $\mathbb{R}^3$ is given explicitly by x = cos z and y = sin z. Sketch the curve and find an implicit and a parametric equation for the curve.

$f: \mathbb{R} \to \mathbb{R}^2$
$(x, y) = f(z) = (\cos z, \sin z)$

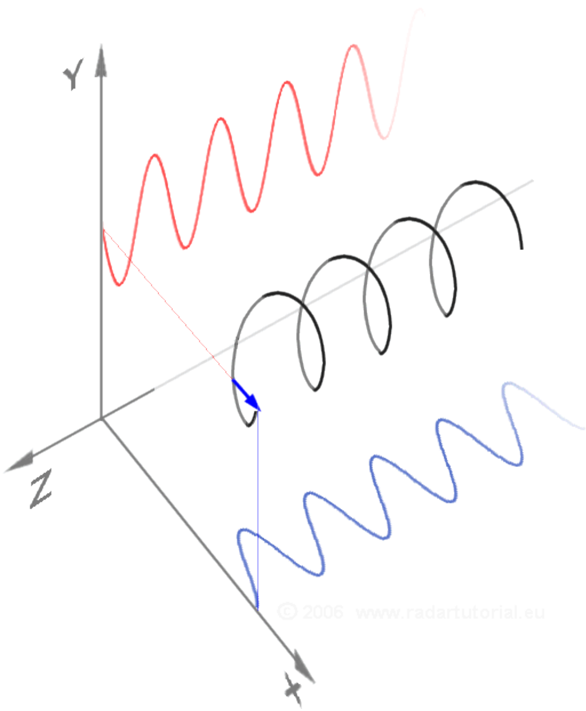


Parametric functions are not unique.

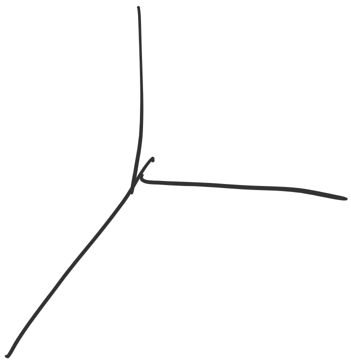
The helix is given parametrically by $(x, y, z) = (\cos t, \sin t, t)$
And it is given implicitly by
$x = \cos z, y = \sin z$.

(So the helix is the null set of Null(g) where $g: \mathbb{R}^3 \to \mathbb{R}^2$ is given by $g(x, y, z) = (x - \cos z, y - \sin z)$

X = Cosine curve.
Y = sine curve

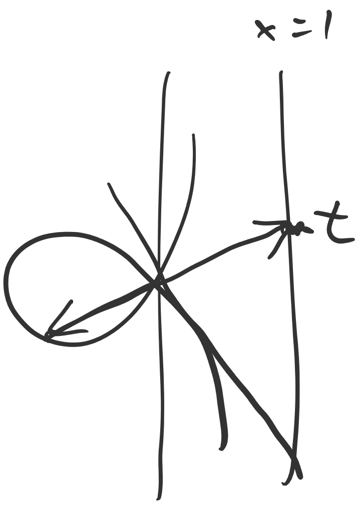


Sine sheet and cosine sheet intersection.



2.9 Exercise: The **alpha curve** is given implicitly by $y^2 = x^3 + x^2 = x^2(x+1)$. Sketch the curve,

find explicit equations for the top and bottom halves of the curve, and find a parametric equation for the entire curve.



The top and bottom halves are given by $y = \pm\sqrt{x^3 + x^2}$

Parametric Equation
Projection

The line from $(0,0)$ to $(1, t)$ is given parametrically by
$(x, y) = S(1, t), \quad s \in \mathbb{R}$
$= (s, st)$

And $(x, y) = (s, st)$ lies on the alpha curve when $y^2 = x^3 + x^2$

$s^2t^2 = s^3 + s^2$

$s = 0$ or $s = t^2 - 1$

When $s = t^2 - 1$ we obtained the point

$s^2t^2 = s^3 + s^2$

$s = 0$ or $s = t^2 - 1$

When $s = t^2 - 1$, we obtained the point
$(x, y) = (s, st) = \left(t^2 - 1, t(t^2 - 1)\right)$
Parametric equation of α curve.

Thus, the alpha curve is parametrically by $(x, y) = \left(t^2 - 1, t(t^2 - 1)\right), t \in \mathbb{R}$

Given (x, y) on the alpha curve, note that t is given by $t = \frac{y}{x}$, unless $x = 0$
Find t, $\frac{y}{x}$

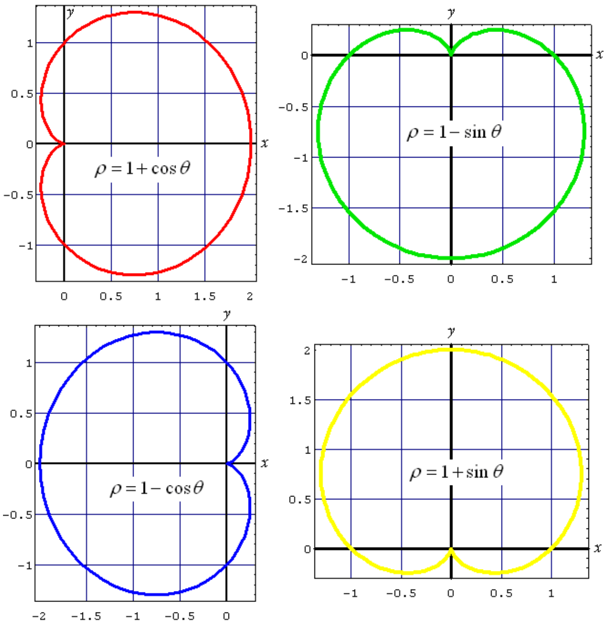
Origin being mapped to two points, so not bijective.

Send rational coordinates to rational coordinates
Proof alpha curve number density

2.10 Exercise: The curve which is given explicitly in polar coordinates by r = r(θ) is given parametrically in Cartesian coordinates by (x; y) = $\alpha(t) = (r(t)\cos t, r(t)\sin t)$.
Sketch the **cardioid** which is given in polar coordinates by r = r(θ) = 1+cos θ, then find an implicit equation for the curve.

Polar coordinates

List a table to sketch the curve.



$r = \cos\theta$

Is a circle.

The cardioid is given parametrically in Cartesian coordinates (taking $t = \theta$ and using $x = r\cos\theta, y = r\sin\theta$)

$(x, y) = (r\cos\theta, r\sin\theta) = \left((1 + \cos t)\cos t, (1 + \cos t)\sin t\right)$

We can obtain an implicit equation in Cartesian coordinates as follows

$x = r\cos\theta, y = r\sin\theta, x^2 + y^2 = r^2$ as follows.

$r = 1 + \cos\theta$
$r^2 = r + r\cos\theta$
$x^2 + y^2 = \sqrt{x^2 + y^2} + x$
Or as

$x^2 + y^2 = (x^2 + y^2 - x)^2$

2.11 Exercise: The **twisted cubic, X,** is given parametrically by $(x, y, z) = \alpha(t) = (t, t^2, t^3)$.
Sketch the curve and find an implicit and an explicit equation for the curve.

(So it is the range of $P: \mathbb{R} \to \mathbb{R}^3$ is given by $(x, y, z) = f(t) = (t, t^2, t^3)$)

X is given explicitly by $y = x^2, z = x^3$
So X is the graph of the function $g: \mathbb{R} \to \mathbb{R}^2$ given by $(y, z) = g(x) = (x^2, x^3)$

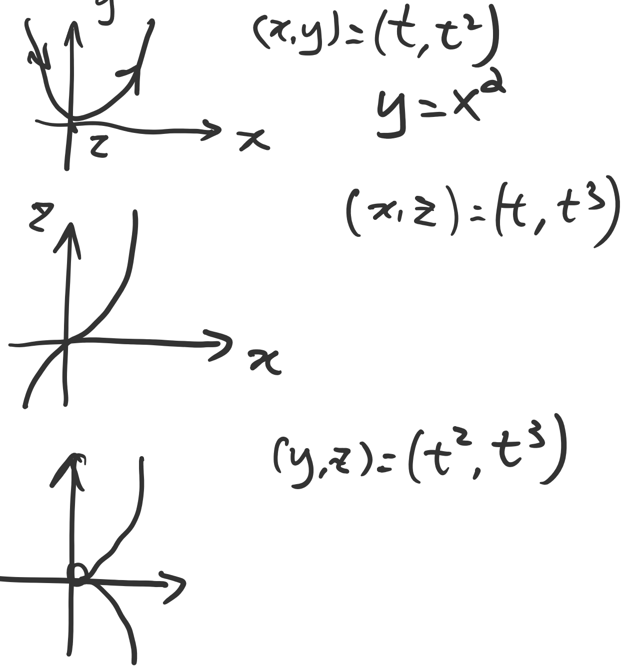
X is given implicitly by $y = x^2, z = x^3$

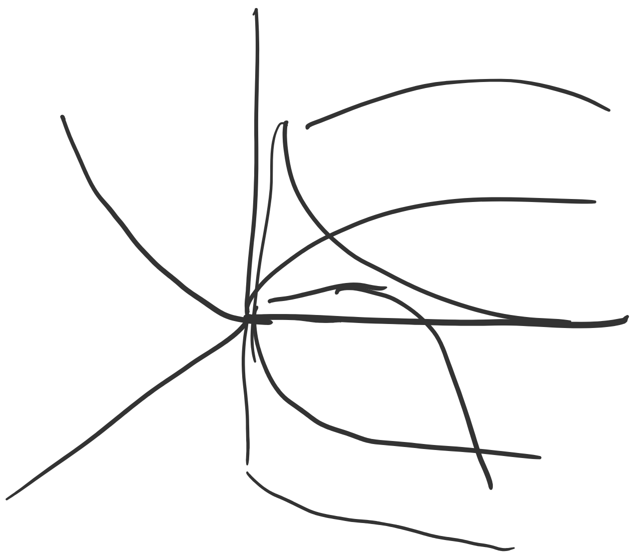
$z = xy$

(so X is the null set of $h: \mathbb{R} \to \mathbb{R}^2$ given by $(u, v) = h(x, y, z) = (y - x^2, z - x^3)$

Sketch

Right handed





Let us verify that $\text{Range}(f) = \text{Null}(h)$

Where $f(t) = (t, t^2, t^3)$
$h(x, y, z) = (y - x^2, z - x^2)$

Proof:

Let $(x, y, z) \in \text{Range}(f)$

Choose $t \in \mathbb{R}$ such that $(x, y, z) = f(t) = (t, t^2, t^3)$

Then $y = t^2 = x^2$
$z = t^3 = x^3$

So $h(x, y, z) = (y - x^2, z - x^3) = (0,0)$

Hence $(x, y, z) \in \text{Null}(h)$

Now let $(x, y, z) = \text{Null}(h)$

So $(y - x^2, z - x^3) = (0,0)$

$y = x^2$ and $z = x^3$

Let $t = x$
Then

$f(t) = (t, t^2, t^3) = (x, x^2, x^3) = (x, y, z)$
Hence
$(x, y, z) \in \text{Range}(f)$

2.12 Remark: In order to sketch a surface which is defined explicitly as a graph $z = f(x, y)$ or implicitly as a level set $g(x, y, z) = k$, it often helps to first sketch curves of intersection of the surface with various planes $x = c$, $y = c$ or $z = c$. The intersection of the graph $z = f(x, y)$ with the plane $z = c$ is given implicitly by $f(x, y) = c$. The intersection of the level set $g(x, y, z) = k$ with the plane $z = c$ is given implicitly by $g(x, y, z) = k$.

2.13 Exercise: Sketch the curve of intersection of the cylinder $x^2 + y^2 = 1$ with the parabolic sheet $z = x^2$ and find implicit, explicit, and parametric equations for the curve.

2.14 Exercise: Sketch the surface $z = x^2 + y^2$.

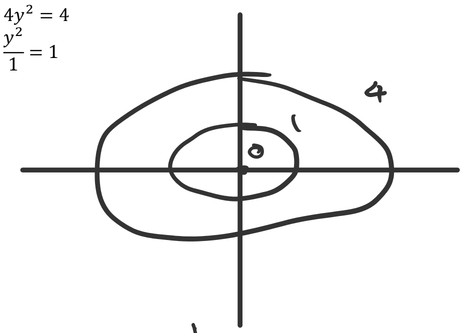
Sketch $z = x^2 + 4y^2$

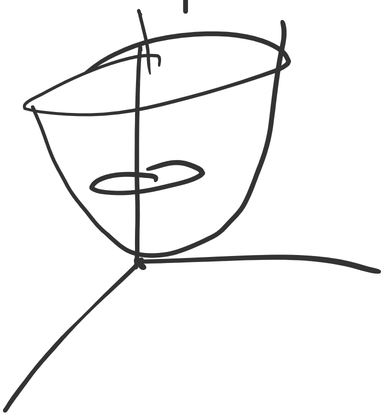
$z = 0$

$x^2 + 4y^2 = 0$
$(x, y) = (0,0)$

$z = 1$
$x^2 + 4y^2 = 1$
$\frac{x^2}{1} + \frac{y^2}{\frac{1}{4}} = 1$

$z = 4$
$x^2 + 4y^2 = 4$
$\frac{x^2}{4} + \frac{y^2}{1} = 1$





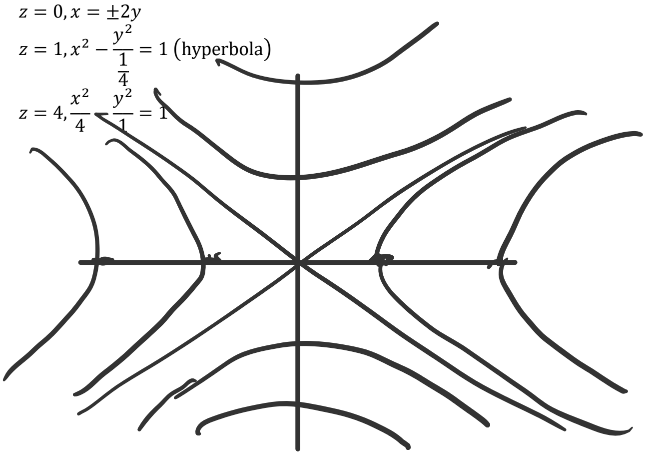
$x = 0, z = 4y^2$
$y = 0, z = x^2$

2.15 Exercise: Sketch the surface $z = 4x^2 - y^2$.

$z = 0, x = \pm 2y$

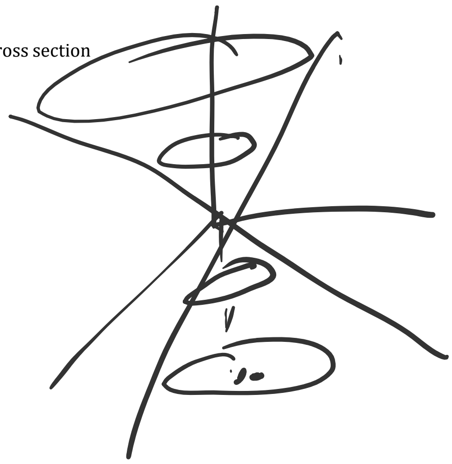
$z = 1, x^2 - \frac{y^2}{\frac{1}{4}} = 1$ (hyperbola)

$z = 4, \frac{x^2}{4} - \frac{y^2}{1} = 1$



2.16 Exercise: Sketch the surface $x^2 + 4y^2 - z^2 = 0$.

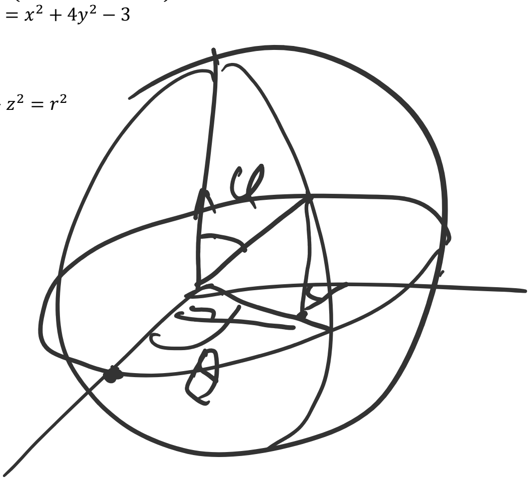
Same cross section



2.17 Exercise: Sketch the surface $(x, y, z) = (u, v, u^2 + 4v^2 - 3)$

$(x, y, z) = (u, v, u^2 + 4v^2 - 3)$
Same as $z = x^2 + 4y^2 - 3$

$x^2 + y^2 + z^2 = r^2$



Math Longitude

$z = r\cos\phi$

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

The sphere
$x^2 + y^2 + z^2 = r^2$
Is the range of $f: \mathbb{R}^2 \to \mathbb{R}^3$
$(x,y,z) = f(\theta,\phi) = (r\sin\phi\cos\theta, r\sin\phi\sin\theta, r\cos\phi)$

2.18 Exercise: Find a parametric equation $(x,y,z) = \sigma(\phi,\theta)$ for the sphere of radius r centred at the origin, where the parameters ϕ and θ are the angles of latitude and longtitude. In other words, find $\sigma(\phi,\theta)$ so that when $(x,y,z) = \sigma(\phi,\theta)$, ϕ is the angle between (0,0,1) and (x,y,z) and θ is the angle from (1,0) counterclockwise to (x,y).

2.19 Exercise: Find implicit and parametric equations for the **torus** which is obtained by rotating the circle $(x,z) = (R + r\cos\theta, r\sin\theta)$ about the z-axis.

2.20 Definition: An **affine space** in $\mathbb{R}^n$ is a set of the form $p + V = \{p + v | v \in V\}$ for some $p \in \mathbb{R}^n$ and some vector space $V \subseteq \mathbb{R}^n$. The **dimension** of the affine space $p + V$ is the same as the dimension of V. The set $p + V$ is called the affine space through p parallel to V, or the affine space through p perpendicular to $V^\perp$, where $V^\perp$ is the **orthogonal complement** of V, given by $V^\perp = \{x \in \mathbb{R}^n | x \cdot v = 0 \text{ for all } v \in V\}$.

2.21 Example:
In $\mathbb{R}^3$, the only zero dimensional vector space is the origin $\{0\}$, the 1-dimensional vector spaces are the lines through the origin, the 2-dimensional spaces are the planes through the origin, and the only 3-dimensional vector space is all of $\mathbb{R}^3$. The 0-dimensional affine spaces are the points in $\mathbb{R}^3$, the 1-dimensional affine spaces are the lines in $\mathbb{R}^3$, the 2-dimensional affine spaces are the planes in $\mathbb{R}^3$, and the only 3-dimensional affine space is all of $\mathbb{R}^3$.

2.22 Definition:
Let $f: \mathbb{R}^n \to \mathbb{R}^m$. The function f is called **linear** when it is of the form $f(x) = Ax$ for some matrix $A \in M_{m\times n}(\mathbb{R})$, and f is called **affine** when it is of the form $f(x) = Ax + b$ for some matrix $A \in M_{m\times n}(\mathbb{R})$ and some vector $b \in \mathbb{R}^m$.

2.23 Note:
Let $A \in M_{m\times n}$ and let $f: \mathbb{R}^n \to \mathbb{R}^m$ be the linear map $f(x) = Ax$. Let $u_1, \dots, u_n$ be the column vectors of A and let $v_1, \dots, v_m$ be the row vectors of A so that we have $A = (u_1, \dots, u_n) = (v_1, \dots, v_m)^T$. Let c be a point in the range of f, say $f(p) = c$. Then

$$\text{Range}(f) = \{Ax | x \in \mathbb{R}^n\} = \left\{\textstyle\sum_{i=1}^n u_i x_i \,\middle|\, \text{each } x_i \in \mathbb{R}\right\} = \text{Span}\{u_1, \dots, u_n\} = \text{Col}(A),$$
$$\text{Null}(f) = \text{Null}(A) = \{x \in \mathbb{R}^n | Ax = 0\} = \{x \in \mathbb{R}^n | v_i \cdot x = 0 \text{ for all } i\} = \text{Row}(A)^\perp,$$
$$f^{-1}(c) = \{x \in \mathbb{R}^n | Ax = c\} = \{x \in \mathbb{R}^n | Ax = Ap\} = \{x \in \mathbb{R}^n | A(x - p) = 0\}$$
$$= \{p + y \in \mathbb{R}^n | Ay = 0\} = p + \text{Null}(A), \text{ and}$$
$$\text{Graph}(f) = \left\{\begin{pmatrix} x \\ Ax \end{pmatrix} \middle| x \in \mathbb{R}^n\right\} = \text{Span}\left\{\begin{pmatrix} e_1 \\ u_1 \end{pmatrix}, \begin{pmatrix} e_2 \\ u_2 \end{pmatrix}, \dots, \begin{pmatrix} e_n \\ u_n \end{pmatrix}\right\} = \text{Col}\begin{pmatrix} I \\ A \end{pmatrix}.$$

It follows that

$$\dim\big(\text{Graph}(f)\big) = n$$
$$\dim\big(\text{Range}(f)\big) = \text{rank}(A) \text{ and}$$
$$\dim\big(\text{Null}(f)\big) = \dim\big(f^{-1}(c)\big) = \text{nullity}(A) = n - \text{rank}(A).$$

2.24 Note: Let $A \in M_{m\times n}(\mathbb{R})$, let $b \in \mathbb{R}^m$ and let $f(x) = Ax + b$. Let c be in the range of f with say $f(p) = c$. Then

$$\text{Graph}(f) = \left\{\begin{pmatrix} x \\ Ax + b \end{pmatrix} \middle| x \in \mathbb{R}^n\right\} = \begin{pmatrix} 0 \\ b \end{pmatrix} + \text{Col}\begin{pmatrix} I \\ A \end{pmatrix},$$
$$\text{Range}(f) = \{Ax + b | x \in \mathbb{R}^n\} = b + \text{Col}(A), \text{ and}$$
$$f^{-1}(c) = \{x \in \mathbb{R}^n | Ax + b = c = Ap + b\} = \{x \in \mathbb{R}^n | A(x - p) = 0\} = p + \text{Null}(A),$$

Note that if $u_1, u_2, \dots, u_n$ are the columns of A and $e_1, e_2, \dots, e_n$ are the standard basis vectors for $\mathbb{R}^n$, then we have $f(0) = b$ and $f(e_i) = Ae_i + b = u_i + b$. If $v_1, \dots, v_m$ are the row vectors of A and $k = c - b$, then since

$$f(x) = c \leftrightarrow Ax + b = c \leftrightarrow Ax = k \leftrightarrow v_i \cdot x = k_i \text{ for all } i$$

It follows that the level set $f(x) = c$ is the intersection of the affine spaces $v_i \cdot x = k_i$, and we note that the space $v_i \cdot x = k_i$ is the affine space in $\mathbb{R}^m$ of dimension $m - 1$ through p perpendicular to v_i.

2.25 Exercise: Define $f: \mathbb{R}^3 \to \mathbb{R}^2$ by $f(x,y,z) = (x + 3y + 2z, 2z + 5y + 3z)$ and let $(a,b) = (1,1)$. Find a parametric equation for the level set $f(x,y,z) = (a,b)$.

2.26 Exercise: Let $A = \begin{pmatrix} 4 & 1 & 1 \\ 1 & 0 & 2 \\ 5 & 2 & -4 \end{pmatrix}$ and $b = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$ and let $f(x) = Ax + b$. Find an implicit equation for the range of f.

Sketch.

Short test.
Level curves.

Easy tests

Half way of computation and proof
Prove theorems.

Know solution to every assignment questions.

Test 1 Preparation

2019年5月23日 5:02

NOTE: The MATH 247 Term Test 1 will be held on
Monday May 27, from 12:30-1:20 pm, in MC 4063.

The test will cover Chapters 1 and 2 (omit sections 1.22-1.29)
and Assignments 1 and 2. There will be 4 questions.
you will be asked to prove 1 of the following 3 theorems:
Theorem 1.19 Part 3 (Closure and Limit Points)
Theorem 1.31 Part 1 (Open Subsets of P)
Theorem 1.32 Part 1 (Connected Subsets of P)
No calculators will be allowed.

来自 <http://www.math.uwaterloo.ca/~snew/math247-2019-S/index.html>

1.19 Theorem: (Properties of Interior, Limit and Boundary Points)
Let $A \subseteq \mathbb{R}^n$.

1. A^0 is equal to the set of all interior points of A.
2. A is closed if and only if $A' \subseteq A$.
3. $\bar{A} = A \cup A'$.
4. $\partial A = \bar{A} \setminus A^0$. or equivalently $\bar{A} = A^0 \cup \partial A$ and $A^0 \cap \partial A = \emptyset$

To prove Part (3) we shall prove that $A \cup A'$ is the smallest closed set which contains A. It is clear that $A \cup A'$ contains A. We claim that $A \cup A'$ is closed, that is $(A \cup A')^c$ is open. Let $a \in (A \cup A')^c$, that is let $a \in \mathbb{R}^n$ with $a \notin A$ and $a \notin A'$. Since $a \notin A'$ we can choose $r > 0$ so that $B(a,r) \cap A = \emptyset$. We claim that because $B(a,r) \cap A = \emptyset$ it follows that $B(a,r) \cap A' = \emptyset$.

(Point b arbitrarity close to a)
Suppose, for a contradiction, that $B(a,r) \cap A' \neq \emptyset$. Choose $b \in B(a,r) \cap A'$. Since $b \in B(a,r)$ and $B(a,r)$ is open, we can choose $s > 0$ so that $B(b,s) \subseteq B(a,r)$. Since $b \in A'$ (b is a limit point) it follows that $B(b,s) \cap A \neq \emptyset$. Choose $x \in B(b,s) \cap A$. Then we have $x \in B(b,s) \subseteq B(a,r)$ and $x \in A$ and so $x \in B(a,r) \cap A$, which contradicts the fact that $B(a,r) \cap A = \emptyset$.

Thus $B(a,r) \cap A' = \emptyset$, as claimed. Since $B(a,r) \cap A = \emptyset$ and $B(a,r) \cap A' = \emptyset$ it follows that $B(a,r) \cap (A \cup A') = \emptyset$ hence $B(a,r) \subseteq (A \cup A')^c$. Thus proves that $(A \cup A')^c$ is open, and hence $A \cup A'$ is closed.

It remains to show that for every closed set K with $A \subseteq K$ we have $A \cup A' \subseteq K$. Let K be a closed set in $\mathbb{R}^n$ with $A \subseteq K$. Note that since $A \subseteq K$ it follows that $A' \subseteq K'$ because if $a \in A'$ then for all $r > 0$ we have $B(a,r) \cap A \neq \emptyset$ hence $B(a,r) \cap K \neq \emptyset$ and so $a \in K'$. Since K is closed we have $K' \subseteq K$ by Parts (2). Since $A' \subseteq K'$ and $K' \subseteq K$ we have $A' \subseteq K$. Since $A \subseteq K$ and $A' \subseteq K$ we have $A \cup A' \subseteq K$, as required. This completes the proof of Part (3).

1.31 Theorem:

Let $A \subseteq P \subseteq \mathbb{R}^n$

1. A is open in P if and only if $A = U \cap P$ for some set $U \subseteq \mathbb{R}^n$ which is open in $\mathbb{R}^n$.
2. A is closed in P if and only if $A = K \cap P$ for some set $K \subseteq \mathbb{R}^n$ which is closed in $\mathbb{R}^n$.

Proof:

1. Suppose A is open in P. For each $a \in A$. Choose $r_a > 0$ such that $B_P(a, r_a) \subseteq A$, that is $B(a, r_a) \cap P \subseteq A$.
   Let $U = \cup_{a \in A} B(a, r_a)$
   Note that U is open in $\mathbb{R}^n$, because it is a union of open sets in $\mathbb{R}^n$.
   Note that $A \subseteq U$ (since for each $a \in A, a \in B(a, r_A) \subseteq U$) and $A \subseteq P$ so $A \subseteq U \cap P$
   Also, note that $U \cap P = \left(\cup_{a \in A} B(a, r_a)\right) \cap P = \cup_{a \in A}\left(B(a, r_a) \cap P\right) = \cup_{a \in A} B_P(a, r_a) \subseteq A$
   Since $B_P(a, r_a) \subseteq A$ for all $a \in A$.

   Thus we have $A = U \cap P$.

   Suppose conversely, that $A = U \cap P$, where $U \subseteq \mathbb{R}^n$ is an open set in $\mathbb{R}^n$
   Let $a \in A$. Since $a \in A$ and $a = U \cap P \subseteq U$. We have $a \in U$.
   Since $a \in U$ and U is open in $\mathbb{R}^n$. We can choose $r > 0$ so that $B(a,r) \subseteq U$.
   Then we have $B_P(a,r) = B(a,r) \cap P \subseteq U \cap P = A$
   That A is open in P.

   End of Part 1.

We constuct $A^o = U$

Note: U here is just another way of describing A.

U here acts like a bridge that get the openness in $\mathbb{R}^n$
To A

Openness of U -> Openness of A in P. (Take the intersection)

1.32 Theorem: Let $A \subseteq P \subseteq \mathbb{R}^n$. Define A to be **connected in** P when there do not exists set $E, F \subseteq P$ which are open in P and which separate A. Define A to be **compact in** P when for every set S of open sets in P such that $A \subseteq$

$\cup S$ there exists a finite subset $T \subseteq S$ such that $A \subset \cup T$. Then

A is connected in $P \leftrightarrow A$ is connected in R^n

Proof:

For the purpose of this theorem alone, we define A to be

Suppose that A is disconnected in $\mathbb{R}^n$. Let U and V be open sets in $\mathbb{R}^n$ which separate A.
Let $E = U \cap P$ and $F = V \cap P$.
Note that E and F are open in $\mathbb{R}^n$.
Verify that E and F separate A

Suppose conversely that A is disconnected in P. Choose open sets $E, F \subseteq P$ which are open in P and separate A.

Choose open sets U and V in $\mathbb{R}^n$
Such that $E = U \cap P$ and $F = V \cap P$.

We have $U \cap A \supseteq E \cap A \neq \emptyset, V \cap A \supseteq F \cap A \neq \emptyset, A \subseteq E \cup F \subseteq U \cup V$, but we might have $U \cap V \neq \emptyset$

For each $a \in U$, choose $r_a > 0$ so that $B_P(a, 2r_a) \subseteq U$ (we can do this since U is open in $\mathbb{R}^n$)

Then let $U_o = \bigcup_{a \in E} B(a, r_a)$

Note that U_o is open in $\mathbb{R}^n$ (since it is a union of open sets)
And $E \subseteq U_o$ (since each $a \in E$ lies in $B(a, r_a) \subseteq U$)

And $U_0 \cap P = \left(\bigcup_{a \in E} B(a, r)\right) \cap P = \bigcup_{a \in E}\left(B(a, r_a) \cap P\right) = \bigcup_{a \in E} B_P(a, r)$

And $E \subseteq P$ so that $E \subseteq U_0 \cap P$
And $U_0 \subseteq U$ (since each $B(a, r) \subseteq U$)

So $U_0 \cap P \subseteq U \cap P = E$

Thus $E = U_0 \cap P$

Similarly, for each $b \in V$ choose $s_b > 0$ so that $B(b, 2s_b) \subseteq V$
Then let $V_0 = \bigcup_{b \in F} B(b, s_b)$
And then we have $V \cap P = F$

We have $E \cap A \subseteq U_0 \cap A \neq \emptyset$
$\emptyset \neq F \cap A \subseteq V_0 \cap A$
And $A \subseteq E \cup F \subseteq U_0 \cup V_0$

We claim that $U_0 \cap V_0 = \emptyset$

Suppose, for a contradiction, that $U_0 \cap V_0 \neq \emptyset$
Choose $c \in U_0 \cap V_0$
$Since\ c \in U_0 = \bigcup_{a \in E} B(a, r)$
We can choose $a \in E$
Such that $c \in B(a, r_a)$
Likewise, we can choose $b \in F$
So that $c \in B(b, s_b)$

Say $r_a \leq S_b$
Then $|a - b| \leq |a - c| + |c - b| < r_a + s_b < 2s_b$
So we have $a \in B(b, 2s_b) \subseteq V$
But then $a \in P$ and $a \in V$ so $a \in V \cap P = F$
And $a \in E$
Which contradicts the fact that $E \cap F = \emptyset$

Topic summary:

Dot product
- Bilinearity
- Symmetry
- Positive Definiteness

Norm / length
- Positive Definiteness
- Scaling
- $|x \pm y|^2 = |x|^2 \pm 2(x \cdot y) + |y|^2$

- The Polarization Identities
- The Cauchy-Schwarz Inequality
- The Triangle Inequality

Distance

- Positive Definiteness
- Symmetry
- The Triangle Inequality

Angle

- $\theta(u, v) = \cos^{-1} \frac{u \cdot v}{|u||v|} \in [0, \pi]$

Orthogonal

- $u \cdot v = 0$

Sphere
Open ball
Closed ball
Punctured ball

Bounded
$A \subseteq B(a, r)$ for some $a \in \mathbb{R}^n$ and some $0 < r < \mathbb{R}$.
Open
For every $a \in A$, there exists $r > 0$ such that $B(a, r) \subseteq A$,
Closed
$A^c = \mathbb{R}^n \setminus A$ is open in $\mathbb{R}^n$.
Basic Properties of Open Sets
Basic Properties of Closed Sets
Interior
The interior of A is the largest open set which is contained in A.
Closure
The closure of A is the smallest closed set which contains A. In other words, $A \subseteq \bar{A}$ and $\bar{A}$ is closed, and for every closed set K with $A \subseteq K$ we have $\bar{A} \subseteq K$

1. 17 Corollary

1. $\left(A^0\right)^0 = A^0$ and $\bar{\bar{A}} = \bar{A}$.
2. A is open if and only if $A = A^0$
3. A is closed if and only if $A = \bar{A}$

Interior point
Some $r > 0, B(a, r) \subseteq A$
Limit point
$B^*(a, r) \cap A \neq \emptyset$
A'
Boundary point
$B(a, r) \cap A \neq \emptyset, B(a, r) \cap A^c \neq \emptyset$.
∂A

Separate

Connected
Disconnected

Open ball in P
Closed ball in P

Connected in P

Function

Domain

Graph

Explicit

Null set

Level set
Inverse image
Implicit

Range / Image

Parametrically

Parameters

Affine space

Dimension
Orthogonal

Linear
Affine

chap3

2019年5月29日 2:21

chap3

Chapter 3. Limits and Continuity

3.1 Definition: For $p \in \mathbf{Z}$, let $\mathbf{Z}_{\geq p} = \{n \in \mathbf{Z} | n \geq p\} = \{p, p+1, p+2, \cdots\}$. For a set A, a **sequence** in A is a function $a : \mathbf{Z}_{\geq p} \to A$ for some $p \in \mathbf{Z}$. We write $(a_n)_{n \geq p}$ to denote the sequence $a : \mathbf{Z}_{\geq p} \to A$ given by $a(n) = a_n$, where $a_n \in A$ for all $n \geq p$. A **subsequence** of the sequence $(a_n)_{n \geq p}$ is a sequence of the form $(b_k)_{k \geq q}$ with $b_k = a_{n_k}$ for some $p \leq n_k < n_{k+1}$ for all $k \geq q$.

3.2 Definition: Let $(a_n)_{n \geq p}$ be a sequence in $\mathbf{R}^m$. We say the sequence $(a_n)_{n \geq p}$ is **bounded** when

$$\exists r > 0 \; \forall n \in \mathbf{Z}_{\geq p} \; |a_n| \leq r.$$

For $b \in \mathbf{R}^m$, we say that the sequence $(a_n)_{n \geq p}$ **converges to** b and write $\lim_{n \to \infty} a_n = b$ (or $a_n \to b$) when

$$\forall \epsilon > 0 \; \exists N \in \mathbf{Z}_{\geq p} \; \forall n \in \mathbf{Z}_{\geq p} (n \geq N \implies |a_n - b| < \epsilon).$$

We say the sequence $(a_n)_{n \geq p}$ **diverges to** ∞ and write $\lim_{n \to \infty} a_n = \infty$ (or $a_n \to \infty$) when

$$\forall r > 0 \; \exists N \in \mathbf{Z}_{\geq p} \; \forall n \in Z_{\geq p} (n \geq N \implies |a_n| \geq r).$$

We say that the sequence $(a_n)_{n \geq p}$ **converges** when it converges to some point $b \in \mathbf{R}^m$ and otherwise we say that it **diverges**.

3.3 Theorem: *(Convergent Sequences are Bounded) Let $(a_n)_{n \geq p}$ be a sequence in $\mathbf{R}^m$. If $(a_n)_{n \geq p}$ converges in $\mathbf{R}^m$ then $(a_n)_{n \geq p}$ is bounded.*

Proof: Suppose that $(a_n)_{n \geq p}$ converges in $\mathbf{R}^m$. Let $u = \lim_{n \to \infty} a_n \in \mathbf{R}^m$. Choose $N \geq p$ such that $n \geq N \implies |a_n - u| < 1$. For $n \geq N$, by the Triangle Inequality we have $|a_n| \leq |a_n - u| + |u| < 1 + |u|$. Thus we can choose $r = \max\{|a_p|, |a_{p+1}|, \cdots, |a_{N-1}|, 1 + |u|\}$ to obtain $|a_n| \leq r$ for all $n \geq p$, and so the sequence $(a_n)_{n \geq p}$ is bounded, as required.

3.4 Theorem: *(Uniqueness of Limits of Sequences) Let $(a_n)_{n \geq p}$ be a sequence in $\mathbf{R}^m$ and let $u, v \in \mathbf{R}^m \cup \{\infty\}$. If $\lim_{n \to \infty} a_n = u$ and $\lim_{n \to \infty} a_n = v$ then $u = v$.*

Proof: We prove the theorem in the case that $u, v \in \mathbf{R}^m$ and leave the case that $u = \infty$ or $v = \infty$ as an exercise. Suppose that $\lim_{n \to \infty} a_n = u \in \mathbf{R}^m$ and $\lim_{n \to \infty} a_n = v \in \mathbf{R}^m$. Suppose, for a contradiction, that $u \neq v$. Choose $N_1 \geq p$ such that $n \geq N_1 \implies |a_n - u| < \frac{|u-v|}{2}$ and choose $N_2 \geq p$ such that $n \geq N_2 \implies |a_n - v| < \frac{|u-v|}{2}$. Let $N = \max\{N_1, N_2\}$. For $n \geq N$ we have $|u - v| \leq |u - a| + |a - v| < \frac{|u-v|}{2} + \frac{|u-v|}{2} = |u - v|$ which is impossible. Thus we must have $u = v$, as required.

3.5 Theorem: *(Limits of Subsequences) Let $(a_n)_{n \geq p}$ be a sequence in $\mathbf{R}^m$ and let $(a_{n_k})_{k \geq q}$ be a subsequence of $(a_n)_{n \geq p}$. If $\lim_{n \to \infty} a_n = u \in \mathbf{R}^m \cup \{\infty\}$ then $\lim_{k \to \infty} a_{n_k} = u$.*

Proof: We give the proof in the case that $u \in \mathbf{R}^m$. Suppose that $\lim_{n \to \infty} a_n = u \in \mathbf{R}^m$ and let $(a_{n_k})_{k \geq q}$ be any subsequence of (a_n). Let $\epsilon > 0$. Choose $N \geq p$ such that $n \geq N \implies |a_n - u| < \epsilon$. Choose $M \geq q$ such that $k \geq M \implies n_k \geq N$ (we can do this since each $n_k \in \mathbf{Z}$ with $n_k < n_{k+1}$ and hence $n_k \to \infty$ as $k \to \infty$). Then for $k \geq M$ we have $n_k \geq N$ and so $|a_{n_k} - u| < \epsilon$. Thus $\lim_{k \to \infty} a_{n_k} = u$, as required.

1

We say that $(x_n)_{n \geq p}$ is Cauchy when
$\forall \epsilon > 0 \; \exists N \in \mathbb{Z}_{\geq p}$
$\forall k, \ell \in \mathbb{Z}_{\geq p}$
$k, \ell \geq N \Rightarrow |x_k - x_\ell| < \epsilon$

3.6 Remark: It follows from the above theorem that the initial index p of a sequence $(a_n)_{n\geq p}$ does not effect whether or not the sequence converges, and it does not influence the value of the limit. For this reason, we often omit the initial index p from our notation and denote the sequence $(a_n)_{n\geq p}$ simply as (a_n).

3.7 Definition: Let $(a_n)_{n\geq p}$ be a sequence in $\mathbf{R}^n$. For $n \geq p$ let $a_n = (a_{n,1}, a_{n,2}, \cdots, a_{n,m})$. For each index k with $1 \leq k \leq m$, the k^{th} **component sequence** of $(a_n)_{n\geq p}$ is the sequence $(a_{n,k})_{n\geq p} = (a_{p,k}, a_{p+1,k} \cdots)$. Note that the sequence $(a_n)_{n\geq p}$ in $\mathbf{R}^m$ determines and is determined by its component sequences $(a_{n,k})_{n\geq p}$.

3.8 Theorem: *(Limits of Component Sequences) Let $(a_n)_{n\geq p}$ be a sequence in $\mathbf{R}^m$, say $a_n = \big(a_{n,1}, a_{n,2}, \cdots, a_{n,m}\big) \in \mathbf{R}^m$.*

(1) $(a_n)_{n\geq p}$ is bounded if and only if $(a_{n,k})_{n\geq p}$ is bounded for all indices k.
(2) For $b = (b_1, \cdots, b_m) \in \mathbf{R}^m$ we have $\lim\limits_{n\to\infty} a_n = b$ if and only if $\lim\limits_{n\to\infty} a_{n,k} = b_k$ for all k.

$(x_n)_{n\geq p}$ converges $\leftrightarrow$ $(x_{n,k})_{n\geq p}$ converges for each index $1 \leq k \leq m$
And for $u \in \mathbb{R}^m$ with $u = (u_1, u_2, \dots, u_m)$
$\lim\limits_{n\to\infty} x_n = u \leftrightarrow \lim\limits_{n\to\infty} x_{n,k} = u_k$ for all $1 \leq k \leq m$.

Proof: Suppose that $(a_n)_{n\geq p}$ is bounded. Choose $r > 0$ such that $|a_n| \leq r$ for all $n \geq p$. Let $n \geq p$ and let $1 \leq k \leq m$. Then $|a_{n,k}| \leq |a_n| \leq r$ and so the sequence $(a_{n,k})_{n\geq p}$ is also bounded. Now suppose, conversely, that $(a_{n,k})_{n\geq p}$ is bounded for all indices k. For each k, chose $r_k > 0$ such that $|a_{n,k}| \leq r_k$ for all $n \geq p$. Let $r = r_1 + \cdots + r_m$. Then for all $n \geq p$, by the Triangle Inequality we have $|a_n| \leq |a_{n,1}| + |a_{n,2}| + \cdots + |a_{n,m}| < r_1 + r_2 + \cdots + r_m = r$ and so the sequence $(a_n)_{n\geq p}$ is bounded. This proves Part (1).

Proof:

Let

To prove Part (2), suppose first that $\lim\limits_{n\to\infty} a_n = b$. Let $\epsilon > 0$ and choose $N \geq p$ so that $n \geq N \Longrightarrow \big|a_n - b\big| < \epsilon$. Let $1 \leq k \leq m$. For $n \geq N$ we have $\big|a_{n,k} - b_k\big| \leq |a_n - b| < \epsilon$ and so $\lim\limits_{n\to\infty} a_{n,k} = b_k$. Now suppose, conversely, that $\lim\limits_{n\to\infty} a_{n,k} = b_k$ for all indices k. Let $\epsilon > 0$. For each index k, choose $N_k \geq p$ such that $n \geq N_k \Longrightarrow \big|a_{n,k} - b_k\big| < \frac{\epsilon}{m}$. Then for $n \geq N$, by the Triangle Inequality we have $|a_n - b| \leq \sum\limits_{k=1}^{m} |a_{n,k} - b_k| < \epsilon$ and so $\lim\limits_{n\to\infty} a_n = b$.

$N = \max(N_1, \dots, N_m)$

$|y| = \left|\sum\limits_{k=1}^{m} y_k e_k\right| \leq \sum\limits_{k=1}^{m} |y_k e_k|$
e^k are the standard basis vectors
By triangle inequality.

3.9 Theorem: *(Operations on Limits of Sequences) Let (a_n) and (b_n) be sequences in $\mathbf{R}^m$ and let $c \in \mathbf{R}$. Suppose that $\lim\limits_{n\to\infty} a_n = u \in \mathbf{R}^n$ and $\lim\limits_{n\to\infty} b_n = v \in \mathbf{R}^n$. Then*

(1) $\lim\limits_{n\to\infty} (a_n + b_n) = u + v$,
(2) $\lim\limits_{n\to\infty} (c\, a_n) = c\, u$,
(3) $\lim\limits_{n\to\infty} |a_n| = |u|$,
(4) $\lim\limits_{n\to\infty} (a_n \cdot b_n) = u \cdot v$, and
(5) if $m = 3$ then $\lim\limits_{n\to\infty} (a_n \times b_n) = u \times v$.

Proof: These follow easily from Part (2) of the above theorem and from known properties of sequences in $\mathbf{R}$. For eample, to prove Part (1), note that

$$\lim_{n\to\infty} (a_n + b_n)_k = \lim_{n\to\infty} (a_{n,k} + b_{n,k}) = \lim_{n\to\infty} a_{n,k} + \lim_{n\to\infty} b_{n,k} = u_k + v_k = (u + v)_k.$$

3.10 Theorem: *(Sequential Characterization of Limit Points) Let $A \subseteq \mathbf{R}^m$ and let $a \in \mathbf{R}^m$. Then $a \in A'$ if and only if there exists a sequence (x_n) in $A \setminus \{a\}$ such that $\lim_{n\to\infty} x_n = a$.*

Tail sequence Thm

Proof: Let $a \in A'$. For each $n \in \mathbf{Z}^+$, since $a \in A'$ we have $B^*\big(a, \frac{1}{n}\big) \cap A \neq \emptyset$ so we can choose an element $x_n \in B^*\big(a, \frac{1}{n}\big) \cap A$ and then we have $x_n \in A \setminus \{a\}$ and $|x_n - a| < \frac{1}{n}$. Given $\epsilon > 0$ we can choose a positive integer $N > \frac{1}{\epsilon}$ and then we have $n \geq N \Longrightarrow |x_n - a| < \frac{1}{n} \leq \frac{1}{N} < \epsilon$. Thus $(x_n)_{n\geq 1}$ is a sequence in $A \setminus \{a\}$ with $\lim_{n\to\infty} x_n = a$.

Suppose, conversely, that $(x_n)_{n\geq p}$ is a sequence in $A \setminus \{a\}$ with $\lim_{n\to\infty} x_n = a$. Let $r > 0$. Since $\lim_{n\to\infty} x_n = a$ we can choose $N \geq p$ so that $n \geq N \Longrightarrow |x_n - a| < r$. Then we have $x_N \in A \setminus A$ and $|x_N - a| < r$ so that $x_N \in B^*(a, r)$, and hence $B^*(a, r) \neq 0$. Since $r > 0$ was arbitrary, it follows that $a \in A'$.

3.11 Theorem: *(Sequential Characterization of Closed Sets) Let $A \subseteq \mathbf{R}^m$. Then A is closed (in $\mathbf{R}^m$) if and only if every for every sequence in A which converges in $\mathbf{R}^m$, the limit of the sequence lies in A.*

Closedness

Proof: Suppose that A is closed. Let $(x_n)_{n\geq p}$ be a sequence in A which converges in $\mathbf{R}^n$. Let $a = \lim_{n\to\infty} x_n$. Suppose, for a contradiction, that $a \notin A$. Since $a \notin A$ we have $A = A \setminus \{a\}$ and so (x_n) is a sequence in $A \setminus \{a\}$. Since (x_n) is a sequence in $A \setminus \{a\}$ with $\lim_{n\to\infty} x_n = a$, we have $a \in A'$ by the Characterization of Limit Points. Since A is closed we have $A' \subseteq A$ and so $a \in A$, giving the desired contradiction.

Suppose, conversely, that for every sequence in A which converges in $\mathbf{R}^n$, the limit of the sequence lies in A. Let $a \in A'$. By the Characterization of Limit Points, we can choose a sequence (x_n) in $A \setminus \{a\}$ such that $\lim_{n\to\infty} x_n = a$. Then (x_n) is a sequence in A which converges in $\mathbf{R}^n$, and so its limit must lie in A, thus we have $a \in A$. Since $a \in A'$ was arbitrary, this proves that $A' \subseteq A$ and so A is closed.

3.12 Theorem: *(Bolzano-Weierstrass) Every bounded sequence in $\mathbf{R}^m$ has a convergent subsequence.*

Keep extracting subsequences

Write elements

Proof: For this proof, we shall label the components of an element in $\mathbf{R}^m$ using superscripts rather than subscripts, se we shall write an element $x \in \mathbf{R}^m$ as $(x^1, x^2, \cdots, x^m)$. Let (x_n) be a bounded sequence in $\mathbf{R}^m$. Then the first component sequence (x_n^1) is a bounded sequence in $\mathbf{R}$. By the Bolzano-Weierstrass Theorem for sequences in $\mathbf{R}$, we can choose a convergent subsequence $(x_{n_\ell}^1)$, where $n_1 < n_2 < \cdots$. Since the second component sequence (x_n^2) is bounded, the subsequence $(x_{n_\ell}^2)$ is also bounded so we can choose a convergent subsequence $(x_{n_{\ell_k}}^2)$, where $\ell_1 < \ell_2 < \cdots$. Note that the sequence $(x_{n_{\ell_k}}^1)$ also converges because it is a subsequence of the convergent subsequence $(x_{n_\ell}^1)$. Since the sequence (x_n^3) is bounded, the subsequence $(x_{n_{\ell_k}}^3)$ is also bounded so we can choose a convergent subsequence $(x_{n_{\ell_{k_j}}}^3)$, where $k_1 < k_2 < \cdots$. We then obtain convergent subsequences of each of the first 3 component sequences (x_n^i) for $i = 1, 2, 3$, namely the subsequences $(x_{n_{\ell_{k_j}}}^i)$. We repeat the procedure until we obtain simultaneous subsequences of all m component sequences (x_n^i), which we can combine to form a subsequence of the original sequence (x_n) in $\mathbf{R}^m$.

3.13 Definition: Let $(a_n)_{n\geq p}$ be a sequence in $\mathbf{R}^n$. We say that (a_n) is **Cauchy** when

$$\forall \epsilon > 0\ \exists N \in \mathbf{Z}_{\geq p}\ \forall k, \ell \in \mathbf{Z}_{\geq p} \Big(k, \ell \geq N \Longrightarrow |a_k - a_\ell| < \epsilon \Big).$$

3.14 Theorem: *(The Completeness of $\mathbf{R}^m$) For every sequence in $\mathbf{R}^m$, the sequence converges if and only if it is Cauchy.*

Proof: Let (x_n) be a sequence in $\mathbf{R}^n$. Suppose that (x_n) converges. Let $a = \lim_{n\to\infty} x_n$. Let $\epsilon > 0$. Choose N so that $n \geq N \Longrightarrow |x_n - a| < \frac{\epsilon}{2}$. Then for $k, \ell \geq N$ we have $|x_k - a| < \frac{\epsilon}{2}$ and $|x_\ell - a| < \frac{\epsilon}{2}$ so $|x_k - x_\ell| \leq |x_k - a| + |a - x_\ell| < \epsilon$. Thus (x_n) is Cauchy.

Now suppose that $(x_n)_{n\geq p}$ is Cauchy. Choose $N \geq p$ so that $k, \ell \geq N \Longrightarrow |x_k - x_\ell| < 1$. Then for all $k \geq N$ we have $|x_k - x_N| < 1$ hence $|x_k| \leq |x_k - x_N| + |x_N| < 1 + |x_N|$, and so (x_n) is bounded by $\max\big\{|x_p|, |x_{p+1}|, \cdots, |x_{N-1}|, 1 + |x_N|\big\}$. Choose a convergent subsequence (x_{n_k}) and let $a = \lim_{k\to\infty} x_{n_k}$. Let $\epsilon > 0$. Since (x_n) is Cauchy we can choose M so that $n, \ell \geq M \Longrightarrow |x_n - x_\ell| < \frac{\epsilon}{2}$. Since $\lim_{k\to\infty} x_{n_k} = a$ we can choose k so that $n_k \geq M$ and $|x_{n_k} - a| < \frac{\epsilon}{2}$. Then for $n \geq M$ we have $|x_n - a| \leq |x_n - x_{n_k}| + |x_{n_k} - a| < \epsilon$.

3.15 Definition: Let $A \subseteq \mathbf{R}^n$ and let $f : A \to \mathbf{R}^m$. When a is a limit point of A and $b \in \mathbf{R}^m$, we say that $f(x)$ **converges to** b as x tends to a, and we write $\lim_{x\to a} f(x) = b$ when

$$\forall \epsilon > 0\ \exists \delta > 0\ \forall x \in A \Big(0 < |x - a| < \delta \Longrightarrow |f(x) - b| < \epsilon \Big).$$

Don't have to be strict...

When a is a limit point of A, we say that $f(x)$ **diverges to** ∞ and we write $\lim_{x\to a} f(x) = \infty$ when

$$\forall r > 0\ \exists \delta > 0\ \forall x \in A \Big(0 < |x - a| < \delta \Longrightarrow |f(x)| \geq r \Big).$$

3.16 Theorem: *(Sequential Characterization of Limits) Let $f : A \subseteq \mathbf{R}^n \to \mathbf{R}^m$, let a be a limit point of A and let $u \in \mathbf{R}^m \cup \{\infty\}$. Then $\lim_{x\to a} f(x) = u$ if and only if $\lim_{n\to\infty} f(x_n) = u$ for every sequence (x_n) in $A \setminus \{a\}$ with $\lim_{n\to\infty} x_n = a$.*

Vacuously True if Not a limit Point

Proof: We give the proof in the case that $u \in \mathbf{R}^m$. Suppose first that $\lim_{x\to a} f(x) = u \in \mathbf{R}^m$. Let (x_n) be a sequence in $A \setminus \{a\}$ with $x_n \to a$. Let $\epsilon > 0$. Since $\lim_{x\to a} f(x) = u$ we can choose $\delta > 0$ so that $0 < |x - a| < \delta \Longrightarrow |f(x) - u| < \epsilon$. Since $x_n \to a$ we can choose N so that $n \geq N \Longrightarrow |x_n - a| < \delta$. For $n \geq N$ we have $|x_n - a| < \delta$ and we have $x_n \neq a$ (since $x_n \in A \setminus \{a\}$) and so $0 < |x_n - a| < \delta$ and hence $|f(x_n) - u| < \epsilon$. Thus $\lim_{n\to\infty} f(x_n) = u$, as required.

Negation of the if then statement

Suppose, conversely, that $\lim_{x\to a} f(x) \neq u$. Choose ϵ such that for every $\delta > 0$ there exists $x \in A$ such that $0 < |x - a| < \delta$ and $|f(x) - u| \geq \epsilon$. For each $n \in \mathbf{Z}^+$, choose $x_n \in A$ such that $0 < |x_n - a| < \frac{1}{n}$ and $|f(x_n) - u| \geq \epsilon$. For each n, since $0 < |x_n - a|$ we have $x_n \neq a$ so the sequence (x_n) lies in $A \setminus \{a\}$. Since $|x_n - a| < \frac{1}{n}$ for all $n \in \mathbf{Z}^+$ it follows that $x_n \to a$. Since $|f(x_n) - u| \geq \epsilon$ for all n, it follows that $\lim_{n\to\infty} f(x_n) \neq u$. Thus we have found a sequence (x_n) in $A \setminus \{a\}$ with $x_n \to a$ such that $\lim_{n\to\infty} f(x_n) \neq u$.

AC

$X: \mathbb{Z}^+ \to A \setminus \{a\}$

3.17 Note: Using the Sequential Characterization of Limits, many properties of limits of sequences immediately imply analogous properties of limits of function. We list some of these properties in the following theorems.

Axiom of Choice......

Axiom — For any set *X* of nonempty sets, there exists a choice function *f* defined on *X*.

来自 <https://en.wikipedia.org/wiki/Axiom_of_choice>

3.18 Theorem: *(Uniqueness of Limits of Functions) Let $f : A \subseteq \mathbf{R}^n \to \mathbf{R}^m$, let $a \in A'$, and let $u, v \in \mathbf{R}^m \cup \{\infty\}$. If $\lim_{x\to a} f(x) = u$ and $\lim_{x\to a} f(x) = v$ then $u = v$.*

Proof: This can be proven by imitating the proof of the Uniqueness of Limits of Sequences. Alternatively, we can use Uniqueness of Limits of Sequences together with the Sequential Characterization of Limits as follows. Since $a \in A'$ we can choose a sequence $(x_n) \in A \setminus \{a\}$ such that $x_n \to a$. By the Sequential Characterization of Limits, since $\lim_{x\to a} f(x) = u$ we have $\lim_{n\to\infty} f(x_n) = u$ and since $\lim_{x\to a} f(x) = v$ we have $\lim_{n\to\infty} f(x_n) = v$. By the Uniqueness of Limits of Sequences, since $\lim_{n\to\infty} f(x_n) = u$ and $\lim_{n\to\infty} f(x_n) = v$ it follows that $u = v$.

3.19 Theorem: *(Local Determination of Limits of Functions) Let $A \subseteq \mathbf{R}^n$, let $a \in A'$, let $B = B^*(a, r) \cap A$ with $r > 0$. Let $f : A \to \mathbf{R}^m$ and let $g : B \to \mathbf{R}^m$ and suppose that $f(x) = g(x)$ for all $x \in B$. Then $\lim_{x\to a} f(x)$ exists in $\mathbf{R}^n \cup \{\infty\}$ if an only if $\lim_{x\to a} g(x)$ exists in $\mathbf{R}^m \cup \{\infty\}$ and, in this case, the limits are equal.*

Proof: We leave the proof as an exercise.

3.20 Definition: Let $f : A \subseteq \mathbf{R}^n \to \mathbf{R}^m$. We can write $f(x) = \big(f_1(x), f_2(x), \cdots, f_m(x)\big)$ where $f_k : A \to \mathbf{R}$ for each index k. Then the function f_k is called the k^{th} component function of f. Note that $f_k = p_k \circ f$ where $p_k : \mathbf{R}^m \to \mathbf{R}$ is the k **projection map** given by $p_k(y_1, \cdots, y_k, \cdots, y_m) = y_k$.

3.21 Theorem: *(Limits of Component Functions) Let $f : A \subseteq \mathbf{R}^n \to \mathbf{R}^m$ be given by $f(x) = \big(f_1(x), \cdots, f_m(x)\big)$, let a be a limit point of A, and let $b = (b_1, b_2, \cdots, b_m) \in \mathbf{R}^m$. Then $\lim_{x\to a} f(x) = b$ if and only if $\lim_{x\to a} f_k(x) = b_k$ for all indices k.*

Proof: Suppose that $\lim_{x\to a} f(x) = b$. Let (x_n) be any sequence in $A \setminus \{a\}$ with $x_n \to a$. By the Sequential Characterization of Limits, we have $\lim_{n\to\infty} f(x_n) = b$. By Limits of Component Sequences, we have $\lim_{n\to\infty} f_k(x_n) = b_k$ for all indices k. By the Sequential Characterization of Limits again, it follows that $\lim_{x\to a} f_k(x) = b_k$ for all indices k.

Suppose conversely that $\lim f_k(x) = b_k$ for all k. Let (x_n) be any sequence in $A \setminus \{a\}$

By the Sequential Characterization of Limits, we have $\lim_{n\to\infty} f(x_n) = b$. By Limits of Component Sequences, we have $\lim_{n\to\infty} f_k(x_n) = b_k$ for all indices k. By the Sequential Characterization of Limits again, it follows that $\lim_{x\to a} f_k(x) = b_k$ for all indices k.

Suppose, conversely, that $\lim_{x\to a} f_k(x) = b_k$ for all k. Let (x_n) be any sequence in $A \setminus \{a\}$ with $x_n \to a$. By the Sequential Characterization of Limits, we have $\lim_{n\to\infty} f_k(x_n) = b_k$ for all k. By Limits of Component Sequences, we have $\lim_{n\to\infty} f(x) = b$. By the Sequential Characterization of Limits again, it follows that $\lim_{x\to a} f(x) = b$.

3.22 Theorem: *(Operations on Limits of Functions) Let $f, g : A \subseteq \mathbf{R}^n \to \mathbf{R}^m$, let $a \in A'$ and let $c \in \mathbf{R}$. Suppose that $\lim_{x\to a} f(x) = u \in \mathbf{R}^m$ and $\lim_{n\to\infty} g(x) = v \in \mathbf{R}^m$. Then*

(1) $\lim_{x\to a} (f+g)(x) = u + v$,
(2) $\lim_{x\to a} (cf)(x) = cu$,
(3) $\lim_{x\to a} |f|(x) = |u|$,
(4) $\lim_{x\to a} (f \cdot g)(x) = u \cdot v$, and
(5) when $m = 3$ we have $\lim_{x\to\infty} (f \times g)(x) = u \times v$.

Proof: This follows from Operations on Limits of Sequences, together with the Sequential Characterization of Limits.

3.23 Theorem: *(Comparison Theorem) Let $f, g : A \subseteq \mathbf{R}^n \to \mathbf{R}$ with $f(x) \le g(x)$ for all $x \in A$ and let $a \in A'$.*
(1) If $\lim_{x\to a} f(x) = u \in \mathbf{R} \cup \{\pm\infty\}$ and $\lim_{x\to a} g(x) = v \in \mathbf{R} \cup \{\pm\infty\}$ then $u \le v$.
(2) If $\lim_{x\to a} f(x) = \infty$ then $\lim_{x\to a} g(x) = \infty$.
(3) If $\lim_{x\to a} g(x) = -\infty$ then $\lim_{x\to\infty} f(x) = -\infty$.

Proof: This follows from the Comparison Theorem for Sequences in $\mathbf{R}$ together with the Sequential Characterization of Limits.

Function take Real values

3.24 Theorem: *(Squeeze Theorem) Let $f, g, h : A \subseteq \mathbf{R}^n \to \mathbf{R}$ with $f(x) \le g(x) \le h(x)$ for all $x \in A$, and let $u \in \mathbf{R} \cup \{\pm\infty\}$. If $\lim_{x\to a} f(x) = u = \lim_{x\to a} h(x)$ then $\lim_{x\to a} g(x) = u$.*

Proof: This follows from the Squeeze Theorem for Sequences in $\mathbf{R}$ together with the Sequential Characterization of Limits.

3.25 Definition: Let $A \subseteq \mathbf{R}^n$, let $B \subseteq \mathbf{R}^m$, and let $f : A \to B$. For $a \in A$, we say that f is **continuous at** a when

$$\forall \epsilon > 0\ \exists \delta > 0\ \forall x \in A\ \big(|x - a| < \delta \Longrightarrow |f(x) - f(a)| < \epsilon\big).$$

$|x - a| > 0$ is not needed
Don't need to be strict

We say that f is **continuous** (in A) when f is continuous at a for every $a \in A$. We say that f is **uniformly continuous** in A when

$$\forall \epsilon > 0\ \exists \delta > 0\ \forall a \in A\ \forall x \in A\ \big(|x - a| < \delta \Longrightarrow |f(x) - f(a)| < \epsilon\big).$$

like an expansion
x and a are symmetric.
So any two points

3.26 Theorem: *(Continuity at Limit Points and Isolated Points) Let $A \subseteq \mathbf{R}^n$ and let $f : A \to \mathbf{R}^m$.*
(1) When a is a limit point of A, f is continuous at $a \iff \lim_{x\to a} f(x) = f(a)$.
(2) When a is an isolated point of A, f is always continuous at a.

Proof: We leave the proof as an exercise.

3.27 Theorem: *(Sequential Characterization of Continuity) Let $A \subseteq \mathbf{R}^n$, let $f : A \to \mathbf{R}^n$, and let $a \in A$. Then f is continuous at a if and only if $\lim_{n\to\infty} f(x_n) = f(a)$ for every sequence $(x_n)_{n \ge p}$ in A with $\lim_{n\to\infty} x_n = a$.* Should be able to proof...

Proof: Suppose f is continuous at a. Let (x_n) be any sequence in A with $x_n \to a$. Let $\epsilon > 0$. Since f is continuous at a we can choose $\delta > 0$ so that $|x - a| < \delta \Longrightarrow |f(x) - f(a)| < \epsilon$. Since $x_n \to a$ we can choose N so that $n \ge N \Longrightarrow |x_n - a| < \delta$. Then for all $n \ge N$ we have $|x_n - a| < \delta$ hence $|f(x_n) - f(a)| < \epsilon$, and so $\lim_{n\to\infty} f(x_n) = f(a)$, as required.

Suppose that f is not continuous at a. Choose $\epsilon > 0$ such that for every $\delta > 0$ there exists $x \in A$ such that $|x - a| < \delta$ and $|f(x) - f(a)| \ge \epsilon$. For each $n \in \mathbf{Z}^+$, choose $x_n \in A$ such that $|x_n - a| < \frac{1}{n}$ and $|f(x_n) - f(a)| \ge \epsilon$. Since $|x_n - a| < \frac{1}{n}$ for all $n \in \mathbf{Z}^+$ it follows that $x_n \to a$. Since $\big|f(x_n) - f(a)\big| \ge \epsilon$ for all n, it follows that $\lim_{n\to\infty} f(x_n) \ne f(a)$. Thus we have found a sequence (x_n) in A with $x_n \to a$ such that $\lim_{x\to a} f(x_n) \ne f(a)$.

3.28 Theorem: *(Local Determination of Continuity) Let $A \subseteq \mathbf{R}^n$, let $a \in A'$, and let $B = B^*(a, r) \cap A$ where $r > 0$. Let $f : A \to \mathbf{R}^n$ and $g : B \to \mathbf{R}^n$ and suppose that $f(x) = g(x)$ for all $x \in B$. Then f is continuous at a if and only if g is continuous at a.*

Proof: The proof is left as an exercise.

surround a with a disk, nothing changes

3.29 Theorem: *(Continuity of Component Functions) Let $A \subseteq \mathbf{R}^n$ and let $f : A \to \mathbf{R}^m$. Then f is continuous at a if and only if f_k is continuous at a for every index k.*

Proof: This can be proven by imitating the proof of Continuity of Component Sequences or by using the result of Continuity of Component Sequences together with the Sequential Characterization of Continuity.

3.30 Theorem: *(Operations on Continuous Functions) Let $A \subseteq \mathbf{R}^n$, let $f, g : A \to \mathbf{R}^m$, let $a \in A$ and let $c \in \mathbf{R}$. If f and g are continuous at a then so are each of the functions $f + g$, cf, $|f|$ and $f \cdot g$, and also $f \times g$ in the case that $m = 3$.*

fg when $m = 1$ or $m = 2$ and $\mathbb{R}^2 = \mathbb{C}$

and $\frac{f}{g}$ when $m = 1$ or 2 and $g(x) \neq 0$ in A.

Proof: This follows from the Sequential Characterization of Continuity along with Operations on Limits of Sequences.

3.31 Theorem: *(Composition and Limits) Let $f : A \subseteq \mathbf{R}^n \to \mathbf{R}^m$, let $g : B \subseteq \mathbf{R}^m \to \mathbf{R}^\ell$ and let $h = g \circ f : C \subseteq \mathbf{R}^n \to \mathbf{R}^\ell$ where $C = A \cap f^{-1}(B)$. Let $a \in C' \subseteq A'$ and let $b \in B'$. Suppose that $\lim_{x \to a} f(x) = b$ and $\lim_{y \to b} g(y) = c \in \mathbf{R}^\ell \cup \{\infty\}$.*

Add: $(C \neq \emptyset)$

(1) If $f(x) \neq b$ for all $x \in C \setminus \{a\}$ then $\lim_{x \to a} h(x) = c$.
(2) If $b \in B$ and g is continuous at b then $\lim_{x \to a} h(x) = g(b) = c$.

Note: For $f: A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and $a \in A$,
if $a \in A'$ then
f is continuous at a
$\leftrightarrow \lim_{x \to a} f(x) = f(a)$

if $a \notin A'$, then f is vacuously continuous at a.
(Isolated points?)

Proof: We leave the proof of Part (1) as an exercise. To prove Part (2), suppose that $b \in B$ and g is continuous at b. Note that since $b \in B'$ and g is continuous at b we have $g(b) = \lim_{y \to b} g(y) = c$ by Theorem 3.26. Let (x_n) be any sequence in $C \setminus \{a\}$ with $x_n \to a$. Since $C \subseteq A$, the sequence (x_n) also lies in $A \setminus \{a\}$. By the Sequential Characterization of Limits of Functions, since $\lim_{x \to a} f(x) = b$ we have $\lim_{n \to \infty} f(x_n) = b$. For each index n we have $x_n \in C = A \cap f^{-1}(B)$ so that $f(x_n) \in B$, and so the sequence $\big(f(x_n)\big)$ lies in B. By the Sequential Characterization of Continuity, since g is continuous at b and $f(x_n) \to b$ we have $\lim_{n \to \infty} g\big(f(x_n)\big) = g(b) = c$, that is $\lim_{n \to \infty} h(x_n) = g(b) = c$. By the Sequential Characterization of Limits, it follows that $\lim_{x \to a} h(x) = g(b) = c$.

Proof (1): Suppose $f(x) \neq b$ for any $x \in C \setminus \{a\}$
Let $\epsilon > 0$
Since $\lim_{y \to b} g(y) = c$, we can choose $\delta_1 > 0$ that for all $y \in B$
$0 < |y - b| < \delta_1 \Rightarrow |g(y) - c| < \epsilon$

Since $\lim_{x \to a} f(a) = b$, we can choose $\delta > 0$ such that for all $x \in C = A \cap f^{-1}(B)$
$0 < |x - a| < \delta \to |f(x) - b| < \delta_1$
Then for any $x \in A$ with $0 < |x - a| < \delta$ we have $|f(x) - b| < \delta$.
and since $0 < |x - a|$ we have $x \in C \setminus \{a\}$ so
$0 < |f(x) - b| < \delta_1$
and hence, by the choice of δ_1
$|g(f(x)) - c| < \epsilon$

We proved that $\forall \epsilon > 0, \exists \delta > 0, \forall x \in C$
$0 < |x - a| < \delta \Rightarrow |g(f(x)) - c| < \epsilon$
By the definition, $\lim_{x \to a} g(f(x)) = c$

3.32 Corollary: *(Composition of Continuous Functions) Let $f : A \subseteq \mathbf{R}^n \to \mathbf{R}^m$, let $g : B \subseteq \mathbf{R}^m \to \mathbf{R}^\ell$, and let $h = g \circ f : C \subseteq \mathbf{R}^n \to \mathbf{R}^\ell$ where $C = A \cap f^{-1}(B)$.*
(1) If f is continuous at $a \in A$ and g is continuous at $b = f(a) \in B$ then h is continuous at a.
(2) If f is continuous in A and g is continuous in B then h is continuous in C.

3.33 Definition: An **elementary function** is a function $f : A \subseteq \mathbf{R}^n \to \mathbf{R}^m$ which can be obtained, using the operations of addition, subtraction, multiplication, division, and composition of functions (whenever those operations are defined) from the following functions, which we call the **basic elementary functions**: and the single-variable, real-valued functions c, x^n, $x^{1/n}$, e^x, $\ln x$, $\sin x$, $\cos x$, $\tan x$, $\sin^{-1} x$, $\cos^{-1} x$ and $\tan^{-1} x$. and the k^{th} inclusion map $I_k : \mathbf{R} \to \mathbf{R}^\ell$ given by $I_k(t) = (0, \cdots, 0, t, 0, \cdots, 0) = t\, e_k$, and the k^{th} projection map $P_k : \mathbf{R}^\ell \to \mathbf{R}$ given by $P_k(x_1, \cdots, x_\ell) = x_k$.

finitely many operations

3.34 Corollary: *Elementary functions are continuous in their domains.*

E.g

$f: \mathbb{R} \to \mathbb{R}, g: \mathbb{R} \to \mathbb{R}$
$f(x) = 1$ for all x
$g(y) = \begin{cases} 0 & \text{if } y \neq 1 \\ 1 & \text{if } y = 1 \end{cases}$

$g(f(x)) = 1$ for all x

$\lim_{x \to 0} f(x) = 1, \lim_{y \to 1} g(y) = 0$

Cool!

3.35 Exercise: Show that $\lim_{(x,y) \to (0,0)} \frac{x^2 - 2y^2}{x^2 + y^2}$, $\lim_{(x,y) \to (0,0)} \frac{xy}{x^2 + y^2}$ and $\lim_{(x,y) \to (0,0)} \frac{xy^2}{x^2 + y^4}$ do not exist, and that $\lim_{(x,y) \to (0,0)} \frac{3x^2 y}{x^2 + 2y^2} = 0$ and $\lim_{(x,y) \to (0,0)} \frac{xy}{\sqrt{x^2 + y^2}} = 0$.

If $\lim_{(x,y) \to (0,0)} \frac{x^2 - 2y^2}{x^2 + y^2}$ exists, $(x, y) = (t, 0)$ limit $= 1$
$(x, y) = (0, t)$
limit = -2

Corollary: (Part 2)
If $f: A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and $g: B \subseteq \mathbb{R}^m \to \mathbb{R}^\ell$, and $C = A \cap f^{-1}(B) \neq \emptyset$ and $h = g \circ f: C \to \mathbb{R}^\ell$

then (1) if f is continuous at $a \in C$ and g is continuous at $b = f(a) \in B$ then h is continuous at a

(2) If f and g are cocntinuous, then so is g.
Corollary,
All elementary functions $f: A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ are continuous.

An elementary function $f: A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is any function which can be obtained from the basic elementary functions.

Used as a definition in more abstract space

3.36 Theorem: *(Topological Characterization of Continuity) Let $A \subseteq \mathbf{R}^n$, let $B \subseteq \mathbf{R}^m$, and let $f : A \to B$.*

Do matter — Don't matter

(1) f is continuous if and only if $f^{-1}(E)$ is open in A for every open set E in B.
(2) f is continuous if and only if $f^{-1}(F)$ is closed in A for every closed set F in B.

Inverse image of open set is open

Inverse image of closed set is closed

Proof: We prove Part (1) and leave the proof of Part (2) as an exercise. Suppose that f is continuous. Let E be an open set in B. Let $a \in f^{-1}(E)$ so we have $f(a) \in E$. Since $f(a) \in E$ and E is open in B we can choose $\epsilon > 0$ so that $B_B\big(f(a), \epsilon\big) \subseteq E$. Since f is continuous at a we can choose $\delta > 0$ so that for all $x \in A$, $|x - a| < \delta \Longrightarrow \big|f(x) - f(a)\big| < \epsilon$. Let $x \in B_A(a, \delta)$, that is let $x \in A$ with $|x - a| < \delta$. Since $x \in A$ and $f : A \to B$ we have $f(x) \in B$. Since $x \in A$ with $|x - a| < \delta$, we have and $\big|f(x) - f(a)\big| < \epsilon$. Since $f(x) \in B$ with $\big|f(x) = f(a)\big| < \epsilon$, we have $f(x) \in B_B\big(f(a), \epsilon\big) \subseteq E$ hence $x \in f^{-1}(E)$. Since $a \in B_A(a, \delta)$ was arbitrary, this shows that $B_A(a, \delta) \subseteq f^{-1}(E)$. Thus $f^{-1}(E)$ is open in A, as required.

Is a rewording of $\epsilon - \delta$ definition in terms of balls

Abstract set
Topological Space
Abstract norm
Metric Space

Suppose, on the other hand, that $f^{-1}(E)$ is open in A for every open set E in B. Let $a \in A$ and let $\epsilon > 0$. The set $E = B_B(f(a), \epsilon)$ is open in B so the set $f^{-1}(E)$ is open in A, and so we can choose $\delta > 0$ such that $B_A(a, \delta) \subseteq f^{-1}(E)$. It follows that for all $x \in B_A(a, \delta)$ we have $f(a) \in E = B_B\big(f(a), \epsilon\big)$. Equivalently, for all $x \in A$, if $|x - a| < \delta$ then $f(x) \in B$ with $\big|f(x) - f(a)\big| < \epsilon$. hus f is continuous at a. Since $a \in A$ was arbitrary, f is continuous (in its domain A).

3.37 Theorem: *(Properties of Continuous Functions) Let $\emptyset \neq A \subseteq \mathbf{R}^n$, let $B \subseteq \mathbf{R}^m$, and let $f : A \to B$ be continuous.*
(1) If A is connected then $f(A)$ is connected.
(2) If A is compact then $f(A)$ is compact.
(3) If A is compact the f is uniformly continuous on A.
(4) If A is compact and $m = 1$ then $f(x)$ attains its maximum and minimum values on A.
(5) if A is compact and f is bijective and continuous then f^{-1} is continuous.

Extreme Value Theorem

Proof: We sketch a proof for Parts (1), (2) and (4) and leave some details, along with the other two parts, as an exercise. To prove Part (1), suppose that $f(A)$ is disconnected. Choose open sets U and V in $\mathbf{R}^m$ which separate $f(A)$. Since f is continuous and U and

Contrapositive.

(5) If A is compact and f is bijective and continuous then f^{-1} is continuous.

Extreme Value Theorem

Proof: We sketch a proof for Parts (1), (2) and (4) and leave some details, along with the other two parts, as an exercise. To prove Part (1), suppose that $f(A)$ is disconnected. Choose open sets U and V in $\mathbf{R}^m$ which separate $f(A)$. Since f is continuous and U and V are open, it follows that $f^{-1}(U)$ and $f^{-1}(V)$ are open in A. Verify that $f^{-1}(U)$ and $f^{-1}(V)$ separate A, so A is disconnected.

Contrapositive.

Simply take the inverse

To prove Part (2), suppose that A is compact. Let $S = \{U_k | k \in K\}$ be an open cover of $f(A)$ (with each U_k open in $\mathbf{R}^n$). For each set $k \in K$, since U_k is open in $\mathbf{R}^m$ and f is continuous, it follows that $f^{-1}(U_k)$ is open in A. Let $T = \{f^{-1}(U_k) | k \in K\}$. Verify that T is an open cover of A (with each set $f^{-1}(U_k)$ open in A). Since A is compact, we can choose a finite subset $J \subseteq K$ such that the set $\{f^{-1}(U_j) | j \in J\}$ is an open cover of A. Verify that the set $\{U_j | j \in J\}$ is an open cover for $f(A)$, so $f(A)$ is compact.

To prove Part (4), suppose that $f : A \subseteq \mathbf{R}^n \to \mathbf{R}$ with A is compact. Since A is compact and f is continuous, $f(A)$ is compact by Part (2). Since $f(A)$ is compact, it is closed and bounded by the Heine Borel Theorem. Since $f(A)$ is bounded and non-empty (since $A \neq \emptyset$) it has a supremum and an infemum in $\mathbf{R}$. Let $u = \sup f(A)$. By the Approximation Property of the Supremum, for each $n \in \mathbf{Z}^+$ we can choose $x_n \in A$ with $u - \frac{1}{n} < f(x_n) \le u$, and it follows that $f(x_n) \to u$ and hence u is a limit point of $f(A)$. Since u is a limit point of $f(A)$ and $f(A)$ is closed, we have $u \in f(A)$. Thus we can choose $a \in A$ such that $f(a) = u = \sup f(A) = \max f(A)$, and then f attains its maximum value at $a \in A$. Similarly, we can choose $b \in A$ such that $f(b) = \inf f(A) = \min f(A)$.

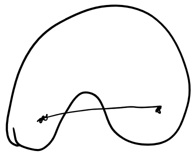
(So $\exists a \in A$ such that $f(a) \ge f(x)$ for all $x \in A$)

Squeeze or Definition

$\bar{A} = A \cup A'$

3.38 Definition: Let $A \subseteq \mathbf{R}^n$ and let $a, b \in A$. A (continuous) **path** from a to b in A is a continuous function $f : [0,1] \to A$ with $f(0) = a$ and $f(1) = b$. We say that A is **path-connected** when for every $a, b \in A$ there exists a continuous path from a to b in A.

infinite



not convex

3.39 Theorem: *(Path-Connectedness and Connectedness) Let $A \subseteq \mathbf{R}^n$.*
(1) If A is path-connected then A is connected.
(2) If A is open and connected then A is path-connected.

Path-connected is stronger than connected but usually easy to prove?

Proof: We prove Part (1) and leave Part (2) as an exercise. Suppose that A is path connected and suppose, for a contradiction, that A is not connected. Let U and V be open sets in $\mathbf{R}^n$ which separate A, that is $U \cap A \neq \emptyset$, $V \cap A \neq \emptyset$, $U \cap V = \emptyset$ and $A \subseteq U \cup V$. Choose $a \in U \cap A$ and $b \in V \cap A$. Since A is path connected we can choose a continuous path $f : [0,1] \to A$ with $f(0) = a$ and $f(1) = b$. Since f is continuous, $f^{-1}(U)$ and $f^{-1}(V)$ are open in $[0,1]$. Since $f(0) = a \in U$ we have $0 \in f^{-1}(U)$ so $f^{-1}(U) \neq \emptyset$. Similarly $1 \in f^{-1}(V)$ so $f^{-1}(V) \neq \emptyset$. Since $U \cap V = \emptyset$ we also have $f^{-1}(U) \cap f^{-1}(V) = \emptyset$ (indeed if we had $t \in f^{-1}(U) \cap f^{-1}(V)$ then we would have $f(t) \in U$ and $f(t) \in V$ so that $f(t) \in U \cap V$). Since $f : [0,1] \to A \subseteq U \cup V$ it follows that $[0,1] = f^{-1}(U) \cup f^{-1}(V)$ (indeed, given $t \in [0,1]$ we have $f(t) \in A \subseteq U \cup V$, so either $f(t) \in U$ or $f(t) \in V$ hence either $t \in f^{-1}(U)$ or $t \in f^{-1}(V)$). Thus the open sets $f^{-1}(U)$ and $f^{-1}(V)$ separate $[0,1]$. This is not possible since $[0,1]$ is connected, so we have obtained the desired contradiction.

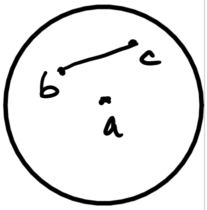
Say A is convex when for all $a, b \in A$ the line segment $[a,b] \subseteq A$ where $[a,b] = \{a + t(b-a) | 0 \le t \le 1\}$

3.40 Exercise: Show that the set $U = \{(x,y) \in \mathbf{R}^2 | y > x^2\}$ is open in $\mathbf{R}^2$.

3.41 Exercise: Show that for $a \in \mathbf{R}^n$ and $r > 0$, the set $B(a,r)$ is connected.

We show that is convex, hence path-connected, hence connected.

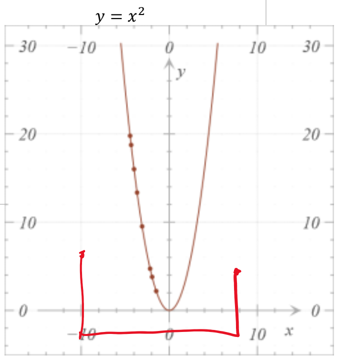
Solution:
Let $b, c \in B(a,r)$
Let $\alpha(t) = b + t(c-b) = (1-t)b + tc$ for $0 \le t \le 1$
For $0 \le t \le 1$, we have
$|\alpha(t) - a| = |b + t(c-b) - a| = |(b-a) + t((c-a) - (b-a))|$
$= |(1-t)(b-a) + t(c-a)|$
$\le (1-t)|b-a| + t|c-a|$ By Triangle Inequality, since $0 \le t \le 1$
$< (1-t)r + t \cdot r$ (subtlety here)
$= r$
So we have $\alpha(t) \in B(a,r)$ for all $0 \le t \le 1$.



$|b - a|$

eg. Let $A = \{(x,y) \in \mathbb{R}^2 | y < x^2\}$
Show that A is open and connected.

Solution:
Let $f: \mathbb{R}^2 \to \mathbb{R}$ be given by $f(x,y) = y - x^2$
Then $A = \{(x,y) | f(x,y) < 0\} = f^{-1}((-\infty, 0))$
A is open because f is continuous (its elementary) and $(-\infty, 0)$ is open.



Since not convex, not easy to do a path-connected.

Given $(a,b) \in A$, so $b < a^2$,
the map
$\alpha(t) = (a,b) + t((a,-1) - (a,b))$
$= (a,b) + t(0, -1-b)$
$= (a, (1-t)b - t)$

is continuous with $\alpha(0) = (a,b)$
and $\alpha(1) = (a,-1)$.
and for $(x,y) = (a, (1-t)b - t)$ with $0 \le t \le 1$.
We have $y = b - t(b+1) \le b < a^2 = x^2$
So that $(x,y) = \alpha(t) \in A$
hence α is a continuous path from (a,b) to $(a,-1)$ in A.
Verify that
$\beta(t) = (a,-1) + t((c,-1) - (a,-1))$
$= (a,-1) + t((c-a), 0)$

$= (a + t(c - a), -1)$
is a continuous path from $(a, -1)$ to $(c, -1)$
Also, as above, we have a continuous path from $(c, d) \in A$ to $(c, -1) \in A$
Since there is a path in A from (a, b) to $(a, -1)$ and ....
$(a, -1) \to (c, -1)$
$(c, d) \to (c, -1)$

It follows from your homework that there is a path from (a, b) to (c, d) in A.
Thus, A is path-connected, hence connected.

eg. For each of the following functions $g(x, y)$, find $\lim_{(x,y)\to(0,0)} g(x, y)$, if it exists. $g: \mathbb{R}^2 \setminus \{(0,0)\} \to \mathbb{R}$

1. $\dfrac{3x^2y^2}{x^2 + 2y^2}$

   $|g(x, y) - 0| \le \dfrac{3x^2|y|}{x^2} = 3|y| \to 0$

   Indeed, given $\epsilon > 0$
   Choose $\delta = \frac{\epsilon}{3}$, then for
   $|(x, y) - (0,0)| < \delta$
   $\Rightarrow \sqrt{x^2 + y^2} < \delta = \dfrac{\epsilon}{3}$
   $\Rightarrow |y| = \sqrt{y^2} \le \sqrt{x^2 + y^2} < \dfrac{\epsilon}{3}$
   $\Rightarrow |g(x, y) - 0| \le 3|y| < \epsilon$
   Hence $\lim_{(x,y)\to(0,0)} g(x, y) = 0$

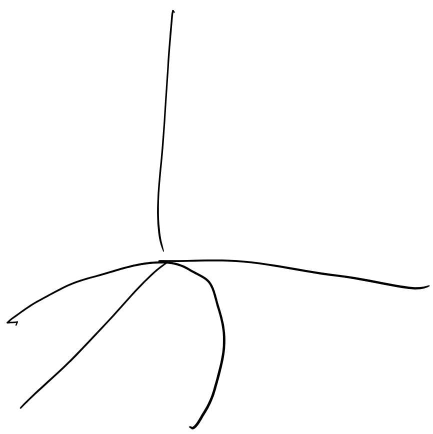
2. $\dfrac{xy}{\sqrt{x^2 + y^2}}$
   $|g(x, y) - 0| = \dfrac{|xy|}{\sqrt{x^2 + y^2}} \le \dfrac{\frac{1}{2}(x^2 + y^2)}{\sqrt{x^2 + y^2}} = \dfrac{1}{2}\sqrt{x^2 + y^2}$
   Note: $2|xy| \le x^2 + y^2$
   Thus, $\lim_{(x,y)\to(0,0)} \frac{xy}{\sqrt{x^2+y^2}} = 0$
   Indeed given $\epsilon > 0$, *choose* $\delta = 2\epsilon$
   then for (x, y) *with* $|(x, y) - (0,0)| < \delta = 2\epsilon$
   We have $\sqrt{x^2 + y^2} < 2\epsilon$
   Hence $|g(x, y) - 0| \le \frac{1}{2}\sqrt{x^2 + y^2} < \epsilon.$
3. $\dfrac{x^2 - 2y^2}{x^2 + y^2}$
   Define $\alpha: \mathbb{R} \to \mathbb{R}^2$
   if we let $\alpha(t) = (t, 0)$
   then if $\lim_{(x,y)\to(0,0)} g(x, y) = u \in \mathbb{R} \cup \{\infty\}$
   Then by the Limits and Composites Theorem
   $u = \lim_{(x,y)\to(0,0)} (x, y) = \lim_{t\to 0} g(\alpha(t)) = \lim_{t\to 0} 1 = 1$
   But if we define $\beta: \mathbb{R} \to \mathbb{R}^2$ by
   $\beta(t) = (0, t)$ then
   $u = \lim_{(x,y)\to(0,0)} g(x, y) = \lim_{t\to 0} g(\beta(t)) = \lim_{t\to 0} -\dfrac{2t^2}{t^2} = -2$
   Thus DNE

   Alternatively,
   Sequential Characterization
   Let $x_n = \left(\frac{1}{n}, 0\right)$ so $x_n \to (0,0)$
   and let $y_n = \left(0, \frac{1}{n}\right)$ so $y_n \to (0,0)$
   So by the Sequential Characterization of Limits
4. $\dfrac{xy}{x^2 + y^2}$
   Define $\alpha: \mathbb{R} \to \mathbb{R}^2$ by $\alpha(t) = (t, 0)$
   and $\beta: \mathbb{R} \to \mathbb{R}^2$ by $\beta(t) = (t, t)$
   Then if $\lim_{(x,y)\to(0,0)} g(x, y) = u$ then
   $u = 0$

   and $v = \lim_{t\to 0} g(\beta(t)) = \frac{1}{2}$
   So the limit cannot exist
5. $\dfrac{xy^2}{x^2 + y^4}$
   For fixed $y \ne 0$ and $h(x) = \frac{xy^2}{x^2+y^2}$ we have $h'(x) = y^2 \cdot \frac{(x^2+y^4)-2x^2}{(x^2+y^4)^2} = \frac{y^2(y^4-x^2)}{(y^4+x^2)^2}$
   So $h'(x) = 0 \Leftrightarrow x^2 = y^4$
   $\Leftrightarrow x = \pm y^2$
   $h(x) = \pm \dfrac{y^4}{2y^4} = \pm \dfrac{1}{2}$

For $\alpha: \mathbb{R} \to \mathbb{R}^2$ given by $(\alpha)(t) = (0, t)$
and $\beta: \mathbb{R} \to \mathbb{R}^2$ given by $\beta(t) = (t^2, t)$
if $\lim_{(x,y)\to(0,0)} g(x, y) = u$
then we must have
$u = \lim_{t\to 0} g(\alpha(t)) = 0$
and $u = \lim_{t\to 0} g(\beta(t)) = \frac{1}{2}$

Then the limit does not exist.



Sketch $z = \frac{xy}{x^2+y^2}$ and $z = \frac{xy}{\sqrt{x^2+y^2}}$ and $z = \frac{xy^2}{x^2+y^4}$

Solution: We use polar coordinate,

$x = r\cos\theta$, $y = r\sin\theta$.
Thus, $z = \frac{xy}{x^2+y^2} = \cos\theta\sin\theta = \frac{1}{2}\sin 2\theta$
$z = \dfrac{xy}{\sqrt{x^2 + y^2}} = \dfrac{1}{2} r \sin 2\theta$

chap4

2019年5月29日 2:22

chap4

Chapter 4. Introduction to Derivatives

4.1 Definition: Let $U \subseteq \mathbf{R}^n$ be open in $\mathbf{R}^n$, let $f : U \subseteq \mathbf{R}^n \to \mathbf{R}$, and let at $a \in U$, say $a = (a_1, \cdots, a_n)$. We define the k^{th} **partial derivative** of f at a to be

$$\frac{\partial f}{\partial x_k}(a) = g_k{}'(a_k) \text{ , where } g_k(t) = f(a_1, \cdots, a_{k-1}, t, a_{k+1}, \cdots a_n) \,,$$

Pretend other variables are constant

or equivalently,

$$\frac{\partial f}{\partial x_k}(a) = h_k{}'(0) \text{ , where } h_k(t) = f(a_1, \cdots, a_{k-1}, a_k + t, a_{k+1}, \cdots a_n) \,,$$

provided that the derivatives exist. Note that g_k and h_k are functions of a single variable. Sometimes $\frac{\partial f}{\partial x_k}$ is written as f_{x_k} or as f_k. When we write $u = f(x)$, we can also write $\frac{\partial f}{\partial x_k}$ as $\frac{\partial u}{\partial x_k}$, u_{x_k} or u_k. When $n = 3$ and we write x, y and z instead of x_1, x_2 and x_3, the partial derivatives $\frac{\partial f}{\partial x_1}$, $\frac{\partial f}{\partial x_2}$ and $\frac{\partial f}{\partial x_3}$ are written as $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ and $\frac{\partial f}{\partial z}$, or as f_x, f_y and f_z. When $n = 1$ so there is only one variable $x = x_1$ we have $\frac{\partial f}{\partial x}(a) = \frac{df}{dx}(a) = f'(a)$.

Eg For $f(x,y) = \frac{e^{y-x^2}}{\sqrt{2xy}}$

we have

$$\frac{\partial f}{\partial x}(x,y) = \frac{-2xe^{y-x^2}\sqrt{2xy} - e^{y-x^2}\frac{y}{\sqrt{2xy}}}{2xy} = \cdots$$

$$\frac{\partial f}{\partial y}(x,y) = e^{y-x^2}\sqrt{2xy} - = \cdots$$

4.2 Note: To calculate the partial derivative $\frac{\partial f}{\partial x_k}(x)$, we can treat the variables x_i with $i \neq k$ as constants, and differentiate f as if it were a function of the single variable x_k.

4.3 Exercise: Let $f(x,y) = x^3y + 2xy^2$. Find $\frac{\partial f}{\partial x}(1,2)$ and $\frac{\partial f}{\partial y}(1,2)$.

4.4 Exercise: Let $f(x,y,z) = (x - z^2)\sin(x^2y + z)$. Find $\frac{\partial f}{\partial x}(x,y,z)$ and $\frac{\partial f}{\partial x}(3, \frac{\pi}{2}, 0)$.

4.5 Definition: Let $U \subseteq \mathbf{R}^n$ be open in $\mathbf{R}^n$, let $f : U \subseteq \mathbf{R}^n \to \mathbf{R}^m$ and let $a \in U$. Write $u = f(x) = \big(f_1(x), f_2(x), \cdots, f_m(x)\big)^T$ with $x = (x_1, x_2, \cdots, x_n)^T$. We define the **derivative matrix**, or the **Jacobian matrix**, of f at a to be the matrix

$$Df(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \cdots & \frac{\partial f_2}{\partial x_n}(a) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \frac{\partial f_m}{\partial x_2}(a) & \cdots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix}$$

and we define the **linearization** of f at a to be the affine map $L : \mathbf{R}^n \to \mathbf{R}^m$ given by

$$L(x) = f(a) + Df(a)(x - a)$$

$x \in \mathbb{R}^n, a \in \mathbb{R}^n, f(a) \in \mathbb{R}^m, Df(a) \in M_{m\times n}(\mathbb{R})$

provided that all the partial derivatives $\frac{\partial f_k}{\partial x_l}(a)$ exist.

4.6 Definition: Let U be open in $\mathbf{R}^n$ and let $f : U \subseteq \mathbf{R}^n \to \mathbf{R}^m$. We say that f is $\mathcal{C}^1$ in U when all the partial derivatives $\frac{\partial f_k}{\partial f_l}$ exist and are continuous in U. The **second order partial derivatives** of f are the functions

$$\frac{\partial^2 f_j}{\partial x_k \partial x_l} = \frac{\partial\left(\frac{\partial f_j}{\partial x_l}\right)}{\partial x_k}.$$

We also write $\frac{\partial^2 f_j}{\partial {x_k}^2} = \frac{\partial^2 f_j}{\partial x_k \partial x_k}$. We say that f is $\mathcal{C}^2$ when all the partial derivatives $\frac{\partial^2 f_j}{\partial x_k \partial x_l}$ exist and are continuous in U. Higher order derivatives can be defined similarly, and we say f is $\mathcal{C}^k$ when all the k^{th} order derivatives $\frac{\partial^k f_j}{\partial x_{i_1} \partial x_{i_2} \cdots \partial x_{i_k}}$ exist and are continuous in U.

4.7 Definition: Let $a \in U$ where U is an open set in $\mathbf{R}$, and let $f : U \subseteq \mathbf{R} \to \mathbf{R}^m$, say $x = f(t) = \big(x_1(t), x_2(t), \cdots, x_m(t)\big)$. Then we write $f'(a) = Df(a)$ and we have

$$f'(a) = Df(a) = \begin{pmatrix} \frac{\partial x_1}{\partial t}(a) \\ \vdots \\ \frac{\partial x_m}{\partial t}(a) \end{pmatrix} = \begin{pmatrix} x_1{}'(a) \\ \vdots \\ x_m{}'(a) \end{pmatrix}.$$

The vector $f'(a)$ is called the **tangent vector** to the curve $x = f(t)$ at the point $f(a)$. In the case that t represents time and $f(t)$ represents the position of a moving point, $f'(a)$ is also called the **velocity** of the moving point at time $t = a$.

Derivative matrix -> affine map

Graph of affine map is affine space.

4.8 Definition: Let $a \in U$ where U is an open set in $\mathbf{R}^n$ and let $f : U \subseteq \mathbf{R}^n \to \mathbf{R}$. We define the **gradient** of f at a to be the vector

$$\nabla f(a) = Df(a)^T = \left(\frac{\partial f}{\partial x_1}(a), \cdots, \frac{\partial f}{\partial x_n}(a)\right)^T = \begin{pmatrix} \frac{\partial f}{\partial x_1}(a) \\ \vdots \\ \frac{\partial f}{\partial x_n}(a) \end{pmatrix}.$$

4.9 Definition: Let $U \subseteq \mathbf{R}^n$ be open in $\mathbf{R}^n$, let $f : U \subseteq \mathbf{R}^n \to \mathbf{R}^m$, and let $a \in U$. We say that f is **differentiable** at a when there exists an affine map $L : \mathbf{R}^n \to \mathbf{R}^n$ such that

$$\forall \epsilon > 0\ \exists \delta > 0\ \forall x \in U \Big(|x - a| \le \delta \Longrightarrow \big|f(x) - L(x)\big| \le \epsilon |x - a|\Big).$$

We say that f is differentiable in U when f is differentiable at every point $a \in U$.

$\exists m \in \mathbb{R}$.

True with equal sign stuck in

Linearization
The function $f(x) = f(a) + f'(a)(x - a)$ is the linearization of f at a.

Tells that when $x \approx a$, we have $f(x) \approx l(x)$.

Definition: For a function $f : U \subseteq R^n \to \mathbb{R}^m$ with U open in $\mathbb{R}^n$, and for $a \in U$, we say that f is differentiable at $a \in U$ when $\exists A \in M_{m \times n}(\mathbb{R})$
$\forall \epsilon > 0, \exists \delta > 0, \forall x \in U$
$|x - a| \le \delta \Rightarrow \big|f(x) - \big(f(a) + A(x - a)\big)\big| \le \epsilon |x - a|$

4.10 Theorem: *Let $U \subseteq \mathbf{R}^n$ be open, let $f : U \subseteq \mathbf{R}^n \to \mathbf{R}^m$ and let $a \in U$. Then*
(1) If f is differentiable at a then the partial derivatives of f at a all exist, and the affine map L which appears in the definition of the derivative is the linearization of f at a.
(2) If f is differentiable in U then f is continuous in U.
(3) If f is $\mathcal{C}^1$ in U then f is differentiable in U.
(4) If f is $\mathcal{C}^2$ in U then $\frac{\partial^2 f_j}{\partial x_k \partial x_\ell} = \frac{\partial^2 f_j}{\partial x_\ell \partial x_k}$ for all j, k, ℓ.

Proof: The proof will be given in the next chapter.

4.11 Note: Let $a \in U$ where U is open in $\mathbf{R}^n$ and let $f : U \subseteq \mathbf{R}^n \to \mathbf{R}^m$ be differentiable at a. The definition of the derivative, together with Part (1) of the above theorem, imply that the function f is approximated by its linearization near $x = a$, that is when $x \cong a$ we have

$$f(x) \cong L(x) = f(a) + Df(a)(x - a).$$

Facts,
When f is differentiable at a, the matrix A is unique.
~~Indeed, the partial derivatives $\frac{\partial f_k}{\partial x_\ell}(a)$ all exists and A is the function~~

If the partial derivative $\frac{\partial f_k}{\partial x_\ell}(x)$ exists and are continuous at a. Thus f is differnetiable of a.

The geometric objects (curves and surfaces etc) Graph(f), Null(f), $f^{-1}(k)$ and Range(f) are all approximated by the affine spaces Graph(L), Null(L), $L^{-1}(k)$ and Range(L). Each of these affine spaces is called the (affine) **tangent space** of its corresponding geometric object: the space Graph(L) is called the (affine) tangent space of the set Graph(f) at the point $(a, f(a))$; when $f(a) = 0$, the space Null(L) is called the (affine) tangent space of Null(f) at the point a, and more generally when $f(a) = k$, so that $a \in f^{-1}(k)$, the space $L^{-1}(k)$ is called the (affine) tangent space to $f^{-1}(k)$ at the point a; and the space Range(L) is called the (affine) tangent space of the set Range(f) at the point $f(a)$. When a tangent space is 1-dimensional we call it a **tangent line** and when a tangent space is 2-dimensional we call it a **tangent plane**.

$x = f(t)$
$x = L(t)$

$y = f(x)$
$y = L(x)$

$f(x) = k$
$L(x) = k$

Special Case:
When $f : U \subseteq \mathbb{R}^1 \to \mathbb{R}^m$
and we write $(x_1, x_2, \ldots, x_n) = f(t) = (x(1), \ldots, x$

and we have
$f'(a) = Df(a)$
...

Tangent vector

Derivative

The dimension of the graph $y = f(x)$ at the point $\big(a, f(a)\big)$ is the dimension of the graph $y = L(x)$. (namely n). The dimension of the level set $f'(x) = k$ at the point a is the dimension of the level set $L(x) = k$. (Namely dim Null $Df(a)$ = nullity $Df(a)$).
The dimension of the image $X = f(t)$ at the point $f(a)$ is the dimension of the image $X = L(t)$. (namely dim Col $Df(a)$ = rank $Df(a)$).

4.12 Exercise: Find an explicit, an implicit and a parametric equation for the tangent line to the curve in $\mathbf{R}^2$ which is defined explicitly by the equation $y = f(x)$, implicitly by the equation $g(x, y) = k$, and parametrically by the equation $(x, y) = \alpha(t) = \big(x(t), y(t)\big)$.

4.13 Exercise: Find an explicit, an implicit, and a parametric equation for the tangent line to the curve in $\mathbf{R}^3$ which is defined explicitly by $(x, y) = f(z) = \big(x(z), y(z)\big)$, implicitly by $u(x, y, z) = k$ and $v(x, y, z) = l$, and parametrically by $(x, y, z) = \alpha(t) = \big(x(t), y(t), z(t)\big)$.

4.14 Exercise: Find an explicit, an implicit and a parametric equation for the tangent plane to the surface in $\mathbf{R}^3$ which is defined explicitly by $z = f(x, y)$, implicitly by $g(x, y, z) = k$, and parametrically by $(x, y, z) = \sigma(s, t) = \big(x(s, t), y(s, t), z(s, t)\big)$.

4.15 Exercise: Find a parametric equation for the tangent line to the helix given by $(x, y, z) = (2\cos t, 2\sin t, t)$ at the point where $t = \frac{\pi}{3}$, and find the point where this tangent line crosses the xz-plane.

4.16 Exercise: Find an explicit equation for the tangent plane to the surface $z = \dfrac{e^{x^2+2xy}}{\sqrt{2+y}}$ at the point $(2, -1)$.

4.17 Exercise: Find an implicit equation for the tangent line to the curve given by $2\sqrt{y + x^2} + \ln(y - x^2) = 6$ at the point $(2, 5)$.

4.18 Exercise: Find a parametric equation for the tangent line to the curve of intersection of the paraboloid $z = 2 - x^2 - y^2$ with the cone $y = \sqrt{x^2 + z^2}$ at the point $p = (1, 1, 0)$.

4.19 Exercise: Find an explicit equation for the tangent plane to the surface given by $(x, y, z) = \big(r\cos t, r\sin t, \frac{3}{1+r^2}\big)$ at the point where $(r, t) = \big(\sqrt{2}, \frac{\pi}{4}\big)$.

4.20 Theorem: *(The Chain Rule) Let $f : U \subseteq \mathbf{R}^n \to V \subseteq \mathbf{R}^m$, let $g : V \subseteq \mathbf{R}^m \to \mathbf{R}^l$, and let $h(x) = g(f(x))$. If f is differentiable at a and g is differentiable at $f(a)$ then h is differentiable at a and $Dh(a) = Dg(f(a))Df(a)$.*

Proof: A proof will be given in the next chapter.

4.21 Exercise: Let $z = f(x, y) = 4x^2 - 8xy + 5y^2$, $(u, v) = g(z) = \big(\sqrt{z - 1}, 5\ln z\big)$ and $h(x, y) = g\big(f(x, y)\big)$. Find $Dh(2, 1)$.

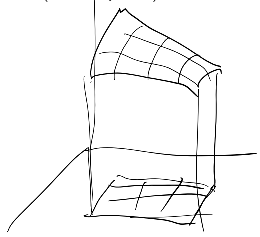
4.22 Exercise: Let $(x, y) = f(r, \theta) = \big(r\cos\theta, r\sin\theta\big)$, let $z = g(x, y)$ and let $z = h(r, \theta) = g\big(f(r, \theta)\big)$. If $h(r, \theta) = r^2 e^{\sqrt{3}(\theta - \frac{\pi}{6})}$ then find $\nabla g(\sqrt{3}, 1)$.

4.23 Exercise: Let $(x, y, z) = f(s, t)$ and $(u, v) = g(x, y, z)$. Find a formula for $\frac{\partial u}{\partial t}$.

Special Case 2:
For $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^1$

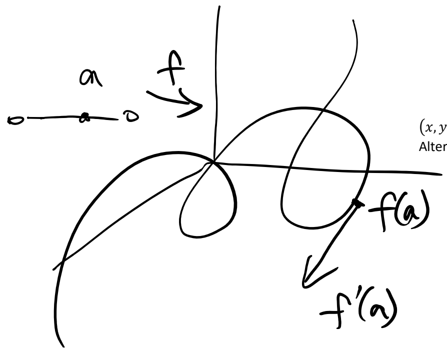
gradient.

eg. For $f : U \subseteq \mathbb{R}^2 \to \mathbb{R}$ writing $z = f(x, y)$
we have $Df(a, b) = \left(\frac{\partial f}{\partial x}(a, b), \frac{\partial f}{\partial y}(a, b)\right)$



$L(x, y) = f(a, b) + Df(a, b)\begin{pmatrix} x - a \\ y - b \end{pmatrix}$
$= f(a, b) + \frac{\partial f}{\partial x}(a, b)(x - a) + \frac{\partial f}{\partial y}(a, b)(y - b)$

eg. For $f : U \subseteq \mathbb{R}^1 \to \mathbb{R}^3$
writing $(x, y, z)^T = f(t) = \big(x(t), y(t), z(t)\big)^T$
we have $Df(a) = f'(a) = \big(x'(a), y'(a), z'(a)\big)^T$



$(x, y, z) = L(t) = f(a) + f'(a)(t - a)$
Alternatively, by $(x, y, z) = f(a) + f'(a)t$

4.24 Definition: Let $a \in U$ where U is an open set in $\mathbf{R}^n$, let $f : U \subseteq \mathbf{R}^n \to \mathbf{R}$ be differentiable at a, and let $v \in \mathbf{R}^n$. We define the **directional derivative of** f **at** a **with respect to** v, written as $D_v f(a)$, as follows: pick any differentiable curve $\alpha(t)$ with $\alpha(0) = a$ and $\alpha'(0) = v$ (for example, we could pick $\alpha(t) = a + v\,t$), and define $D_v f(a)$ to be the rate of change of the function f at $t = 0$ as we move along the curve α. To be precise, let $\beta(t) = f\big(\alpha(t)\big)$, note that $\beta'(t) = Df\big(\alpha(t)\big)\alpha'(t)$, and then define $D_v f(a)$ to be

$$\begin{aligned} D_v f(a) &= \beta'(0) \\ &= Df(\alpha(0))\,\alpha'(0) \\ &= Df(a)\,v \\ &= \nabla f(a) \cdot v\,. \end{aligned}$$

Notice that the formula for $D_v f(a)$ does not depend on the choice of the curve $\alpha(t)$. The (directional) **derivative of** f **in the direction of** v is defined to be the $D_w f(a)$ where w is the unit vector in the direction of v, that is $w = \frac{v}{|v|}$.

4.25 Exercise: Let $f(x, y, z) = x \sin(y^2 - 2xz)$ and let $\alpha(t) = \left(\sqrt{t}\,,\, \frac{1}{2}t\,,\, e^{(t-4)/4}\right)$. Find the rate of change of f as we move along the curve $\alpha(t)$ when $t = 4$.

4.26 Theorem: *Let $f : U \subseteq \mathbf{R}^n \to \mathbf{R}$ be differentiable at $a \in U$. Say $f(a) = b$. The gradient $\nabla f(a)$ is perpendicular to the level set $f(x) = b$, it is in the direction in which f increases most rapidly, and its length is the rate of increase of f in that direction.*

Proof: The proof will be given in the next chapter.

4.27 Note: Let $a \in U$ where U is an open set in $\mathbf{R}^n$, and let $f : U \subseteq \mathbf{R}^n \to \mathbf{R}^m$ be differentiable. The k^{th} column vector of the derivative matrix $Df(a)$ is the vector

$$f_{x_k}(a) = \frac{\partial f}{\partial x_k}(a) = \left(\frac{\partial f_1}{\partial x_k}(a), \cdots, \frac{\partial f_m}{\partial x_k}(a)\right)^T \in \mathbf{R}^m,$$

which is the tangent vector to the curve $\beta_k(t) = f\big(\alpha_k(t)\big)$ at $t = 0$, where α_k is the curve through a in the direction of the standard basis vector e_k given by $\alpha_k(t) = a + te_k$.

The ℓ^{th} column vector of the derivative matrix $Df(a)$ is the vector

$$\nabla f_\ell(a) = \left(\frac{\partial f_\ell}{\partial x_1}(a), \cdots, \frac{\partial f_\ell}{\partial x_n}(a)\right)^T$$

which is orthogonal to the level set $f_\ell(x) = f_\ell(a)$, and points in the direction in which f_ℓ increases most rapidly, and its length is the rate of increase of f_ℓ in that direction.

Find the tangent line at (1,1,0) to the curve of intersection of $z = 2 - x^2 - y^2$ and $y = \sqrt{x^2 + z^2}$ (equivalently, $y^2 = x^2 + z^2, y \geq 0$)

For $g: \mathbb{R}^3 \to \mathbb{R}^2$

$$\begin{pmatrix} u \\ v \end{pmatrix} = g(x, y, z) = \begin{pmatrix} x^2 + y^2 - z \\ x^2 - y^2 + z^2 \end{pmatrix}$$

The curve is the level set

$$\begin{pmatrix} u \\ v \end{pmatrix} = g(x, y, z) = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

We have $g(1,1,0) = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$

and $Dg(x, y, z) = \cdots = \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \end{pmatrix}$

$$Dg(1,1,0) = \begin{pmatrix} 2 & 2 & 1 \\ 2 & -2 & 0 \end{pmatrix}$$

$$L(x, y, z) = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

$$g(1,1,0) + Dg(1,1,0)\begin{pmatrix} x-1 \\ y-1 \\ z-0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

That is
$2(x-1) + 2(y-1) + z = 0$
$2(x-1) - 2(y-1) = 0$

or
$2x + 2y + z = 4$
$x - y = 0$

Continue at chap4_new

parametric / implicit representation not unique.

eg. For $f: U \subseteq \mathbb{R}^3 \to \mathbb{R}^1$
written as $u = f(x, y, z)$
we have

$$Df(a,b,c) = \left(\frac{\partial f}{\partial x}(a,b,c), \frac{\partial f}{\partial y}(a,b,c), \frac{\partial f}{\partial z}(a,b,c)\right) = \nabla f(a,b,c)^T$$

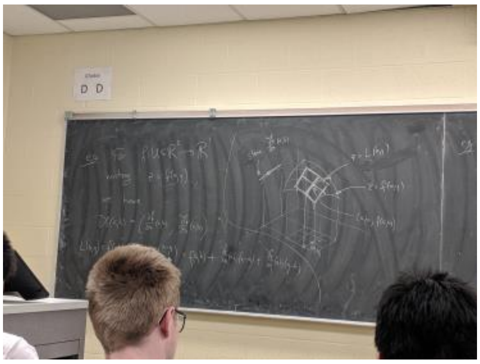
$$f(a,b,c) + Df(a,b,c)\begin{pmatrix} x-a \\ y-b \\ z-c \end{pmatrix} = k$$

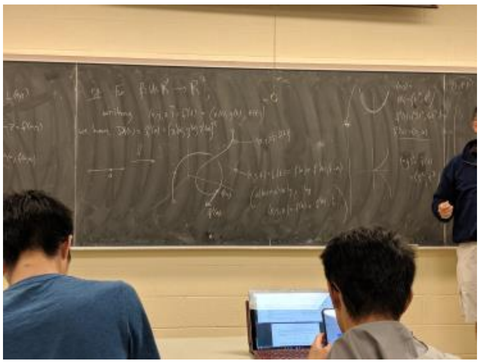
$$Df(a,b,c)\begin{pmatrix} x-a \\ y-b \\ z-c \end{pmatrix} = 0$$

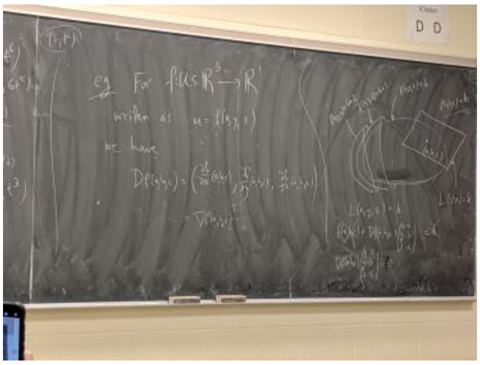
The tangent plane has equation

$$\frac{\partial f}{\partial x}(a,b,c)(x-a) + \frac{\partial f}{\partial y}(a,b,c)(y-b) + \frac{\partial f}{\partial z}(a,b,c)(z-c) = 0$$

normal vector is $\nabla f(a,b,c)$

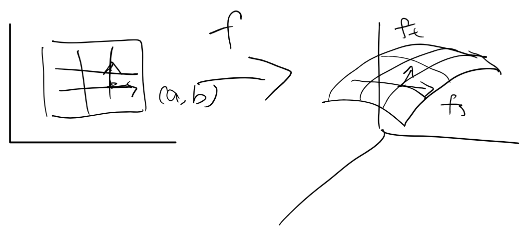






eg For $f: U \subseteq \mathbb{R}^2 \to \mathbb{R}^3$

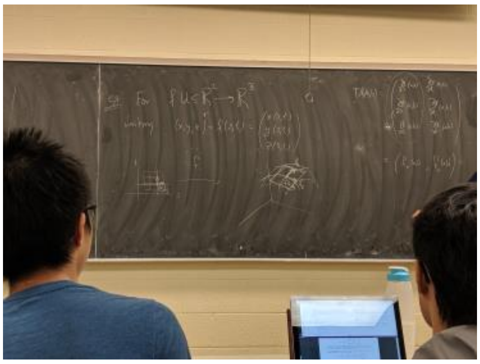
writing $(x, y, z)^T = f(s, t) = \begin{pmatrix} x(s,t) \\ y(s,t) \\ z(s,t) \end{pmatrix}$

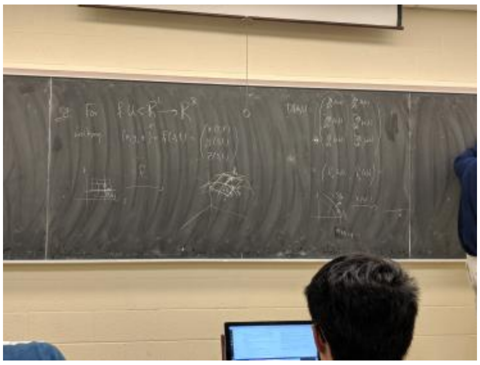


$$Df(a,b) = \begin{pmatrix} \frac{\partial x}{\partial s}(a,b) & \frac{\partial x}{\partial t}(a,b) \\ \frac{\partial y}{\partial s}(a,b) & \frac{\partial y}{\partial t}(a,b) \\ \frac{\partial z}{\partial s}(a,b) & \frac{\partial z}{\partial t}(a,b) \end{pmatrix}$$

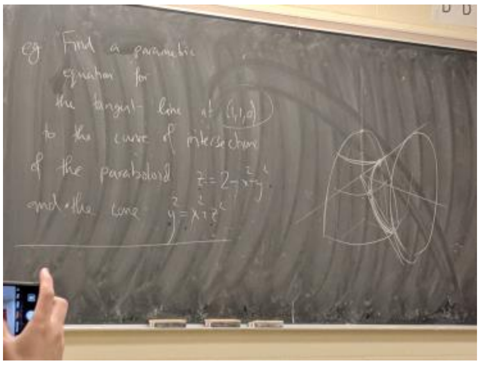
$= (f_s(a,b), f_t(a,b))$

Rows are gradient





eg. Find a parametric equation for the tangent line at (1,1,0) to the curve of intersection of two surfaces (paraboloid) $z = 2 - x^2 - y^2$ and the cone $y^2 = x^2 + z^2$



Chapter 4. Introduction to Derivatives

4.1 Definition: Let $U \subseteq \mathbf{R}^n$ be open in $\mathbf{R}^n$, let $f : U \subseteq \mathbf{R}^n \to \mathbf{R}$, and let at $a \in U$, say $a = (a_1, \cdots, a_n)$. We define the k^{th} **partial derivative** of f at a to be

$$\frac{\partial f}{\partial x_k}(a) = {g_k}'(a_k) \text{ , where } g_k(t) = f(a_1, \cdots, a_{k-1}, t, a_{k+1}, \cdots a_n) ,$$

or equivalently,

$$\frac{\partial f}{\partial x_k}(a) = {h_k}'(0) \text{ , where } h_k(t) = f(a_1, \cdots, a_{k-1}, a_k + t, a_{k+1}, \cdots a_n) ,$$

provided that the derivatives exist. Note that g_k and h_k are functions of a single variable.

Sometimes $\frac{\partial f}{\partial x_k}$ is written as f_{x_k} or as f_k. When we write $u = f(x)$, we can also write $\frac{\partial f}{\partial x_k}$ as $\frac{\partial u}{\partial x_k}$, u_{x_k} or u_k. When $n = 3$ and we write x, y and z instead of x_1, x_2 and x_3, the partial derivatives $\frac{\partial f}{\partial x_1}$, $\frac{\partial f}{\partial x_2}$ and $\frac{\partial f}{\partial x_3}$ are written as $\frac{\partial f}{\partial x}$, $\frac{\partial f}{\partial y}$ and $\frac{\partial f}{\partial z}$, or as f_x, f_y and f_z. When $n = 1$ so there is only one variable $x = x_1$ we have $\frac{\partial f}{\partial x}(a) = \frac{df}{dx}(a) = f'(a)$.

4.2 Example: Let $f(x, y) = x^3 y + 2xy^2$. Find $\frac{\partial f}{\partial x}(1, 2)$ and $\frac{\partial f}{\partial y}(1, 2)$.

Solution: Let $g_1(t) = f(t, 2) = 2t^3 + 8t$. Then $g_1'(t) = 6t^2 + 8$ so $\frac{\partial f}{\partial x}(1, 2) = g'(1) = 14$. Let $g_2(t) = f(1, t) = t + 2t^2$. Then $g_2'(t) = 1 + 4t$ so $\frac{\partial f}{\partial y}(1, 2) = g_2'(2) = 9$.

4.3 Note: Rather than explicitly determining the functions $g_k(t)$ as we did in the above solution, we can calculate the partial derivative $\frac{\partial f}{\partial x_k}(a)$ by simply treating the variables x_i with $i \neq k$ as constants, and differentiating f as if it were a function of the single variable x_k.

4.4 Example: Let $f(x, y, z) = (x - z^2)\sin(x^2 y + z)$. Find $\frac{\partial f}{\partial x}(x, y, z)$ and $\frac{\partial f}{\partial x}(3, \frac{\pi}{2}, 0)$.

Solution: Treating y and z as constants, we obtain

$$\frac{\partial f}{\partial x}(x, y, z) = \sin(x^2 y + z) + (x - z^2)\cos(x^2 y + z)(2xy)$$

and so $\frac{\partial f}{\partial x}(3, \frac{\pi}{2}, 0) = \sin\frac{9\pi}{2} + 3\cos\frac{9\pi}{2}(3\pi) = 1$.

4.5 Definition: Let $U \subseteq \mathbf{R}^n$ be open in $\mathbf{R}^n$, let $f : U \subseteq \mathbf{R}^n \to \mathbf{R}^m$ and let $a \in U$. Write $u = f(x) = \big(f_1(x), f_2(x), \cdots, f_m(x)\big)^T$ with $x = (x_1, x_2, \cdots, x_n)^T$. We define the **derivative matrix**, or the **Jacobian matrix**, of f at a to be the matrix

$$Df(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \cdots & \frac{\partial f_2}{\partial x_n}(a) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \frac{\partial f_m}{\partial x_2}(a) & \cdots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix}$$

and we define the **linearization** of f at a to be the affine map $L : \mathbf{R}^n \to \mathbf{R}^m$ given by

$$L(x) = f(a) + Df(a)(x - a)$$

provided that all the partial derivatives $\frac{\partial f_k}{\partial x_l}(a)$ exist.

4.6 Definition: Let U be open in $\mathbf{R}^n$ and let $f : U \subseteq \mathbf{R}^n \to \mathbf{R}^m$. We say that f is $\mathcal{C}^1$ in U when all the partial derivatives $\frac{\partial f_k}{\partial f_l}$ exist and are continuous in U. The **second order partial derivatives** of f are the functions

$$\frac{\partial^2 f_j}{\partial x_k \partial x_l} = \frac{\partial\left(\frac{\partial f_j}{\partial x_l}\right)}{\partial x_k}.$$

We also write $\frac{\partial^2 f_j}{\partial {x_k}^2} = \frac{\partial^2 f_j}{\partial x_k \partial x_k}$. We say that f is $\mathcal{C}^2$ when all the partial derivatives $\frac{\partial^2 f_j}{\partial x_k \partial x_l}$ exist and are continuous in U. Higher order derivatives can be defined similarly, and we say f is $\mathcal{C}^k$ when all the k^{th} order derivatives $\frac{\partial^k f_j}{\partial x_{i_1} \partial x_{i_2} \cdots \partial x_{i_k}}$ exist and are continuous in U.

4.7 Definition: Let $a \in U$ where U is an open set in $\mathbf{R}$, and let $f : U \subseteq \mathbf{R} \to \mathbf{R}^m$, say $x = f(t) = \big(x_1(t), x_2(t), \cdots, x_m(t)\big)$. Then we write $f'(a) = Df(a)$ and we have

$$f'(a) = Df(a) = \begin{pmatrix} \frac{\partial x_1}{\partial t}(a) \\ \vdots \\ \frac{\partial x_m}{\partial t}(a) \end{pmatrix} = \begin{pmatrix} {x_1}'(a) \\ \vdots \\ {x_m}'(a) \end{pmatrix}.$$

The vector $f'(a)$ is called the **tangent vector** to the curve $x = f(t)$ at the point $f(a)$. In the case that t represents time and $f(t)$ represents the position of a moving point, $f'(a)$ is also called the **velocity** of the moving point at time $t = a$.

4.8 Definition: Let $a \in U$ where U is an open set in $\mathbf{R}^n$ and let $f : U \subseteq \mathbf{R}^n \to \mathbf{R}$. We define the **gradient** of f at a to be the vector

$$\nabla f(a) = Df(a)^T = \left(\frac{\partial f}{\partial x_1}(a), \cdots, \frac{\partial f}{\partial x_n}(a)\right)^T = \begin{pmatrix} \frac{\partial f}{\partial x_1}(a) \\ \vdots \\ \frac{\partial f}{\partial x_n}(a) \end{pmatrix}.$$

4.9 Definition: Let $U \subseteq \mathbf{R}^n$ be open in $\mathbf{R}^n$, let $f : U \subseteq \mathbf{R}^n \to \mathbf{R}^m$, and let $a \in U$. We say that f is **differentiable** at a when there exists an affine map $L : \mathbf{R}^n \to \mathbf{R}^n$ such that

$$\forall \epsilon > 0\ \exists \delta > 0\ \forall x \in U \Big(|x - a| \leq \delta \Longrightarrow \big|f(x) - L(x)\big| \leq \epsilon |x - a| \Big).$$

We say that f is differentiable in U when f is differentiable at every point $a \in U$.

4.10 Theorem: *Let $U \subseteq \mathbf{R}^n$ be open, let $f : U \subseteq \mathbf{R}^n \to \mathbf{R}^m$ and let $a \in U$. Then*
(1) If f is differentiable at a then the partial derivatives of f at a all exist, and the affine map L which appears in the definition of the derivative is the linearization of f at a.
(2) If f is differentiable in U then f is continuous in U.
(3) If f is $\mathcal{C}^1$ in U then f is differentiable in U.
(4) If f is $\mathcal{C}^2$ in U then $\frac{\partial^2 f_j}{\partial x_k \partial x_\ell} = \frac{\partial^2 f_j}{\partial x_\ell \partial x_k}$ for all j, k, ℓ.

Proof: The proof will be given in the next chapter.

4.11 Note: Let $a \in U$ where U is open in $\mathbf{R}^n$ and let $f : U \subseteq \mathbf{R}^n \to \mathbf{R}^m$ be differentiable at a. The definition of the derivative, together with Part (1) of the above theorem, imply that the function f is approximated by its linearization near $x = a$, that is when $x \cong a$ we have

$$f(x) \cong L(x) = f(a) + Df(a)(x - a)\,.$$

The geometric objects (curves and surfaces etc) Graph$\,(f)$, Null$\,(f)$, $f^{-1}(k)$ and Range(f) are all approximated by the affine spaces Graph$\,(L)$, Null$\,(L)$, $L^{-1}(k)$ and Range(L). Each of these affine spaces is called the (affine) **tangent space** of its corresponding geometric object: the space Graph$\,(L)$ is called the (affine) tangent space of the set Graph$\,(f)$ at the point $\big(a, f(a)\big)$; when $f(a) = 0$, the space Null$\,(L)$ is called the (affine) tangent space of Null$\,(f)$ at the point a, and more generally when $f(a) = k$, so that $a \in f^{-1}(k)$, the space $L^{-1}(k)$ is called the (affine) tangent space to $f^{-1}(k)$ at the point a; and the space Range(L) is called the (affine) tangent space of the set Range(f) at the point $f(a)$. When a tangent space is 1-dimensional we call it a **tangent line** and when a tangent space is 2-dimensional we call it a **tangent plane**.

4.12 Example: Find an explicit, an implicit and a parametric equation for the tangent line to the curve in $\mathbf{R}^2$ which is defined explicitly by the equation $y = f(x)$, implicitly by the equation $g(x, y) = k$, and parametrically by the equation $(x, y) = \alpha(t) = \big(x(t), y(t)\big)$.

Solution: The curve in $\mathbf{R}^2$ defined explicitly by $y = f(x)$ has a tangent line at the point $(a, f(a))$ which is given explicitly by $y = L(x)$, that is

$$y = f(a) + f'(a)(x - a).$$

When $g(a, b) = k$, the curve in $\mathbf{R}^2$ defined implicitly by the equation $g(x, y) = k$ has a tangent line at the point (a, b) which is given implicitly by the equation $L(x, y) = k$, that is by $f(a, b) + \big(\frac{\partial f}{\partial x}(a, b)\,,\, \frac{\partial f}{\partial y}(a, b)\big)(x - a, y - b)^T = k$, or equivalently by

$$\tfrac{\partial f}{\partial x}(a, b)(x - a) + \tfrac{\partial f}{\partial y}(a, b)(y - b) = 0.$$

The curve in $\mathbf{R}^2$ defined parametrically by $(x, y) = \alpha(t) = \big(x(t), y(t)\big)$ or, more accurately, by $(x, y)^T = \alpha(t) = \big(x(t), y(t)\big)^T$ has a tangent line at the point $\alpha(a) = \big(x(a), y(a)\big)^T$ which is given parametrically by $(x, y)^T = L(t) = \alpha(a) + \alpha'(a)(t - a)$, that is

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x(a) \\ y(a) \end{pmatrix} + \begin{pmatrix} x'(a) \\ y'(a) \end{pmatrix}(t - a).$$

4.13 Example: Find an explicit, an implicit, and a parametric equation for the tangent line to the curve in $\mathbf{R}^3$ which is defined explicitly by $(x, y) = f(z) = \big(x(z), y(z)\big)$, implicitly by $u(x, y, z) = k$ and $v(x, y, z) = l$, and parametrically by $(x, y, z) = \alpha(t) = \big(x(t), y(t), z(t)\big)$.

Solution: The curve in $\mathbf{R}^3$ given explicitly by $(x, y) = f(z) = \big(x(z), y(z)\big)$ or, more accurately, by $(x, y)^T = f(z) = \big(x(z), y(z)\big)^T$, has a tangent plane at the point $(x(c), y(c), c)$ which is given explicitly by $(x, y)^T = L(z) = f(c) + Df(c)(z - c)$, that is by

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x(c) \\ y(c) \end{pmatrix} + \begin{pmatrix} x'(c) \\ y'(c) \end{pmatrix}(z - c)$$

When $u(a, b, c) = k$ and $v(a, b, c) = \ell$ and we write $g(x, y, z) = \big(u(x, y, z), v(x, y, z)\big)^T$, the curve in $\mathbf{R}^3$ given implicitly by $g(x, y, z) = (k, \ell)^T$, has a tangent line at (a, b, c) given implicitly by $L(x,y,z)=(k,\ell)^T$, that is b $g(a,b,c)+Dg(a,b,c)\big(x-a,y-b,z-c\big)^T=(k,\ell)^T$, or equivalently by

$$\begin{pmatrix} \frac{\partial u}{\partial x}(a,b,c) & \frac{\partial u}{\partial y}(a,b,c) & \frac{\partial u}{\partial z}(a,b,c) \\ \frac{\partial v}{\partial x}(a,b,c) & \frac{\partial v}{\partial y}(a,b,c) & \frac{\partial v}{\partial z}(a,b,c) \end{pmatrix} \begin{pmatrix} x-a \\ y-b \\ z-c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

The curve in $\mathbf{R}^3$ given parametrically by $(x,y,z)^T=\alpha(a)=\big(x(a),y(a),z(a)\big)^T$ has a tangent line at $\alpha(a)$ which is given parametrically by $(x,y,z)^T=L(t)=\alpha(a)+\alpha'(a)(t-a)$, that is

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x(a) \\ y(a) \\ z(a) \end{pmatrix} + \begin{pmatrix} x'(a) \\ y'(a) \\ z'(a) \end{pmatrix} (t-a).$$

4.14 Example: Find an explicit, an implicit and a parametric equation for the tangent plane to the surface in $\mathbf{R}^3$ which is defined explicitly by $z=f(x,y)$, implicitly by $g(x,y,z)=k$, and parametrically by $(x,y,z)=\sigma(s,t)=\big(x(s,t),y(s,t),z(s,t)\big)$.

Solution: The surface in $\mathbf{R}^3$ given explicitly by $z=f(x,y)$ has a tangent plane at the point i$g(a,b,f(a,b))$ given eplicitly by $z=L(x,y)=f(a,b)+Df(a,b)\big(x-a,y-b\big)^T$, that is

$$z=f(a,b)+\tfrac{\partial f}{\partial x}(a,b)(x-a)+\tfrac{\partial f}{\partial y}(a,b)(y-b).$$

When $g(a,b,c)=k$, the surface in $\mathbf{R}^3$ given implicitly by $g(x,y,z)=k$ has tangent plane at (a,b,c) given implicitly by $L(x,y,z)=k$, that is $g(a,b,c)+Dg(a,b,c)\big(x-a,y-b,z-c\big)^T=k$ or equivalenty

$$\tfrac{\partial g}{\partial x}(a,b,c)(x-a)+\tfrac{\partial g}{\partial y}(a,b,c)(y-b)+\tfrac{\partial g}{\partial z}(a,b,c)(z-c)=0.$$

The surface in $\mathbf{R}^3$ defined parametrically by $(x,y,z)=\sigma(s,t)=\big(x(s,t),y(s,t),z(s,t)\big)$ or, more accurately, by $(x,y,z)^T=\sigma(s,t)=\big(x(s,t),y(s,t),z(s,t)\big)^T$ has a tangent plane at $\sigma(a,b)$ which is given parametrically by $(x,y,z)^T=L(s,t)=\sigma(a,b)+D\sigma(a,b)\big(s-a,t-b\big)^T$, that is

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x(a,b) \\ y(a,b) \\ z(a,b) \end{pmatrix} + \begin{pmatrix} \frac{\partial x}{\partial s}(a,b) & \frac{\partial x}{\partial t}(a,b) \\ \frac{\partial y}{\partial s}(a.b) & \frac{\partial y}{\partial t}(a,b) \\ \frac{\partial z}{\partial s}(a,b) & \frac{\partial z}{\partial t}(a,b) \end{pmatrix} \begin{pmatrix} s-a \\ t-b \end{pmatrix}.$$

4.15 Example: Find a parametric equation for the tangent line to the helix given by $(x,y,z)=(2\cos t,2\sin t,t)$ at the point where $t=\frac{\pi}{3}$, and find the point where this tangent line crosses the xz-plane.

Solution: Let $f(t)=(2\cos t,2\sin t,t)$ and note that $f'(t)=\big(-2\sin t,2\cos t,1\big)$. We have $f\big(\frac{\pi}{3}\big)=\big(1,\sqrt{3},\frac{\pi}{3}\big)$ and $f'\big(\frac{\pi}{3}\big)=(-\sqrt{3},1,1)$ and so the tangent line at the point $f\big(\frac{\pi}{3}\big)$ is given parametrically by $(x,y,z)=L(t)=\big(1,\sqrt{3},\frac{\pi}{3}\big)+\big(-\sqrt{3},1,1\big)\big(t-\frac{\pi}{3}\big)$. The point of intersection with the xz-plane occurs when $y=0$, that is when $\sqrt{3}+t-\frac{\pi}{3}=0$, so we take $t=\frac{\pi}{3}-\sqrt{3}$ to obtain $(x,y,z)=L\big(\frac{\pi}{3}-\sqrt{3}\big)=\big(1,\sqrt{3},\frac{\pi}{3}\big)-\sqrt{3}\big(-\sqrt{3},1,1\big)=\big(4,0,\frac{\pi}{3}-\sqrt{3}\big)$.

4.16 Example: Find an explicit equation for the tangent plane to the surface $z = \dfrac{e^{x^2+2xy}}{\sqrt{2+y}}$ at the point $(2,-1)$.

Solution: Let $f(x,y) = \frac{e^{x^2+2xy}}{\sqrt{2+y}}$. Then

$$\frac{\partial f}{\partial x}(x,y) = \frac{e^{x^2+2y}(2x+2y)}{\sqrt{2+y}}$$
$$\frac{\partial f}{\partial y}(x,y) = \frac{e^{x^2+2y}(2x)\sqrt{2+y} - e^{x^2+2xy}\frac{1}{2\sqrt{2+y}}}{2+y}$$

so we have $f(2,-1)=1$, and $\frac{\partial f}{\partial x}(2,-1)=2$ and $\frac{\partial f}{\partial y}(2,-1)=\frac{7}{2}$. Thus the equation to the tangent plane is $z = 1 + 2(x-2) + \frac{7}{2}(y+1)$, or equivalently $4x+7y-2z=-1$.

4.17 Example: Find an implicit equation for the tangent line to the curve given by $2\sqrt{y+x^2} + \ln(y-x^2) = 6$ at the point $(2,5)$.

Solution: Let $g(x,y) = 2\sqrt{y+x^2} + \ln(y-x^2)$ and note that $g(2,5) = 2\sqrt{9} + \ln 1 = 6$. We have $\frac{\partial g}{\partial x}(x,y) = \frac{2x}{\sqrt{y+x^2}} - \frac{2x}{y-x^2}$ and $\frac{\partial g}{\partial y}(x,y) = \frac{1}{\sqrt{y+x^2}} + \frac{1}{y-x^2}$ so that $\frac{\partial g}{\partial x}(2,5) = \frac{4}{3} - \frac{4}{1} = -\frac{8}{3}$ and $\frac{\partial g}{\partial y}(2,5) = \frac{1}{3} + \frac{1}{1} = \frac{4}{3}$, so the tangent line at $(2,5)$ is given by $-\frac{8}{3}(x-2) + \frac{4}{3}(y-5) = 0$ or, equivalently, by $2(x-2) = (y-5)$ or by $y = 2x+1$.

4.18 Example: Find a parametric equation for the tangent line to the curve of intersection of the paraboloid $z = 2 - x^2 - y^2$ with the cone $y = \sqrt{x^2+z^2}$ at the point $p = (1,1,0)$.

Solution: Note that the paraboloid is given by $x^2+y^2+z=2$ and the cone is given by $x^2-y^2+z^2=0$, with $y \ge 0$. Let $u(x,y,z) = x^2+y^2+z$ and $v(x,y,z) = x^2-y^2+z^2$ and let $g(x,y,z) = \big(u(x,y,z), v(x,y,z)\big)^T$ so that the curve of intersection is given implicitly by $g(x,y,z) = (2,0)^T$. Note that $g(1,1,0) = (2,0)^T$ and

$$Dg(x,y,z) = \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \end{pmatrix} = \begin{pmatrix} 2x & 2y & 1 \\ 2x & -2y & 2z \end{pmatrix}$$
$$Dg(1,1,0) = \begin{pmatrix} 2 & 2 & 1 \\ 2 & -2 & 0 \end{pmatrix}$$

The tangent line at $(1,1,0)$ is given implicitly by $Dg(1,1,0)(x-1,y-1,z)^T = (0,0)^T$ that is

$$\begin{pmatrix} 2 & 2 & 1 \\ 2 & -2 & 0 \end{pmatrix}\begin{pmatrix} x-1 \\ y-1 \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

This is equivalent to the pair of equations $2(x-1)+2(y-1)+z=0$ and $2(x-1)-2(y-1)=0$. We remark that these are the equations of the tangent planes to the two given surfaces at $(1,1,0)$. The two equations are equivalent to $2x+2y+z=4$ and $x-y=0$. We let $y=t$, then the second equation gives $x=y=t$, and the first equation gives $z = 4-2x-2y = 4-4t$, so the line is given parametrically by $(x,y,z) = (0,0,4) + t(1,1,-4)$.

4.19 Exercise: Find an explicit equation for the tangent plane to the surface given by $(x,y,z) = \big(r\cos t\,,\, r\sin t\,,\, \frac{3}{1+r^2}\big)$ at the point where $(r,t) = \big(\sqrt{2}, \frac{\pi}{4}\big)$.

Alternate presentation
The paraboloid is given by $z = 2 - x^2 - y^2$

Take the tangent space (plane) first and linear algebra

$y = \sqrt{x^2+z^2}$

The tangent plane to the cone $y = \sqrt{(x^2+z^2)}$ at the point $(x,y,z) = (1,1,0)$ (that is when $(x,z) = (1,0)$) …
Solve it get $x = y$

Take the normal vector.
So it is in the direction of
$$w = u \times v = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} \times \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 4 \end{pmatrix}$$

Thus, the tangent line (which goes through $(1,1,0)^T$) is given by
$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} 1 \\ 1 \\ -4 \end{pmatrix}$$

Different version. Parameter shifted.

Alternate Solution / (Not recommended)

Substitute we get
$2x^2 + z^2 + 2 = 2$
Complete the square
$$2x^2 + \left(z + \frac{1}{2}\right)^2 = \frac{9}{4}$$
$$\frac{x^2}{\frac{9}{8}} + \frac{\left(z+\frac{1}{2}\right)^2}{\frac{9}{4}} = 1$$

The curve of intersection is given parametrically by

$x = \frac{3}{2\sqrt{2}}\cos t$
$z = -\frac{1}{2} + \frac{3}{2}\sin t$
and $y = \sqrt{x^2+z^2}$

It is the image
$(x,y,z) = \alpha(t) = (\ldots,\ldots,\ldots)$
Find s such that $\alpha(s) = (1,1,0)$
Then calculate $\alpha'(s)$
Then the equation is the tangent line is given by $(x,y,z) = \alpha(s) + \alpha'(s)\cdot t$

(or by)
$(x,y,z) = \alpha(s) + \alpha'(s)(t-s)$

4.20 Theorem: *(The Chain Rule) Let $f : U \subseteq \mathbf{R}^n \to V \subseteq \mathbf{R}^m$, let $g : V \subseteq \mathbf{R}^m \to \mathbf{R}^l$, and let $h(x) = g(f(x))$. If f is differentiable at a and g is differentiable at $f(a)$ then h is differentiable at a and $Dh(a) = Dg(f(a))Df(a)$.*

Proof: A proof will be given in the next chapter.

4.21 Exercise: Let $z = f(x,y) = 4x^2 - 8xy + 5y^2$, $(u,v) = g(z) = \big(\sqrt{z-1}\,,\, 5\ln z\big)$ and $h(x,y) = g\big(f(x,y)\big)$. Find $Dh(2,1)$.

4.22 Exercise: Let $(x,y) = f(r,\theta) = \big(r\cos\theta\,,\, r\sin\theta\big)$, let $z = g(x,y)$ and let $z = h(r,\theta) = g\big(f(r,\theta)\big)$. If $h(r,\theta) = r^2 e^{\sqrt{3}(\theta - \frac{\pi}{6})}$ then find $\nabla g(\sqrt{3},1)$.

Explanation:

For $x \approx a$, and $y = f(x) \approx f(a) = b$

$f(x) \approx f(a) + Df(a)(x-a)$
$h(x) = g\big(f(x)\big) = g(y) \approx g(b) + Dg(b)(y-b)$
$= g\big(f(a)\big) + Dg\big(f(a)\big)\big(f(x) - f(a)\big)$
$\approx h(a) + Dh\big(f(a)\big)Df(a)(x-a)$

Verify these approximate do fall in the interval

$h(x,y) = g\big(f(x,y)\big)$. Find $Dh(2,1)$.

4.22 Exercise: Let $(x,y) = f(r,\theta) = \big(r\cos\theta\,,\, r\sin\theta\big)$, let $z = g(x,y)$ and let $z = h(r,\theta) = g\big(f(r,\theta)\big)$. If $h(r,\theta) = r^2 e^{\sqrt{3}(\theta - \frac{\pi}{6})}$ then find $\nabla g(\sqrt{3},1)$.

4.23 Exercise: Let $(x,y,z) = f(s,t)$ and $(u,v) = g(x,y,z)$. Find a formula for $\frac{\partial u}{\partial t}$.

4.24 Definition: Let $a \in U$ where U is an open set in $\mathbf{R}^n$, let $f : U \subseteq \mathbf{R}^n \to \mathbf{R}$ be differentiable at a, and let $v \in \mathbf{R}^n$. We define the **directional derivative of f at a with respect to** v, written as $D_v f(a)$, as follows: pick any differentiable curve $\alpha(t)$ with $\alpha(0) = a$ and $\alpha'(0) = v$ (for example, we could pick $\alpha(t) = a + v\,t$), and define $D_v f(a)$ to be the rate of change of the function f at $t = 0$ as we move along the curve α. To be precise, let $\beta(t) = f\big(\alpha(t)\big)$, note that $\beta'(t) = Df\big(\alpha(t)\big)\alpha'(t)$, and then define $D_v f(a)$ to be

$$\begin{aligned} D_v f(a) &= \beta'(0) \\ &= Df(\alpha(0))\,\alpha'(0) \\ &= Df(a)\,v \\ &= \nabla f(a) \cdot v\,. \end{aligned}$$

Notice that the formula for $D_v f(a)$ does not depend on the choice of the curve $\alpha(t)$. The (directional) **derivative of f in the direction of** v is defined to be the $D_w f(a)$ where w is the unit vector in the direction of v, that is $w = \frac{v}{|v|}$.

4.25 Exercise: Let $f(x,y,z) = x\sin(y^2 - 2xz)$ and let $\alpha(t) = \big(\sqrt{t}\,,\, \frac{1}{2}t\,,\, e^{(t-4)/4}\big)$. Find the rate of change of f as we move along the curve $\alpha(t)$ when $t = 4$.

4.26 Theorem: *Let $f : U \subseteq \mathbf{R}^n \to \mathbf{R}$ be differentiable at $a \in U$. Say $f(a) = b$. The gradient $\nabla f(a)$ is perpendicular to the level set $f(x) = b$, it is in the direction in which f increases most rapidly, and its length is the rate of increase of f in that direction.*

Proof: The proof will be given in the next chapter.

4.27 Note: Let $a \in U$ where U is an open set in $\mathbf{R}^n$, and let $f : U \subseteq \mathbf{R}^n \to \mathbf{R}^m$ be differentiable. The k^{th} column vector of the derivative matrix $Df(a)$ is the vector

$$f_{x_k}(a) = \frac{\partial f}{\partial x_k}(a) = \left(\frac{\partial f_1}{\partial x_k}(a), \cdots, \frac{\partial f_m}{\partial x_k}(a)\right)^T \in \mathbf{R}^m,$$

which is the tangent vector to the curve $\beta_k(t) = f\big(\alpha_k(t)\big)$ at $t = 0$, where α_k is the curve through a in the direction of the standard basis vector e_k given by $\alpha_k(t) = a + te_k$.

The ℓ^{th} column vector of the derivative matrix $Df(a)$ is the vector

$$\nabla f_\ell(a) = \left(\frac{\partial f_\ell}{\partial x_1}(a), \cdots, \frac{\partial f_\ell}{\partial x_n}(a)\right)^T$$

which is orthogonal to the level set $f_\ell(x) = f_\ell(a)$, and points in the direction in which f_ℓ increases most rapidly, and its length is the rate of increase of f_ℓ in that direction.

6

$= g\big(f(a)\big) + Dg\big(f(a)\big)\big(f(x) - f(a)\big)$
$\approx h(a) + Dh\big(f(a)\big)Df(a)(x - a)$

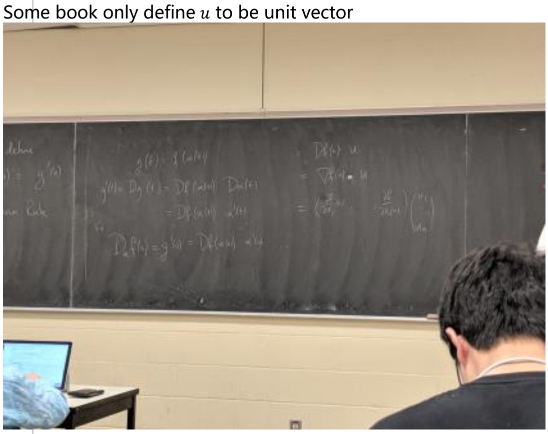
Verify these approximate do fall in the interval

Def. Let $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ be differentiable at $a \in U$ and let $u \in \mathbb{R}^n$.
We define the **directional derivative** $D_u f(a)$ of f at a with respect to u as follows.
Choose a differentiable function (path) $\alpha: (-\epsilon, \epsilon) \subseteq \mathbb{R} \to U \subseteq \mathbb{R}^n$ with $\alpha(0) = a$ and $\alpha'(0) = u$
(for example we could choose $\alpha(t) = a + ut$)

We let $g: (-\epsilon, \epsilon) \subseteq \mathbb{R} \to \mathbb{R}$
is given by $g(t) = f\left(\alpha(t_1)\right)$ and we define
$D_u f(a) = g'(0)$
By the Chain Rule
$g(t) = f\big(\alpha(t)\big)$
$g'(t) = Df\big(\alpha(t)\big)\alpha'(t)$

$D_u f(a) = g'(0) = Df\big(\alpha(0)\big)\alpha'(0) = Df(a) \cdot u = \nabla f(a) \cdot u$
When function only have one variable, usually use prime notation.

HW 4
Some book only define u to be unit vector



Proof:
When we say that $\nabla f(a)$ is perpendicular to the level set $f(x) = k$. We mean that $\nabla f(a)$ is perpendicular to the tangent vector $\alpha'(0)$ of every differentiable curve given by
$\alpha: (-\epsilon, \epsilon) \subseteq \mathbb{R} \to f^{-1}(h) = \{x \in U | f(x) = k\}$
with $\alpha(0) = a$.

Let α be any such differentiable map
Since $\alpha(t) \in f^{-1}(k)$ for all $t \in (-\epsilon, \epsilon)$
We have $f\big(\alpha(t)\big) = k$ for all $t \in (-\epsilon, \epsilon)$ [Near zero]
By the Chain Rule
$Df\big(\alpha(t)\big)\alpha'(t) = 0$
for all $t \in (-\epsilon, \epsilon)$

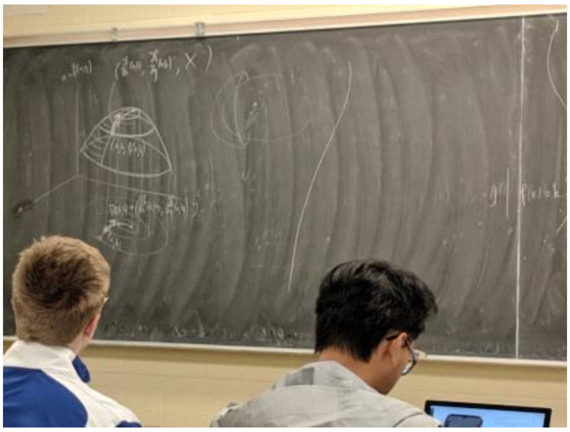
$Df\big(\alpha(0)\big)\alpha'(0) = 0$
$Df(a)\alpha'(0) = 0$
$\nabla f(a) \cdot \alpha'(0) = 0$

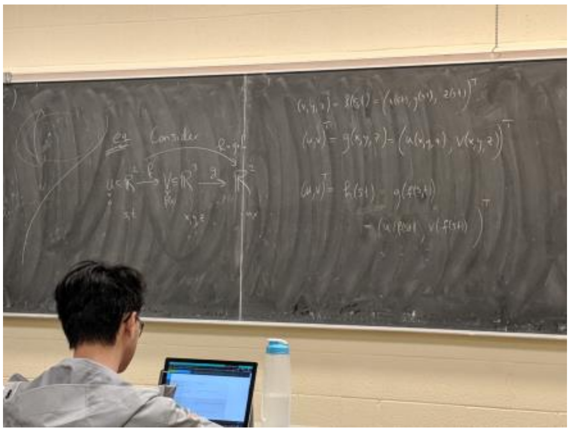
Given any unit vector $u \in \mathbb{R}^n$
If we let $\theta = \theta\big(u, \nabla f(a)\big) \in [0, \pi]$ be the angle between u and $\nabla f(a)$
(Assuming $\nabla f(a) \neq 0$).

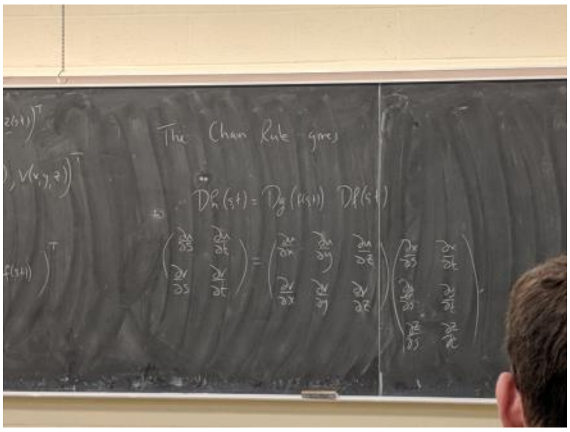
Then,
$D_u f(a) = \nabla f(a) \cdot u = |\nabla f(a)| \cos\theta$
So $D_u f(a)$ has maximum value $|\nabla f(a)|$
which is attained when $\cos\theta = 1$ that is when $\theta = 0$
that is when u is in the direction of $\nabla f(a)$

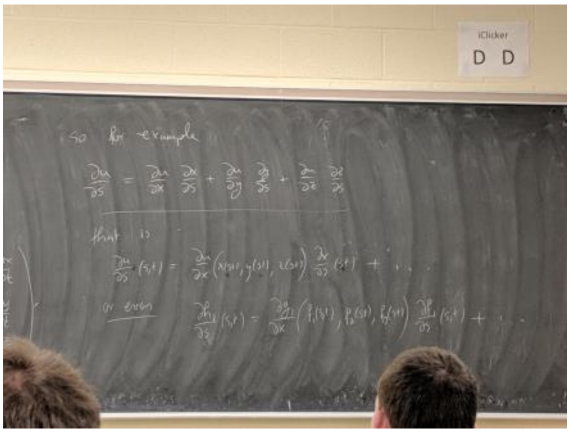
$$u = \frac{\nabla f(a)}{|\nabla f(a)|}$$

Thm:









Examples (Given some functions $f(x,y)$ determine whether they are differentiable at a point (a,b) (usually $(a,b)=(0,0)$)
eg.

For

$$f(x,y)=\begin{cases}\left(\dfrac{xy}{x^2+y^2}\right) & (x,y)\neq(0,0)\\ 0 & (x,y)=(0,0)\end{cases}$$

We saw that for $x=r\cos\theta\,,y=r\sin\theta$

$f(x,y) = \frac{1}{2}\cos 2\theta$

and f is not continuous at (0,0)

(hence as we shall prove f cannot be differentiable at (0,0)).

Also, notice that
$\frac{\partial f}{\partial x}(0,0) = g'(0)$

where $g(t) = f(t,0) = \begin{cases} \frac{0}{t^2} & t \neq 0 \\ 0 & t = 0 \end{cases} = 0$
for all t

So $g'(t) = 0$ for all t.
$\frac{\partial f}{\partial x}(0,0) = g'(0) = 0$

Similarly
$\frac{\partial f}{\partial y}(0,0) = 0;$

---

eg
Let $f(x,y) = |xy|$

Then $|xy| \leq \frac{1}{2}(x^2+y^2) = \frac{1}{2}\left(\sqrt{x^2+y^2}\right)\left(\sqrt{x^2+y^2}\right)$

We claim that f is differentiable at (0,0) with $Df(0,0) = (0,0)$

Let $\epsilon > 0$.
Choose $\delta = 2\epsilon$
Let (x,y) be arbitrary.
Suppose $|(x,y) - (0,0)| < \delta$
that $\sqrt{x^2+y^2} < \delta = 2\epsilon$

$$\left|f(x,y) - f(0,0) - (0,0)\begin{pmatrix} x-0 \\ y-0 \end{pmatrix}\right| = |f(xy)| = |xy| \leq \frac{1}{2}\sqrt{x^2+y^2}\,\sqrt{x^2+y^2} \leq \epsilon|(x,y)-(0,0)|$$

Thus, $Df(0,0) = 0 = (0,0)$

~~$\frac{x^3 - xy^2}{x^2+y^2}$~~

Change it to
$f(x,y) = \frac{x^3 - 3xy^2}{x^2+y^2}$

Eg. Let $f(x,y) = \sqrt{|xy|}$

If $f(x,y)$ was differentiable at (0,0)
then $D_{(1,1)}f(0,0)$ would exist:
For $g(t) = f(t,t) = \sqrt{t^2} = |t|$
We would have $D_{(1,1)}f(0,0) = g'(0)$
But $g'(0)$ does not exist.

(We remark that $\frac{\partial f}{\partial x}(0,0)$ and $\frac{\partial f}{\partial y}(0,0)$ both exists and are zero).

Eg. Try $f(x,y) = \begin{cases} \frac{x^3}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$

Eg. Let $f(x,y) = \begin{cases} \frac{x^3-3xy^2}{x^2+y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$

For $x = r\cos\theta, y = r\sin\theta$.
$f(x,y) = r(\cos^3\theta - 3\cos\theta(\sin\theta))$
$= r(4\cos^3\theta - 3\cos\theta)$
$= r\cos 3\theta$

Not differentiable

Not align with the tangent plane.
(Not smooth)

Suppose $f(x,y)$ is differentiable at (0,0)
For $g(t) = f(t,0) = \frac{t^3-0}{t^2+0} = t$

So $g'(t) = 1$ for all t
and $\frac{\partial f}{\partial x}(0,0) = g'(0) = 1$

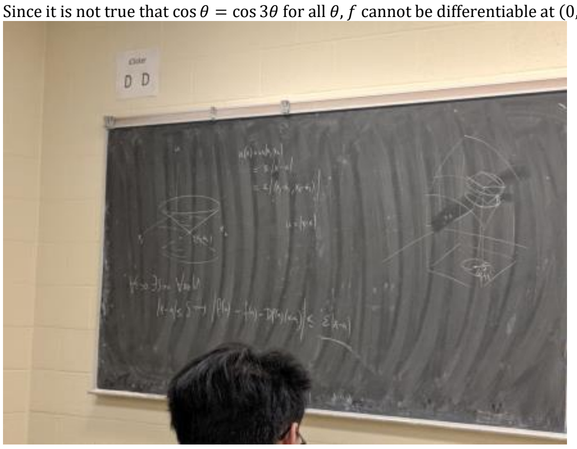
For $g(t) = f(0,t) = 0$
We have $g'(t) = 0$
So $\frac{\partial f}{\partial y}(0,0) = g'(0) = 0$
If f was differentiable at (0,0) we could have $D_{(u,v)}f(0,0) = Df(0,0)\begin{pmatrix} u \\ v \end{pmatrix} = (1,0)\begin{pmatrix} u \\ v \end{pmatrix} = u$

But for example, for fixed θ. and let $(u,v) = (\cos\theta, \sin\theta)$
For $g(t) = f(t\cos\theta, t\sin\theta) = t\cos 3\theta$
So $g'(t) = \cos 3\theta$
So we should have
$D_{(\cos\theta,\sin\theta)}f(0,0) = g'(0) = \cos 3\theta$

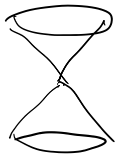
But for example, for fixed θ. and let $(u, v) = (\cos\theta, \sin\theta)$
For $g(t) = f(t\cos\theta, t\sin\theta) = t\cos 3\theta$
So $g'(t) = \cos 3\theta$
So we should have
$D_{(\cos\theta,\sin\theta)} f(0,0) = g'(0) = \cos 3\theta$

And
$D_{(\cos\theta,\sin\theta)} f(0,0) = \cos\theta$.

Since it is not true that $\cos\theta = \cos 3\theta$ for all θ, f cannot be differentiable at (0,0).



$$u(x) = \varepsilon\left|(x_1 - a_1, x_2 - a_2)\right|$$



Term Test 2 Preparation

2019年6月15日 9:56

NOTE: The MATH 247 Term Test 2 will be held on
Tuesday June 18, from 4:30-5:20 in MC 4021.

It will cover Sections 1.22-1.29, Chapter 3, Sections 4.1-4.19,
and Assignments 3 and 5 (but not Assignment 4).

You will be asked to prove 1 of the following 3 theorems.
Theorem 1.29 (The Heine Borel Theorem)
Theorem 3.12 (The Bolzano Weierstrass Theorem)
Theorem 3.37 Parts 2 and 4 (The Extreme Value Theorem)

来自 <http://www.math.uwaterloo.ca/~snew/math247-2019-S/>

chap5

Chapter 5. Differentiation

5.1 Remark: We now begin a more detailed and theoretical presentation of differentiation in Euclidean space. We repeat some of our definitions and restate our theorems in a different order and provide rigorous proofs for the theorems which were not proven earlier.

5.2 Note: Recall that for $f : U \subseteq \mathbf{R} \to \mathbf{R}$ and $a \in U$,

$$f \text{ is differentiable at } a \iff \lim_{x \to a} \frac{f(x) - f(a)}{x - a} \text{ exists}$$

$$\iff \exists m \in \mathbf{R}\ \forall \epsilon > 0\ \exists \delta > 0\ \forall x \in U\ \ 0 < |x - a| < \delta \implies \left| \frac{f(x) - f(a)}{x - a} - m \right| < \epsilon$$

$$\iff \exists m \in \mathbf{R}\ \forall \epsilon > 0\ \exists \delta > 0\ \forall x \in U\ \ 0 < |x - a| < \delta \implies |f(x) - f(a) - m(x - a)| < \epsilon\, |x - a|$$

$$\iff \exists m \in \mathbf{R}\ \forall \epsilon > 0\ \exists \delta > 0\ \forall x \in U\ \ |x - a| \le \delta \implies \big|f(x) - \big(f(a) + m(x - a)\big)\big| \le \epsilon\, |x - a|.$$

In this case, the number $m \in \mathbf{R}$ is unique, we call it the **derivative** of f at a and denote it by $f'(a)$, and the map $\ell(x) = f(a) + f'(a)(x - a)$ is called the **linearization** of f at a.

5.3 Definition: Let $f : U \subseteq \mathbf{R}^n \to \mathbf{R}^m$, where U is open. We say f is **differentiable** at $a \in U$ if there is an $m \times n$ matrix A such that

$$\forall \epsilon > 0\ \exists \delta > 0\ \forall x \in U\ \Big(|x - a| \le \delta \implies \big|f(x) - (f(a) + A(x - a))\big| \le \epsilon |x - a| \Big).$$

We show below that the matrix A is unique, we call it the **derivative** (matrix) of f at a, and we denote it by $Df(a)$. The affine map $L : \mathbf{R}^n \to \mathbf{R}^m$ given by $L(x) = f(a) + Df(a)(x - a)$, which approximates $f(x)$, is called the **linearization** of f at a. We say f is **differentiable** in U when it is differentiable at every point $a \in U$.

5.4 Example: If f is the affine map $f(x) = Ax + b$, then we have $Df(a) = A$ for all a. Indeed given $\epsilon > 0$ we can choose $\delta > 0$ to be anything we like, and then for all x we have

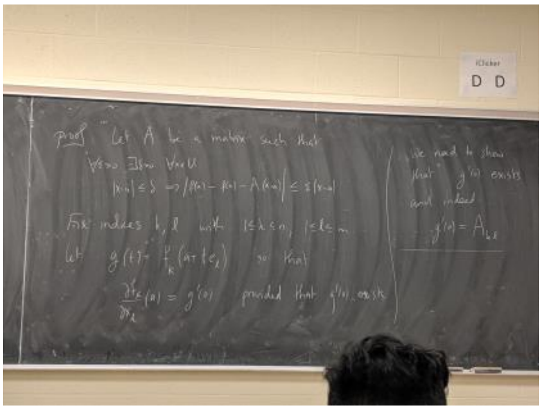
$$\big|f(x) - f(a) - A(x - a)\big| = \big|Ax + b - Aa - b - Ax + Aa\big| = 0 \le \epsilon |x - a|.$$

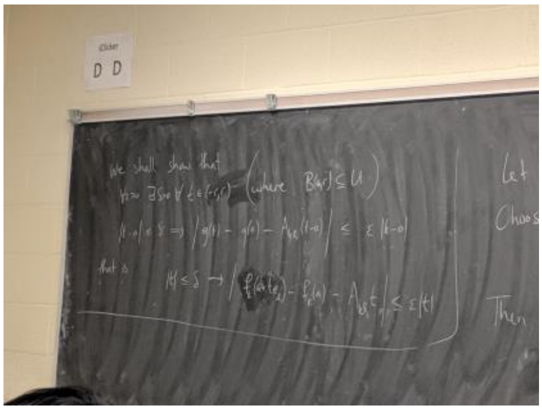
5.5 Theorem: *(The Derivative is the Jacobian) Let $f : U \subseteq \mathbf{R}^n \to \mathbf{R}^m$ and let $a \in U$. If f is differentiable at a then the partial derivatives $\frac{\partial f_k}{\partial x_\ell}(a)$ all exist and the matrix A which appears in the definition of the derivative is equal to the Jacobian matrix $Df(a)$.*

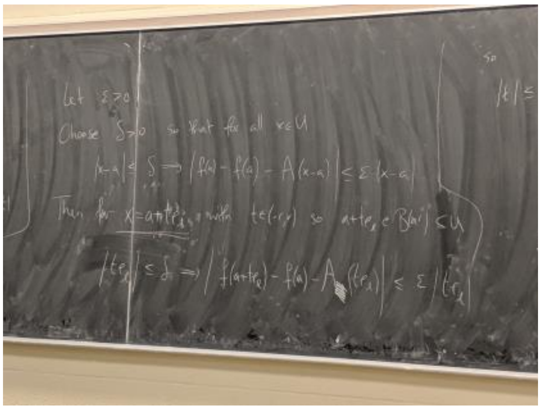
Proof: Suppose that f is differentiable at a. Fix indices k and ℓ and let $g(t) = f_k(a + te_\ell)$ so that $\frac{\partial f_k}{\partial x_\ell}(a) = g'(0)$ provided that the derivative $g'(0)$ exists. Let A be a matrix as in the definition of differentiability. Let $\epsilon > 0$. Choose $\delta > 0$ such that for all $x \in U$ with $|x - a| \le \delta$ we have $\big|f(x) - f(a) - A(x - a)\big| \le \epsilon\, |x - a|$. Let $t \in \mathbf{R}$ with $|t| \le \delta$. Let $x = a + t\, e_\ell$. Then we have $|x - a| = |te_\ell| = |t| \le \delta$ and so $\big|f(x) - f(a) - A(x - a)\big| \le \epsilon\, |x - a|$. Since for any vector $u \in \mathbf{R}^m$ we have $|u_k| \le |u|$, we have

$$\begin{aligned} \big|g(t) - g(0) - A_{k,\ell}\, t\big| &= \big|f_k(a + te_\ell) - f_k(a) - \big(A(te_\ell)\big)_k\big| \\ &\le \big|f(a + te_\ell) - f(a) - A(te_\ell)\big| \\ &= \big|f(x) - f(a) - A(x - a)\big| \\ &\le \epsilon\, |x - a| = \epsilon\, |t|. \end{aligned}$$

It follows that $A_{k,\ell} = g'(0) = \frac{\partial f_k}{\partial x_\ell}(a)$, as required.







The maximum does exist by the Extreme Value Theorem since the sphere is compact $S^{n-1} = \{x \in \mathbb{R}^n \mid |x| = 1\}$ is compact. (S^{n-1} is closed since $S^{n-1} = g^{-1}(\{1\})$ where $g: \mathbb{R}^n \to \mathbb{R}$ is the continuous function $g(x) = |x| = \sqrt{x^2}$) and since the function $f: S^{n-1} \to \mathbb{R}$ given by $f(x) = |Ax|$ is continuous.

Elementary function.

$$Ax = A\left(\sum_{l=1}^{n} x_\ell e_\ell\right) = \sum_{\ell=1}^{n} x_\ell A e_\ell$$

$$|Ax| = \sum_{\ell=1}^{n} |x_\ell A e_\ell| = \sum_{\ell=1}^{n} |Ae_\ell|$$

Ae_ℓ is the ℓ^{th} column of A.

Also for $A, B \in M_{m\times n}(\mathbb{R})$

$||A+B|| \leq ||A|| + ||B||$

5.6 Definition: Let $A \in M_{m\times n}(\mathbf{R})$ and let $S = \{x \in \mathbf{R}^n \,|\, |x| = 1\}$. Since S is compact, by the Extreme Value Theorem, the continuous function $f : \mathbf{R}^n \to \mathbf{R}$ given by $f(x) = |Ax|$ attains its maximum value on S. We define the **norm** of the matrix A to be

$$||A|| = \max\{|Ax| \,\big|\, |x| = 1\}.$$

5.7 Lemma: *(Properties of the Matrix Norm) Let $A \in M_{m\times n}(\mathbf{R})$. Then*
(1) $|Ax| \le ||A||\,|x|$ for all $x \in \mathbf{R}^n$,
(2) $||A|| \le \sum_{k=1}^{m}\sum_{\ell=1}^{n} |A_{k,\ell}|$, and
(3) $||A||$ is equal to the square root of the largest eigenvalue of the matrix A^TA.

Proof: When $x = 0 \in \mathbf{R}^n$ we have $|Ax| = 0 = ||A||\,|x|$ and when $0 \ne x \in \mathbf{R}^n$ we have

$$|Ax| = \Big||x|\,A\tfrac{x}{|x|}\Big| = |x|\big|A\tfrac{x}{|x|}\big| \le |x|\,||A||.$$

This proves Part (1). To prove Part (2), let $x \in \mathbf{R}^n$ with $|x| = 1$. Then $|x_\ell| \le |x| \le 1$ for all ℓ so

$$|Ax| = \Big|\sum_{k=1}^{m}(Ax)_k e_k\Big| \le \sum_{k=1}^{m}\big|(Ax)_k\big| = \sum_{k=1}^{m}\Big|\sum_{\ell=1}^{n}A_{k,\ell}x_\ell\Big| \le \sum_{k=1}^{m}\sum_{\ell=1}^{n}|A_{k,\ell}|\,|x_\ell| \le \sum_{k=1}^{m}\sum_{\ell=1}^{n}|A_{k,\ell}|.$$

We omit the proof of Part (3), which we shall not use (it is often proven in a linear algebra course).

5.8 Theorem: *(Differentiability Implies Continuity) Let $f : U \subseteq \mathbf{R}^n \to \mathbf{R}^m$. If f is differentiable at $a \in U$, then f is continuous at a.* U is open in $\mathbb{R}^n$ with $a \in U$.

Proof: Suppose f is differentiable at a. Note that for all $x \in U$ we have

$$\begin{aligned}|f(x) - f(a)| &= \big|f(x) - f(a) - Df(a)(x-a) + Df(a)(x-a)\big| \\ &\le \big|f(x) - f(a) - Df(a)(x-a)\big| + |Df(a)\,(x-a)| \\ &\le \big|f(x) - f(a) - Df(a)(x-a)\big| + ||Df(a)||\,|x-a|\end{aligned}$$

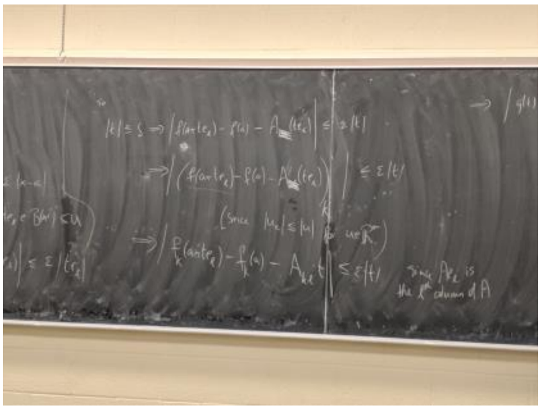
Let $\epsilon > 0$. Since f is differentiable at a we can choose δ with $0 < \delta < \frac{\epsilon}{1+||Df(a)||}$ such that

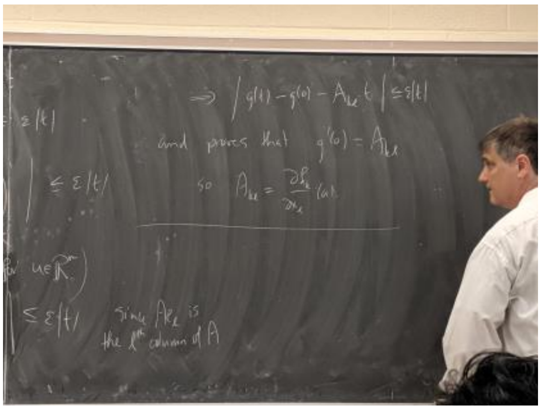
$$|x-a| \le \delta \Longrightarrow \big|f(x) - f(a) - Df(a)(x-a)\big| \le |x-a|$$

and then for $|x-a| \le \delta$ we have

$$\begin{aligned}\big|f(x) - f(a)\big| &\le \big|f(x) - f(a) - Df(a)(x-a)\big| + ||Df(a)||\,|x-a| \\ &\le |x-a| + ||Df(a)||\,|x-a| = \big(1 + ||Df(a)||\big)\,|x-a| \\ &\le \big(1 + ||Df(a)||\big)\,\delta < \epsilon.\end{aligned}$$

2





$\mathbb{R}$ case:

Let $\epsilon > 0$
Choose $\delta > 0$ with $\delta < \frac{\epsilon}{1+f'(a)}$
such that

$|x-a| \le \delta \Rightarrow |f(x) - f(a) - f'(a)(x-a)| \le 1|x-a|$
$\Rightarrow |f(x) - f(a)| \le |f(x) - f(a) - f'(a)(x-a)| + |f'(a)(x-a)|$
$\le |x-a| + |f'(a)||x-a| = (1 + f'(A))|x-a|$
$\le (1 + f'(a)))\delta$

For $\mathbb{R}^n$, simply change $f'(a)$ to $Df(a)$.
And add || || matrix norm.

Theorem: (Continuous Partial Derivatives Implies Continuity)
Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$
with U open in $\mathbb{R}^n$ and $a \in U$
Suppose that all the partial derivatives $\frac{\partial f_k}{\partial x_\ell}(x)$ exist at every $x \in U$ and are continuous at a.
Then, f is differentiable at a.

Remark: $\frac{\partial f_k}{\partial x_\ell}(a) = g'(a_\ell)$ where $g(t) = f(a_1, \dots, a_{\ell-1}, t, a_{\ell+1}, \dots, a_n)$
Hence

$$\frac{\partial f_k}{\partial x_\ell}(a_1, \dots, a_{\ell-1}, t, a_{\ell+1}, \dots a_n) = g'(t)$$

for all t

that is

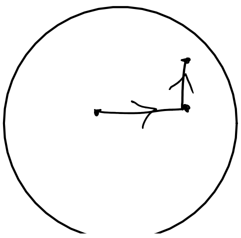
$$\frac{\partial f_k}{\partial x_\ell}(\alpha(t)) = g'(\alpha(t))$$

where $\alpha(t) = (a_1, \dots, a_{\ell-1}, t, a_{\ell+1}, \dots, a_n)$

Proof:

Let $\epsilon > 0$
Choose $\delta > 0$
so that $\bar{B}(a, \delta) \subseteq U$
and so that $\left|\frac{\partial f_k}{\partial x_\ell}(y) - \frac{\partial f_k}{\partial x_\ell}(a)\right| < \frac{\epsilon}{nm}$
for all $y \in \bar{B}(a, \delta)$ and indices k, ℓ.
We want to show that, $\forall x \in U, |x-a| \le \delta \Rightarrow |f(x) - f(a) - Df(a)(x-a)| \le \epsilon|x-a|$
Let $U_\ell = (x_1, \dots, x_\ell, a_{\ell+1}, \dots, a_n)$
for $0 \le \ell \le n$ with $u_o = a, u_n = x$
and note that each $u_\ell \in \bar{B}(a, \delta)$
Since $\bar{B}(a, \delta)$ is convex)

For $1 \le \ell \le n$, let
$\alpha_\ell(t) = (x_1, \dots, x_{\ell+1}, t, a_{\ell+1}, \dots, a_n)$
for t between a_ℓ and x_ℓ
so that $\alpha_\ell(a_\ell) = U_{\ell-1}$
and $\alpha_\ell(x_\ell) = U_\ell$
For $1 \le \ell \le n, 1 \le k \le m$
Let $g_{k,\ell}(t) = f_k(\alpha_\ell(t))$
so that



$(a_1, a_2, a_3) \to (x_1, a_2, a_3) \to (x_1, x_2, a_3)$

$$g'_{k,\ell}(t) = \frac{\partial f_k}{\partial x_\ell}(\alpha_\ell(t))$$

For $1 \le i \le n$, $1 \le k \le m$
Let $g_{k,\ell}(t) = f_k(\alpha_\ell(t))$
so that

$$g'_{k,\ell}(t) = \frac{\partial f_k}{\partial x_\ell}(\alpha_\ell(t))$$

for t between a_ℓ, x_ℓ.

By the Mean Value Theorem
(applied to $g_{k,\ell}(t)$ for t between a_ℓ and x_ℓ)
We can choose $S_{k,\ell}$ between a_ℓ and x_ℓ so that
$(x_\ell - a_\ell)g'_{(k,\ell)}(S_{k,\ell}) = g_{k,\ell}(x_\ell) - g_{k,\ell}(a_\ell)$
that is

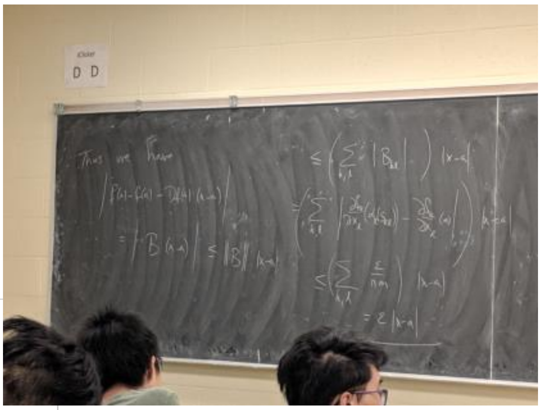
$$\frac{\partial f_k}{\partial x_\ell}\left(\alpha_\ell(S_{k,\ell})\right)(x_\ell - a_\ell) = f_k(U_\ell) - f_k(U_{\ell-1})$$

So we have

$$f_k(x) - f_k(a) = f_k(U_n) - f_k(U_o) = \sum_{i=1}^{n} f_k(U_\ell) - f_k(U_{\ell-1}) = \sum_{i=1}^{n} \frac{\partial f_k}{\partial x_\ell}\left(\alpha_\ell(S_{k,\ell})\right)\cdot(x_\ell - a_\ell)$$

$$f_k(x) - f_k(a) - \left(Df(a)(x-a)\right)_k = \sum_{\ell=1}^{n}\frac{\partial f_k}{\partial x_\ell}\left(\alpha_{\ell(S_{k,\ell})}\right)(x_\ell - a_\ell) - \sum_{\ell=1}^{n}\frac{\partial f_k}{\partial x_\ell}(a)(x_\ell - a_\ell)$$

$= \left(B(x-a)\right)_k$
where B is the matrix with entries $B_{k,\ell} = \frac{\partial f_k}{\partial x_\ell}\left(\alpha_\ell(S_{k,\ell})\right) - \frac{\partial f_k}{\partial x_\ell}(a)$
Continue right side



5.9 Theorem: *(The Chain Rule) Let $f : U \subseteq \mathbf{R}^n \to V \subseteq \mathbf{R}^m$, let $g : V \subseteq \mathbf{R}^m \to \mathbf{R}^\ell$, and let $h(x) = g(f(x))$. If f is differentiable at a and g is differentiable at $f(a)$ then h is differentiable at a and $Dh(a) = Dg(f(a))Df(a)$.*

Proof: Suppose f is differentiable at a and g is differentiable at $f(a)$. Write $y = f(x)$ and $b = f(a)$. We have

$$\begin{aligned}\left|h(x) - h(a) - Dg(f(a))Df(a)(x-a)\right| &= \left|g(y) - g(b) - Dg(b)Df(a)(x-a)\right| \\ &= \left|g(y) - g(b) - Dg(b)(y-b) + Dg(b)(y-b) - Dg(b)Df(a)(x-a)\right| \\ &\le \left|g(y) - g(b) - Dg(b)(y-b)\right| + ||Dg(b)||\,\left|y - b - Df(a)(x-a)\right| \\ &= \underbrace{\left|g(y) - g(b) - Dg(b)(y-b)\right|} + \underbrace{\left(1 + ||Dg(b)||\right)\left|f(x) - f(a) - Df(a)(x-a)\right|}\end{aligned}$$

Triangle Inequality

make this $\le \epsilon|y-b|$

make this $\le \epsilon|x-a|$

and

$$\begin{aligned}|y-b| &= |f(x) - f(a)| \\ &= \left|f(x) - f(a) - Df(a)(x-a) + Df(a)(x-a)\right| \\ &\le \left|f(x) - f(a) - Df(a)(x-a)\right| + ||Df(a)||\,|x-a|\,.\end{aligned}$$

Let $\epsilon > 0$ be given. Since g is differentiable at b we can choose $\delta_0 > 0$ so that

$$|y-b| \le \delta_0 \Longrightarrow \left|g(y) - g(b) - Dg(b)(y-b)\right| \le \frac{\epsilon}{2(1+||Df(a)||)}\,|y-b|\,.$$

Since f is continuous at a we can choose $\delta_1 > 0$ so that

$$|x-a| \le \delta_1 \Longrightarrow |y-b| = |f(x) - f(a)| \le \delta_0$$

$$\delta_2 \le \frac{\delta_1}{1 + \left|\left|Df(a)\right|\right|}$$

Since f is differentiable at a we can choose $\delta_2 > 0$ so that

$$|x-a| \le \delta_2 \Longrightarrow \left|f(x) - f(a) - Df(a)(x-a)\right| \le |x-a|$$

and we can choose $\delta_3 > 0$ so that

$$|x-a| \le \delta_3 \Longrightarrow \left|f(x) - f(a) - Df(a)(x-a)\right| \le \frac{\epsilon}{2(1+||Dg(a)||)}\,|x-a|\,.$$

Let $\delta = \min\{\delta_1, \delta_2, \delta_3\}$. Then for $|x-a| \le \delta$ we have

$$\begin{aligned}|y-b| &\le \left|f(x) - f(a) - Df(a)(x-a)\right| + \left|Df(a)(x-a)\right| \\ &\le |x-a| + ||Df(a)||\,|x-a| \\ &= \left(1 + ||Df(a)||\right)|x-a|\end{aligned}$$

so

$$\left|g(y) - g(b) - Dg(b)(y-b)\right| \le \frac{\epsilon}{2(1+||Df(a)||)}\,|y-b| \le \frac{\epsilon}{2}\,|x-a|$$

and we have

$$\left(1 + ||Dg(b)||\right)\left|f(x) - f(a) - Df(a)(x-a)\right| \le \frac{\epsilon}{2}\,|x-a|$$

and so

$$\left|h(x) - h(a) - Dg(f(a))Df(a)(x-a)\right| \le \frac{\epsilon}{2}\,|x-a| + \frac{\epsilon}{2}\,|x-a| = \epsilon|x-a|.$$

Thus h is differentiable at a with derivative $Dh(a) = Dg(f(a))Df(a)$, as required.

屌

很屌

3

Corollary: If $f: U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is e^1 in U, which means that all the parital derivatives $\frac{\partial f_k}{\partial x_\ell}(x)$ exists and are continuous in U.
Then f is differentiable in U.

Elementary functions are differentiable assuming open domain

$f(x,y) = |xy|$
f is continuous everywhere for all $(x,y) \in \mathbb{R}^2$
and f is not differentiable for all (x,y) for which $|xy| \neq 0$.
$\sqrt{u}$ is continuous. $u \ge 0$
$\sqrt{u}$ is differentiable when $u > 0$

If $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ is e^1 in the open set U.
Then f is differentiable.
The basic elementary functions are the single-variable functions, c, x, x^n where $\mathbb{Z}^+$

$x^{\frac{1}{n}}$ for $x \ge 0$
$e^x, \ln x$ for $x > 0$
$\sin x$ and $\sin^{-1} x$ for $-1 \le x \le 1$
with the k^{th} inclusion
$I_k: \mathbb{R} \to \mathbb{R}^n, I_k(t) = te_k = (0, \dots, 0, t, 0, \dots, 0)$

and the k^{th} projection $P_k: \mathbb{R}^n \to \mathbb{R}$
$P_k(x) = x_k$

The basic open-domain elementary functions are the same functions but restricted to open domains, so
$x^{\frac{1}{n}}$ defined for $x > 0$
$\sin^{-1} x$ defined for $-1 < x < 1$

$f: U \subseteq \mathbb{R}^n \to \mathbb{R}^m$
$g: V \subseteq \mathbb{R}^m \to \mathbb{R}^\ell$

$h = g \circ f: U \cap f^{-1}(V) \subseteq \mathbb{R}^n \to \mathbb{R}$

If $f: A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is elementary, then f is continuous (in A)
If $f: U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is an open-domain elementary function (obtained by applying operations of functions to the basic open-domain elementary functions)
Then f is differentiable, C^1 (actually C^∞) in its domain, hence differentiable in U.

eg. For $f(x,y) = |xy|$, f is continuous at every $(x,y) \in \mathbb{R}^2$ because $f(x,y) = \sqrt{x^2y^2}$.
Also, f is C^1 hence differentiable in $U = \{(x,y) \in \mathbb{R}^n | xy \neq 0\}$
because f is open-domain elementary in U, indeed.
If $p(x,y) = x, g(x,y) = y$
$s(u) = u^2$
$r(u) = \sqrt{u}$ for $u > 0$
then $f = r \circ \left(s \circ (p \cdot q)\right)$

We also saw that $f(x,y) = |xy|$ is differentiable at $(0,0)$

Note: That when $a \neq 0$, f is not differentiable at $(a, 0)$.
Because for $g(t) = f(a,t)$ we have $g(t) = |at| = |a|\cdot|t|$ which is not differentiable at $t = 0$ so $\frac{\partial f}{\partial y}(a,0)$ does not exist

Similarly, $\frac{\partial f}{\partial x}(0,a)$ does not exist so f is not differentiable at $(0,a)$ when $a \neq 0$.

We plan to show that when $f: U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is C^2 in U, which means that all second order partial derivatives

$$\frac{\partial^2 f_j}{\partial x_k \partial x_\ell}(x) = \frac{\partial}{\partial x_k}\left(\frac{\partial f_j}{\partial x_\ell}\right)(x)$$

exists and are continuous at every $x \in U$, then

5.10 Definition: Let $f : U \subseteq \mathbf{R}^n \to R$, let $a \in \mathbf{R}^n$ and let $v \in \mathbf{R}^n$. We define the **directional derivative** of f at a with respect to v, written as $D_v f(a)$, as follows: pick any differentiable function $\alpha : (-\epsilon, \epsilon) \subseteq \mathbf{R} \to U \subseteq \mathbf{R}^n$, where $\epsilon > 0$, such that $\alpha(0) = a$ and $\alpha'(0) = v$ (for example, we could pick $\alpha(t) = a + v\,t$), let $g(t) = f(\alpha(t))$, note that by the Chain Rule we have $g'(t) = Df(\alpha(t))\alpha'(t)$, and then define

$$\begin{aligned}D_v f(a) &= g'(0) \\ &= Df(\alpha(0))\,\alpha'(0) \\ &= Df(a)\,v \\ &= \nabla f(a) \cdot v\,.\end{aligned}$$

Notice that the formula for $D_v f(a)$ does not depend on the choice of the function $\alpha(t)$.

5.11 Remark: Some books only define the directional derivative in the case that vector is a unit vector.

Notice that the formula for $D_v f(a)$ does not depend on the choice of the function $\alpha(t)$.

5.11 Remark: Some books only define the directional derivative in the case that vector is a unit vector.

5.12 Theorem: *Let $f : U \subseteq \mathbf{R}^n \to \mathbf{R}$ be differentiable at $a \in U$. Say $f(a) = b$. The gradient $\nabla f(a)$ is perpendicular to the level set $f(x) = b$, it is in the direction in which f increases most rapidly, and its length is the rate of increase of f in that direction.*

Proof: Let $\alpha(t)$ be any curve in the level set $f(x) = b$, with $\alpha(0) = a$. We wish to show that $\nabla f(a) \perp \alpha'(0)$. Since $\alpha(t)$ lies in the level set $f(x) = b$, we have $f(\alpha(t)) = b$ for all t. Take the derivative of both sides to get $Df(\alpha(t))\alpha'(t) = 0$. Put in $t = 0$ to get $Df(a)\,\alpha'(0) = 0$, that is $\nabla f(a) \cdot \alpha'(0) = 0$. Thus $\nabla f(a)$ is perpendicular to the level set $f(x) = b$.

Next, let u be a unit vector. Then $D_u f(a) = \nabla f(a) \cdot u = |\nabla f(a)| \cos\theta$, where θ is the angle between u and $\nabla f(a)$. So the maximum possible value of $D_u f(a)$ is $|\nabla f(a)|$, and this occurs when $\cos\theta = 1$, that is when $\theta = 0$, which happens when u is in the direction of $\nabla f(a)$.

5.13 Theorem: *Let $U \subseteq \mathbf{R}^n$ be open, let $f : U \subseteq \mathbf{R}^n \to \mathbf{R}^m$ and let $a \in U$. If the partial derivatives $\frac{\partial f_k}{\partial x_\ell}(x)$ are all continuous at a then f is differentiable at a. If f is $\mathcal{C}^1$ in U then f is differentiable in U.*

Proof: Suppose that f is $\mathcal{C}^1$ in U and let $a \in U$. Since each function $\frac{\partial f_k}{\partial f_l}$ is continuous at a, we can choose $r > 0$ so that $B(a,r) \subseteq U$ and $\left|\frac{\partial f_k}{\partial f_\ell}(y) - \frac{\partial f_k}{\partial f_\ell}(a)\right| < \frac{\epsilon}{nm}$ for all $y \in B(a,r)$. Let $x \in B(a,r)$. Let $g(t) = f\big(a + t(x-a)\big)$. By the Chain Rule, $g'(t) = Df\big(a + t(x-a)\big)(x-a)$. By the Mean Value Theorem (for a real-valued function of a single variable) we can choose $s \in [0,1]$ such that $g'(s) = g(1) - g(0) = f(x) - f(a)$. Then, letting $y = a + s(x-a)$ so that $g'(s) = Df(y)(x-a)$, and using Properties of the Matrix Norm, we have

$$
\begin{aligned}
\big|f(x) - f(a) - Df(a)(x-a)\big| &\leq \big|f(x) - f(a) - g'(s)\big| + \big|g'(s) - Df(a)(x-a)\big| \\
&= 0 + \big|g'(s) - Df(a)(x-a)\big| = \big|\big(Df(y) - Df(a)\big)(x-a) \\
&\leq \big\|Df(y) - Df(a)\big\|\,|x-a| \\
&\leq \sum_{k=1}^{m}\sum_{\ell=1}^{n} \left|\tfrac{\partial f_k}{\partial x_\ell}(y) - \tfrac{\partial f_k}{\partial x_\ell}(a)\right| |x-a| \\
&\leq \sum_{k=1}^{m}\sum_{\ell=1}^{n} \tfrac{\epsilon}{nm}\,|x-a| = \epsilon\,|x-a|.
\end{aligned}
$$

4

$$\frac{\partial^2 f_j}{\partial x_k \partial x_\ell}(x) = \frac{\partial}{\partial x_k}\left(\frac{\partial f_j}{\partial x_\ell}\right)(x)$$

exists and are continuous at every $x \in U$, then

$$\frac{\partial^2 f_j}{\partial x_k \partial x_\ell}(x) = \frac{\partial}{\partial x_k}\left(\frac{\partial f_j}{\partial x_\ell}\right)(x)$$

for all $x \in U$

We first prove a lemma

Lemma (Iterated Limits)

Let $I \subseteq \mathbb{R}$ and $J \subseteq \mathbb{R}$ be open intervals in $\mathbb{R}$ with $a \in I$ and $b \in J$

Let $U = (I \times J) \setminus \{(a,b)\}$

and let $f : U \subseteq \mathbb{R}^2 \to \mathbb{R}$

Suppose that $\lim_{(x,y)\to(a,b)} f(x,y) = u \in \mathbb{R}$

and suppose that for every $x \in I \setminus \{a\}$

$\lim\limits_{y\to b} f(x,y)$ exists

Then $\lim_{x\to a} \lim_{y\to b} f(x,y) = u$

eg For $f(x,y) = \frac{x^2}{x^2+y^2}$

$$\lim_{x\to 0} f(x,y) = \begin{cases} 0 & \text{if } y \neq 0 \\ 1 & \text{if } y = 0 \end{cases}$$

and

$$\lim_{y\to 0} f(x,y) = \begin{cases} 1 & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

So $\lim_{x\to 0} \lim_{y\to 0} f(x,y) = 1$

But $\lim_{y\to 0} \lim_{x\to 0} f(x,y) = 0$

5.14 Corollary: *Every function $f : U \subseteq \mathbf{R}^n \to \mathbf{R}^m$, which can be obtained by applying the standard operations (such as multiplication and composition) on functions to basic elementary functions defined on open domains, is differentiable in U.*

5.15 Exercise: For each of the following functions $f : \mathbf{R}^2 \setminus \{(0,0)\} \to \mathbf{R}$, extend the domain of $f(x,y)$ to all of $\mathbf{R}^2$ by defining $f(0,0) = 0$ and then determine whether the partial derivatives of f exist at $90,0)$ and whether f is differential at $(0,0)$.

(a) $f(x,y) = \frac{xy}{x^2+y^2}$ (b) $f(x,y) = |xy|$ (c) $f(x,y) = \sqrt{|xy|}$
(d) $f(x,y) = \frac{x^3}{x^2+y^2}$ (e) $f(x,y) = \frac{x}{(x^2+y^2)^{1/3}}$ (f) $f(x,y) = \frac{x^3-xy^2}{x^2+y^2}$

5.16 Definition: For $a, b \in \mathbf{R}^n$, we define the **line segment** from a to b to be the set

$$[a,b] = \big\{a + t(b-a)\big|0 \le t \le 1\big\}.$$

For $A \subseteq \mathbf{R}^n$ we say the A is **convex** when for all $a, b \in A$ we have $[a,b] \subseteq A$.

5.17 Exercise: Show, using the triangle inequality, that $B(a,r)$ is convex for al $a \in \mathbf{R}^n$ and $r > 0$.

5.18 Theorem: *(The Mean Value Theorem) Let $f : U \subseteq \mathbf{R}^n \to \mathbf{R}^m$ with U open in $\mathbf{R}^n$. Suppose that f is differentiable in U. Let $u \in \mathbf{R}^m$ and let $a, b \in U$ with $[a,b] \subseteq U$. Then there exists $c \in [a,b]$ such that*

$$Df(c)(b-a) \cdot u = \big(f(b) - f(a)\big) \cdot u.$$

Proof: Let $\alpha(t) = a + t(b-a)$ and define $g : [0,1] \to \mathbf{R}$ by $g(t) = f\big(\alpha(t)\big) \cdot u$. By the Chain Rule, we have $g'(t) = \big(Df(\alpha(t))\alpha'(t)\big) \cdot u = \big(Df(\alpha(t))(b-a)\big) \cdot u$. By the Mean Value Theorem (for a real-valued function of a single variable) we can choose $s \in [0,1]$ such that $g'(s) = g(1) - g(0)$, that is $\big(Df(\alpha(s))(b-a)\big) \cdot u = f(b) \cdot u - f(a) \cdot u = \big(f(b) - f(a)\big) \cdot u$. Thus we can take $c = \alpha(s) \in [a,b]$ to get $Df(c)(b-a) \cdot u = \big(f(b) - f(a)\big) \cdot u$.

5.19 Corollary: *(Vanishing Derivative) Let $U \subseteq \mathbf{R}^n$ be open and connected and let $f : U \to \mathbf{R}^n$ be differentiable with $Df(x) = O$ for all $x \in U$. Then f is constant in U.*

Proof: Let $a \in U$ and let $A = \big\{x \in U\big|f(x) = f(a)\big\}$. We claim that A is open (both in $\mathbf{R}^n$ and in U). Let $b \in A$, that is let $b \in U$ with $f(b) = f(a)$. Since U is open we can choose $r > 0$ so that $B(b,r) \subseteq U$. Let $c \in B(b,r)$. Since $B(b,r)$ is convex we have $[b,c] \subseteq B(b,r) \subseteq U$. Let $u = f(c) - f(b)$ and choose $d \in [b,c]$, as in the Mean Value Theorem, so that $\big(Df(d)(c-b)\big) \cdot u = \big(f(c) - f(b)\big) \cdot u$. Then we have

$$\big|f(c) - f(b)\big|^2 = \big(f(c) - f(b)\big) \cdot u = \big(Df(d)(c-b)\big) \cdot u = 0$$

since $Df(d) = O$. Since $\big|f(c) - f(b)\big| = 0$ we have $f(c) = f(b) = f(a)$, and so $c \in A$. Thus $B(b,r) \subseteq A$ and so A is open, as claimed. A similar argument shows that if $b \in U \setminus A$ and we chose $r > 0$ so that $B(b,r) \subseteq U$ then we have $f(c) = f(b)$ for all $c \in B(b,r)$ hence $B(b,r) \subseteq U \setminus A$ and hence $U \setminus A$ is also open. Note that A is non-empty since $a \in A$. If $U \setminus A$ was also non-empty then U would be the union of the two non-empty open sets A and $U \setminus A$, and this is not possible since U is connected. Thus $U \setminus A = \emptyset$ so $U = A$. Since $U = A = \big\{x \in U\big|f(x) = f(a)\big\}$ we have $f(x) = f(a)$ for all $x \in U$, so f is constant in U.

chap6

2019年7月2日 12:35

chap6

Chapter 6. Higher Order Derivatives

6.1 Lemma: *(Iterated Limits) Let I and J be open intervals in $\mathbf{R}$ with $a \in I$ and $b \in J$, let $U = (I \times J) \setminus \{(a,b)\}$, and let $f : U \to \mathbf{R}$. Suppose that $\lim_{y \to b} f(x,y)$ exists for every $x \in I$ and that $\lim_{(x,y)\to(a,b)} f(x,y) = u \in \mathbf{R}$. Then $\lim_{x\to a} \lim_{t \to b} f(x,y) = u$.*

Proof: Define $g : I \to \mathbf{R}$ by $g(x) = \lim_{y\to b} f(x,y)$. Let $\epsilon > 0$. Since $\lim_{(x,y)\to(a,b)} f(x,y) = u$ we can choose $\delta > 0$ such that for all $(x,y) \in U$ with $0 < \left|(x,y) - (a,b)\right| \le 2\delta$ we have $\left|f(x,y) - u\right| \le \epsilon$. Let $x \in I$ with $0 < |x - a| \le \delta$. For all $y \in J$ with $0 < |y - b| \le \delta$ we have $0 < \left|(x,y) - (a,b)\right| \le |x - a| + |y - b| \le 2\delta$ and so $\left|f(x,y) - u\right| \le \epsilon$ and hence

$$|g(x) - u| \le \left|g(x) - f(x,y)\right| + \left|f(x,y) - u\right| \le \left|g(x) - f(x,y)\right| + \epsilon.$$

Take the limit as $y \to b$ on both sides to get $|g(x) - u| \le \epsilon$. Thus $\lim_{x\to a} g(x) = u$, as required.

6.2 Theorem: *(Mixed Partials Commute) Let $f : U \subseteq \mathbf{R}^n \to \mathbf{R}$ where U is open in $\mathbf{R}^n$ with $a \in U$, and let $k, \ell \in \{1, \cdots, n\}$. Suppose $\frac{\partial^2 f}{\partial x_k \partial x_\ell}(x)$ exists in U and is continuous at a, $\frac{\partial f}{\partial x_k}(x)$ exists and is continuous in U, and $\frac{\partial^2 f}{\partial x_\ell \partial x_k}(a)$ exists. Then $\frac{\partial^2 f}{\partial x_\ell \partial x_k}(a) = \frac{\partial^2 f}{\partial x_k \partial x_\ell}(a)$.*

Proof: When $k = \ell$ there is nothing to prove, so suppose that $k \ne \ell$. Choose $r > 0$ so that $B(a, 2r) \subseteq U$. For $|x| < r$ and $|y| < r$ note that the points a, $a + xe_k$, $a + ye_\ell$ and $a + xe_k + ye_\ell$ all lie in $B(a, 2r)$. For $|X| < r$ and $|y| < r$, define

$$g(x,y) = f(a + xe_k + ye_\ell) - f(a + xe_k) - f(a + ye_\ell) + f(a).$$

By the Mean Value Theorem, applied to the function $f(a + xe_k + ye_\ell) - f(a + ye_\ell)$ as a function of y, we can choose t between 0 and y such that

$$y\big(\tfrac{\partial f}{\partial x_\ell}(a + xe_k + te_\ell) - \tfrac{\partial f}{\partial x_\ell}(a + te_\ell)\big) = g(x,y).$$

By the Mean Value Theorem, applied to the function $\frac{\partial f}{\partial x_\ell}(a + xe_k + te_\ell)$ as a function of x, we can choose s between 0 and x such that

$$x\,\tfrac{\partial^2 f}{\partial x_k \partial x_\ell}(a + se_k + te_\ell) = \tfrac{\partial f}{\partial x_\ell}(a + xe_k + te_\ell) - \tfrac{\partial f}{\partial x_\ell}(a + te_\ell).$$

Also by the Mean Value Theorem, applied to the function $f(a + xe_k + ye_\ell) - f(a + xe_k)$ as a function of x, we can choose r between 0 and x such that

$$x\big(\tfrac{\partial f}{\partial x_k}(a + re_k + ye_\ell) - \tfrac{\partial f}{\partial x_k}(a + re_\ell)\big) = g(x,y).$$

Then for $|x| < r$ and $0 < |y| < r$ we have

$$\frac{\frac{\partial f}{\partial x_k}(a + re_k + ye_\ell) - \frac{\partial f}{\partial x_k}(a + re_k)}{y} = \frac{\partial^2 f}{\partial x_k \partial x_\ell}(a + se_k + te_\ell).$$

Since $\frac{\partial^2 f}{\partial x_k \partial x_\ell}$ is continuous, the limit on the right as $(x,y) \to (0,0)$ is equal to $\frac{\partial^2 f}{\partial x_k \partial x_\ell}(a)$, and since $\frac{\partial f}{\partial x_k}$ is continuous, the limit as $y \to 0$ of the limit as $x \to 0$ on the left is equal to $\frac{\partial^2 f}{\partial x_\ell \partial x_k}(a)$, so the desired result follows from the above lemma.

6.3 Corollary: *If $U \subseteq \mathbf{R}^n$ is open and $f : U \subseteq \mathbf{R}^n \to \mathbf{R}$ is $\mathcal{C}^2$ in U then we have $\frac{\partial^2 f}{\partial x_\ell \partial x_k}(x) = \frac{\partial^2 f}{\partial x_k \partial x_\ell}(x)$ for all $x \in U$ and for all k, ℓ.*

1

Maybe $\bar{B}(a, r) \subseteq U$
For $u, v \in \mathbb{R}$ with $|u| < r, |v| < r$ we have $a + se_k + te_\ell \in \bar{B}(a, 2r) \subseteq U$ for all U.

Let $g(u,v) = f(a + ue_k + ve_\ell) - f(a + ue_k) - f(a + ve_\ell) + f(a)$

By the MVT, applied to the function $f(a + ue_k + ve_\ell) - f(a + ve_\ell)$ using the variable v (with u fixed), we can choose t between 0 and v such that

$$v\left(\frac{\partial f}{\partial x_\ell}(a + ue_k + te_\ell) - \frac{\partial f}{\partial x_\ell}(a + te_\ell)\right)$$
$$= \big(f(a + ue_k + ve_\ell) - f(a + ve_\ell)\big) - \big(f(a + ue_k) - f(a)\big)$$
$$= g(u,v)$$

By the MVT applied to the function $\frac{\partial f}{\partial x_\ell}(a + ue_k + te_\ell)$ as a function of u (with t fixed), we can choose s between 0 and u such that

$$u\frac{\partial^2 f}{\partial x_k \partial x_\ell}(a + se_k + te_\ell) = \frac{\partial f}{\partial x_\ell}(a + ue_k + te_\ell) - \frac{\partial f}{\partial x_\ell}(a + te_\ell)$$

So we have $uv\left(\frac{\partial^2 f}{\partial x_k \partial x_\ell}(a + se_k + te_\ell)\right) = g(u,v)$

Apply the MVT again, to the function

$$f(a + ue_k + ve_\ell) - f(a + ue_k)$$

Using the variable u, we can choose r between 0 and u such that

$$u\left(\frac{\partial f}{\partial x_k}(a + re_k + ve_\ell) - f(a + re_k)\right) = g(u,v)$$

$$= uv\frac{\partial^2 f}{\partial x_k \partial x_\ell}(a + se_k + te_\ell)$$

For $0 < |u| < r, 0 < |v| < r$

$$\frac{\frac{\partial f}{\partial x_k}(a + re_k + ve_\ell) - \frac{\partial f}{\partial x_k}(a + re_k)}{v} = \frac{\partial^2 f}{\partial x_k \partial x_\ell}(a + se_k + te_\ell)$$

Since $\frac{\partial^2 f}{\partial x_k \partial x_\ell}(x)$ is continuous at a and since $|s| \le |u|, |t| \le |v|$. We have

$$\lim_{(u,v)\to(0,0)} \frac{\partial^2 f}{\partial x_k \partial x_\ell}(a + se_k + te_\ell) = \frac{\partial^2 f}{\partial x_k \partial x_\ell}(a)$$

Since $\frac{\partial f}{\partial x_k}(x)$ is continuous everywhere (hence at a and at $a + ve_\ell$)

We have $\lim_{u\to 0} \frac{\frac{\partial f}{\partial x_k}(a+re_k+ve_\ell) - \frac{\partial f}{\partial x_k}(a+re_k)}{v}$

By the above lemma, equal to $\frac{\partial^2 f}{\partial x_k \partial x_\ell}(a)$

6.4 Exercise: Verify that for $f(x,y) = \frac{x^2}{x^2+y^2}$ we have $\lim_{x\to 0}\lim_{y\to 0} f(x,y) \neq \lim_{y\to 0}\lim_{x\to 0} f(x,y)$.

6.5 Exercise: Let $f(x,y) = \begin{cases} \dfrac{xy(x^2-y^2)}{x^2+y^2} & \text{, if } (x,y)\neq(0,0) \\ 0 & \text{, if } (x,y)=(0,0) \end{cases}$. Verify that the mixed partial derivatives $\frac{\partial^2 f}{\partial x \partial y}(0,0)$ and $\frac{\partial^2 f}{\partial y \partial x}(0,0)$ both exist, but they are not equal.

6.6 Definition: for $f : U \subseteq \mathbf{R}^n \to \mathbf{R}$, where U is open in $\mathbf{R}^n$ with $a \in U$, we define $D^0 f(a) = f(a)$ and for $\ell \in \mathbf{Z}^+$ we define the ℓ^{th} **total differential** of f at a to be the map $D^\ell f(a) : \mathbf{R}^n \to \mathbf{R}$ given by

$$D^\ell f(a)(u) = \sum_{k_1=1}^{n}\sum_{k_2=1}^{n}\cdots\sum_{k_\ell=1}^{n} \frac{\partial^\ell f}{\partial x_{k_1}\partial x_{k_2}\cdots\partial x_{k_\ell}}(a)\, u_{k_1}u_{k_2}\cdots u_{k_\ell}$$

provided that all of the ℓ^{th} order partial derivatives exist at a.

6.7 Example: When $f : U \subseteq \mathbf{R}^2 \to \mathbf{R}$ is $\mathcal{C}^2$ (so the mixed partial derivatives commute) we have

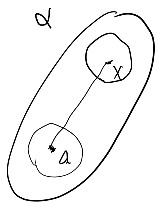
$$D^0 f(u,v) = f(a,b)$$
$$D^1 f(a,b)(u,v) = \tfrac{\partial f}{\partial x}(a,b)\,u + \tfrac{\partial f}{\partial y}(a,b)\,v$$
$$D^2 f(a,b)(u,v) = \tfrac{\partial f}{\partial x^2}(a,b)\,u^2 + 2\tfrac{\partial f}{\partial x\partial y}(a,b)\,uv + \tfrac{\partial f}{\partial y^2}(a,b)\,v^2.$$

6.8 Theorem: *(Taylor's Theorem) Let $f : U \subseteq \mathbf{R}^n \to \mathbf{R}$ where U is open in $\mathbf{R}^n$. Suppose that the m^{th} oder partial derivatives of f all exist in U. Then for all $a, x \in U$ such that $[a,x] \subseteq U$ there exists $c \in [a,x]$ such that*

$$f(x) = \sum_{\ell=0}^{m-1} \tfrac{1}{\ell!} D^\ell f(a)(x-a) \; + \; \tfrac{1}{m!} D^m f(c)(x-a).$$

f is $\mathcal{C}^m$

What this looks like at two variables.



Proof: Let $a, x \in U$ with $[a,x] \subseteq U$. Let $\alpha(t) = a + t(x-a)$ for all $t \in \mathbf{R}$ and note that $\alpha(t) \in U$ for $0 \le t \le 1$. Since U is open and α is continuous, we can choose $\delta > 0$ so that $\alpha(t) \in U$ for all $t \in I = (-\delta, 1+\delta)$. Define $g : I \to \mathbf{R}$ by $g(t) = f(\alpha(t))$. By the Chain Rule, we have

$$g'(t) = Df(\alpha(t))\alpha'(t) = Df(\alpha(t))(x-a) = \sum_{i=1}^{n} \tfrac{\partial f}{\partial x_i}(\alpha(t))(x_i - a_i) = D^1 f(\alpha(t))(x-a).$$

By the Chain Rule again, we have

$$g''(t) = \sum_{i=1}^{n}\Big(\sum_{j=1}^{n} \tfrac{\partial^2 f}{\partial x_j \partial x_i}(\alpha(t))(x_j - a_j)\Big)(x_i - a_i) = D^2 f(\alpha(t))(x-a).$$

An induction argument shows that

$$g^{(\ell)}(t) = D^\ell f(\alpha(t))(x-a).$$

By Taylor's Theorem, applied to the function $g(t)$ on the interval $[0,1]$, we can choose $s \in [0,1]$ such that $g(1) = \sum_{\ell=0}^{m-1} \frac{1}{\ell!} g^{(\ell)}(0) \; + \; \frac{1}{m!} g^{(m)}(s)$, that is

$$f(x) = \sum_{\ell=0}^{m-1} \tfrac{1}{\ell!} D^\ell f(a)(x-a) \; + \; \tfrac{1}{m!} D^m f(\alpha(s))(x-a).$$

c

Thus we can choose $c = \alpha(s) \in [a,x]$.

2

6.9 Definition: For $f : U \subseteq \mathbf{R}^n \to \mathbf{R}$, where U is open in $\mathbf{R}^n$ with $a \in U$, we define the m^{th} **Taylor polynomial** of f at a to be the polynomial

$$T^m f(a)(x) = \sum_{\ell=0}^{m} \tfrac{1}{\ell!} D^\ell f(a)(x-a)$$

provided that all the m^{th} order partial derivatives exist at a. When f is $\mathcal{C}^2$ in U (so that the mixed partial derivatives commute) we have

$$T^2 f(a)(x) = f(a) + Df(a)(x-a) + (x-a)^T Hf(a)\,(x-a)$$

where $Hf(a) \in M_{n\times n}(\mathbf{R})$ is the symmetric matrix with entries $Hf(a)_{k,\ell} = \frac{\partial^2 f}{\partial x_k \partial x_\ell}(a)$. The matrix $Hf(a)$ is called the **Hessian matrix** of f at a.

$D^0 f(a,b)(u,v) = f(a)$

$D^1 f(a,b)(u,v) = \frac{\partial f}{\partial x}(a,b)\, u + \frac{\partial f}{\partial x}(a,b)\, v$

$D^2 f(a,b)(u,v) = \frac{\partial^2 f}{\partial x^2}(a,b)\,u^2 + \frac{\partial^2 f}{\partial x \partial y}(a,b)\,u\cdot v + \frac{\partial^2 f}{\partial x \partial y}(a,b)\, v\cdot u + \frac{\partial^2 f}{\partial y^2}(a,b)\,v^2$

$= \cdots\cdots$ if is C^2

$2\frac{\partial f}{\partial x \partial y}(a,b)\,uv$

$T^0 f(a)(x) = f(a)$
$T^1 f(a)(x) = f(a) + Df(a)(x-a)$
(The linearization of f at a)

and

$T^2 f(a)(x) = f(a) + Df(a)(x-a) + \frac{1}{2!}(x-a)^T Hf(a)(x-a)$

where

$$Hf(a) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(a) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1 (a)} & \cdots & \cdots & \cdots \end{pmatrix}$$

That is $Hf(a) \in M_n(\mathbb{R})$
with $Hf(a)_{k,\ell} = \frac{\partial^2 f}{\partial x_k \partial x_\ell}(a)$

This matrix is called the Hessian matrix of f at a.

Ex1. Define local maximum and minimum values for $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$

Ex2. Show that if $Df(a) \neq (0, \dots, 0)$
Then f does not have max or mim at a.

6.10 Definition: Let $A \in M_n(\mathbf{R})$ be a symmetric matrix. We say that

(1) A is **positive-definite** when $u^T A u > 0$ for all $0 \neq u \in \mathbf{R}^n$,
(2) A is **negative-definite** when $u^T A u < 0$ for all $0 \neq u \in \mathbf{R}^n$, and
(3) A is **indefinite** when there exist $0 \neq u, v \in \mathbf{R}^n$ with $u^T A u > 0$ and $v^T A v < 0$.

6.11 Theorem: *(Characterization of Positive-Definiteness by Eigenvalues) Let $A \in M_n(\mathbf{R})$ be symmetric. Then*

(1) A is positive-definite if and only if all of the eigenvalues of A are positive,
(2) A is negative-definite if and only if all of the eigenvalues of A are negative, and
(3) A is indefinite if and only if A has a positive eigenvalue and a negative eigenvalue.

Proof: Suppose that A is positive definite. Let λ be an eigenvalue of A and let u be a unit eigenvector for λ. Then $\lambda = \lambda|u|^2 = \lambda(u \cdot u) = \lambda u \cdot u = Au \cdot u = u^T A u > 0$. Conversely, suppose that all of the eigenvalues of A are positive. Since A is symmetric, we can orthogonally diagonalize A. Choose a matrix $P \in M_n(\mathbf{R})$ with $P^T = P$ so that $P^T A P = D = \text{diag}(\lambda_1, \cdots, \lambda_n)$. Given $0 \neq u \in \mathbf{R}^n$, let $v = P^T u$. Note that $v \neq 0$ since P^T is invertible. Thus $u^T A u = u^T P D P^T u = v^T D v = \sum_{i=1}^{n} \lambda_i {v_i}^2 > 0$ since every $\lambda_i > 0$ and some $v_i \neq 0$. This proves Part (1). The proofs of Parts (2) and (3) are fairly similar.

6.12 Theorem: *(Characterization of Positive-Definiteness by Determinant) Let $A \in M_n(\mathbf{R})$ be symmetric. For each k with $1 \le k \le n$, let $A^{(k)}$ denote the upper-left $k \times k$ sub matrix of A. Then*

(1) A is positive-definite if and only if $\det(A^{(k)}) > 0$ for all k with $1 \le k \le n$, and
(2) A is negative-definite if and only if $(-1)^k \det(A^{(k)}) > 0$ for all k with $1 \le k \le n$.

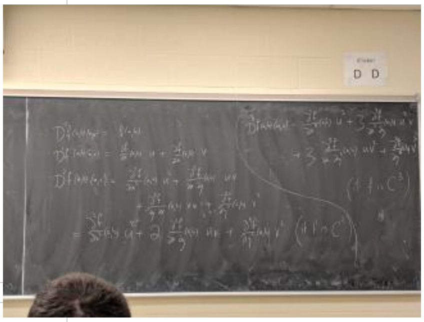
Proof: Part (2) follows easily from Part (1) by noting that A is negative-definite if and only if $-A$ is positive-definite. We shall prove one direction of Part (1). Suppose that A is positive-definite. Let $1 \le k \le n$. Since $u^T A u > 0$ for all $0 \neq u \in \mathbf{R}^n$, we have $(u^T \;\; 0)\, A \begin{pmatrix} u \\ 0 \end{pmatrix} = 0$, or equivalently $u^T A^{(k)} u > 0$, for all $0 \neq u \in \mathbf{R}^k$. This shows that $A^{(k)}$ is positive definite. By the previous theorem, all of the eigenvalues of $A^{(k)}$ are positive. Since $\det(A^{(k)})$ is equal to the product of its eigenvalues, we see that $\det(A^{(k)}) > 0$.

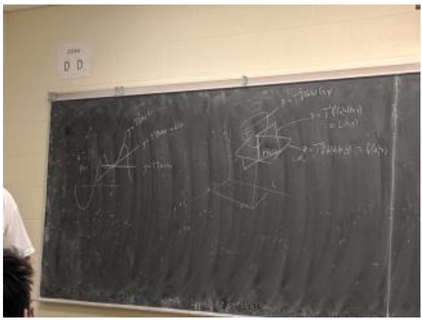
The proof of the other direction of Part (1) is more difficult. We shall not use it and we shall omit the proof. It is often proven in a linear algebra course.

Proof: $P^T A P = \begin{pmatrix} \lambda_1 & \dots & \dots \\ \vdots & \vdots & \vdots \\ \cdots & \cdots & \lambda_n \end{pmatrix}$ Let

3

$T^3 f(a,b)(x,y) = f(a,b) + \frac{1}{1!}(D^2 f(a,b)$





Global maximum at a when $f(x) \le f(a)$ for all $x \in A$.

6.13 Exercise: Let $A = \begin{pmatrix} 3 & -1 & 2 \\ -1 & 2 & 1 \\ 2 & 1 & 5 \end{pmatrix}$. Determine whether A is positive-definite, negative-definite, or indefinite.

6.14 Definition: Let $f : A \subseteq \mathbf{R}^n \to \mathbf{R}$ and let $a \in A$. We say that f has a **local maximum value** at a when there exists $r > 0$ such that $f(a) \ge f(x)$ for all $x \in B_A(a,r)$.

6.13 Exercise: Let $A = \begin{pmatrix} -1 & 2 & 1 \\ 2 & 1 & 5 \end{pmatrix}$. Determine whether A is positive-definite, negative-definite, or indefinite.

6.14 Definition: Let $f : A \subseteq \mathbf{R}^n \to \mathbf{R}$ and let $a \in A$. We say that f has a **local maximum value** at a when there exists $r > 0$ such that $f(a) \geq f(x)$ for all $x \in B_A(a,r)$. We say that f has a **local minimum value** at a when there exists $r > 0$ such that $f(a) \leq x$ for all $x \in B_A(a,r)$.

6.15 Exercise: Show that when $f : U \subseteq \mathbf{R}^n \to \mathbf{R}$ where U is open in $\mathbf{R}^n$ with $a \in U$, if f has a local maximum or minimum value at a then either $Df(a) = 0$ or $Df(a)$ does not exist (that is one of the partial derivatives $\frac{\partial f}{\partial x_k}(a)$ does not exist).

6.16 Definition: Let $f : U \subseteq \mathbf{R}^n \to \mathbf{R}$ where U is open in $\mathbf{R}^n$. For $a \in U$, we say that a is a **critical point** of f when either $Df(a) = 0$ or $Df(a)$ does not exist. When $a \in U$ is a critical point of f but f does not have a local maximum or minimum value at a, we say that a is a **saddle point** of f.

6.17 Theorem: *(The Second Derivative Test) Let $f : U \subseteq \mathbf{R}^n \to \mathbf{R}$ with U open in $\mathbf{R}^n$ and let $a \in U$. Suppose that f is $\mathcal{C}^2$ in U with $Df(a) = 0$. Then*

(1) if $Hf(a)$ is positive definite then f has a local minimum value at a, follows by Taylor's Theorem

(2) if $Hf(a)$ is negative definite then f has a local maximum value at a, and

(3) if $Hf(a)$ is indefinite then f has a saddle point at a.

Proof: Suppose that $Hf(a)$ is positive-definite. Then $\det\left(Hf(a)^{(k)}\right) > 0$ for $1 \leq k \leq n$. Since each determinant function $\det(A^{(k)})$ is continuous as a function in the entries of the matrix A, the set $V = \left\{x \in U \,\middle|\, Hf(a)^{(k)} > 0 \text{ for } k = 1,2,\cdots,n\right\}$ is open. Choose $r > 0$ so that $D(a,r) \subseteq V$. Then we have $u^T Hf(c)\, u > 0$ for all $0 \neq u \in \mathbf{R}^n$ and all $c \in D(a,r)$. Let $x \in D(a,r)$ with $x \neq a$. By Taylor's Theorem, we have

$$f(x) - f(a) - Df(a)(x-a) = (x-a)^T Hf(c)\,(x-a)$$

for some $c \in [a,x]$. Since $Df(a) = 0$ and $Hf(c)$ is positive-definite, we have $f(x) - f(a) > 0$. Thus f has a local minimum value at a. This proves Part (1) and Part (2) is similar.

Let us prove Part (3). Suppose there exists $0 \neq u \in \mathbf{R}^n$ such that $u^T Hf(a)\, u > 0$. Given $r > 0$, scale the vector u if necessary so that $[a,u] \subseteq D(a,r) \cap U$. Let $\alpha(t) = a + tu$ and let $g(t) = f\big(\alpha(t)\big)$ for $0 \leq t \leq 1$. As in the proof of Taylor's Theorem, we have

(or for $-\delta < t < 1 + \delta$) for some $\delta > 0$.

$$g'(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\big(\alpha(t)\big)\, u_i = Df\big(\alpha(t)\big)\, u \text{ , and}$$

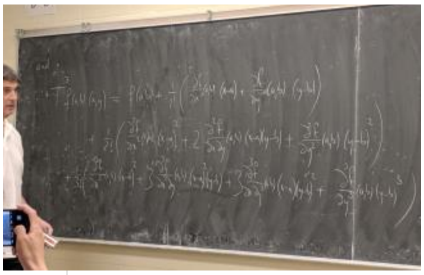
$$g''(t) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}\big(\alpha(t)\big)\, u_i u_j = u^T Hf\big(\alpha(t)\big)\, u\,.$$

Since $g(0) = f(a)$, $g'(0) = Df(a)\, u = 0$ and $g''(0) = u^T Hf(a)\, u > 0$, it follows from single-variable calculus that we can choose t_0 with $0 < t_0 < 1$ so that $g(t_0) > g(0)$. When $x = \alpha(t_0)$ we have $x \in D(a,r) \cap U$ and $f(x) = f\big(\alpha(t_0)\big) = g(t_0) > g(0) = f(a)$, and so f does not have a local maximum value at a. Similarly, if there exists $0 \neq v \in \mathbf{R}^n$ such that $v^T Hf(a)\, v < 0$ then f does not have a local minimum value at a. Thus when $Hf(a)$ is indefinite, f has a saddle point at a.

6.18 Exercise: Find and classify the critical points of the following functions $f : \mathbf{R}^2 \to \mathbf{R}$.

(a) $f(x,y) = x^3 + 2xy + y^2$ (b) $f(x,y) = x^3 + 3x^2y - 6y^2$ (c) $f(x,y) = x^2 y\, e^{-x^2 - 2y^2}$

4



When $A \in M_n(\mathbb{R})$ is symmetric (meaning that $A^T = A$)
all the eigenvalues of A are real. A is orthogonally diagonalizable so there exists an orthogonal matrix $P \in O_n(\mathbb{R}) = O(n,\mathbb{R})$ (so $P^T P = I$) (ie $P^T = P^{-1}$)
such that $P^T A P = D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$
where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A.

A is positive-definite
$\leftrightarrow u^T A u > 0$ for all $0 \neq u \in \mathbb{R}^n$
$\leftrightarrow$ all the eigenvalues of A are positive

A is positive-semidefinite
$\leftrightarrow u^T A u \geq 0$ for all $u \in \mathbb{R}^n$
$\leftrightarrow$ all the eigenvalues of A are non-negative
A is negative-definite $\leftrightarrow$...

A is indefinite
$\leftrightarrow \exists\, u, v \in \mathbb{R}^n$ such that $u^T A u > 0$ and $v^T A v < 0$
$\leftrightarrow$ A has at least one positive eigenvalue and at least one negative eigenvalue

$F(u,v) = v^T A u$

$Q(u) = F(u,u) = u^T A u$

For linear maps
$A \mapsto P^{-1} A P$

Similar matrix

For symmetric bilinear form (and quadratic forms)
$A \mapsto P^T A P$
(where P is invertible)

Congruent

Choose $r > 0$ so $B(a,r) \subseteq U$ and $Hf(x)$ is positive definite at every $x \in B(a,r)$

eg. For $f(x,y) = \frac{4}{2+x+y^2}$ defined when $2 + x + y^2 \neq 0$. Find $T^2 f_{(0,0)}(x,y)$

Solution:

$\frac{\partial f}{\partial x} = -\frac{4}{(2+x+y^2)^2}$

$\frac{\partial f}{\partial y} = -\frac{8y}{(2+x+y^2)^2}$

$\frac{\partial^2 f}{\partial x^2} = \left(\frac{8}{(2+x+y^2)^3}\right)$

$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = \frac{16y}{(2+x+y^2)^3}$

$\frac{\partial^2 f}{\partial y^2} = \cdots =$

So $f(0,0) = 2$
$\frac{\partial f}{\partial x}(0,0) = -1$
$\frac{\partial f}{\partial y}(0,0) = 0$

$\frac{\partial^2 f}{\partial x^2}(0,0) = 1$

$\frac{\partial^2 f}{\partial x \partial y}(0,0) = \frac{\partial^2 f}{\partial y \partial x}(0,0) = 0$

$\frac{\partial^2 f}{\partial y^2}(0,0) = -2$

Hence
$T^2 f(0,0)(x,y) = 2 - x + \frac{1}{2!}\left((x-0)^2 + 2 \cdot 0 \cdot xy - 2y^2\right)$
$= 2 - x + \frac{1}{2}x^2 - y^2$

Eg.
Find and classify the critical points of $f(x,y) = x^3 + 3x^2y - 6y^2$

Solution:
Critical point : Definition __

Classify: determine whether max or min or saddle point

$Df(x,y) = \begin{pmatrix} 3x^2 + 6xy & 3x^2 - 12y \end{pmatrix}$

$Hf = \begin{pmatrix} 6x + 6y & 6x \\ 6x & -12 \end{pmatrix}$

We have
$Df(x,y) = (0,0) \Leftrightarrow x(x+2y) = 0$ and $x^2 = 4y$
$\Leftrightarrow \big((x = 0 \text{ or } x = -2y) \text{ and } x^2 = 4y\big)$
$\Leftrightarrow (x = 0 \text{ and } x^2 = 4y)$ or $(x = -2y \text{ and } x^2 = 4y)$

$\Leftrightarrow (x,y) = (0,0)$ or $(-2,1)$

Thus, f has critical points at $(0,0)$ and $(-2,1)$.
We have
$Hf(0,0) = \begin{pmatrix} 0 & 0 \\ 0 & -12 \end{pmatrix}$ and $Hf(-2,1) = \begin{pmatrix} -6 & -12 \\ -12 & -12 \end{pmatrix}$

Fails the test

For $A = \begin{pmatrix} -6 & -12 \\ -12 & -12 \end{pmatrix}$
$\det A^1 = -6$
$\det A^2 = 36(-2)$

Since $\det A^2 \neq 0$, the eigenvalues of A are non-zero. Since A neither positive definite nor negative definite. A must hae one positive eigenvalue and one negative eigenvalue.
Alternatively,
$A = -6B, B = \begin{pmatrix} 1 & 2 \\ 2 & 2 \end{pmatrix}$

$f_B(x) = \det\begin{pmatrix} x-1 & -2 \\ -2 & x-2 \end{pmatrix} = x^2 - 3x - 2$

So the eigenvalues are $\lambda = \frac{3 \pm \sqrt{9+4}}{2}$

So B, hence A, has one positive and one negative eigenvalue.

Thus, f has a saddle point at $(-2,1)$
The second Derivative Test gives no information about the point $(0,0)$
We have
$f(x,y) = x^3 + 3x^2y - 6y^2$

So for $\alpha(t) = (0,t)$
and $g(t) = f(\alpha(t))$ we have
$g(t) = f(0,t) = -6t^2$
So $g(t)$ has a local max at $t = 0$.
and for $\beta(t) = (t,0)$ and $h(t) = f(\beta(t))$ we have $h(t) = f(t,0) = t^3$

Thus, neither. Saddle point at $(0,0)$

eg. Find the absolute max and min values of $f(x,y) = 4xy - x^4 - 2y^2$ on the compact set $\bar{B}((0,0),2)$

Solution:

$Df = (4y - 4x^3, 4x - 4y)$
$Hf = \begin{pmatrix} -12x^2 & 4 \\ 4 & -4 \end{pmatrix}$

$Df(x,y) = (0,0) \Leftrightarrow y = x^3$ and $x = y$
$\Leftrightarrow x = x^3$ and $x = y$
$\Leftrightarrow ((x$ or $x = \pm 1)$ and $y = x)$
$\Leftrightarrow (x,y) = (0,0)$ or $\pm(1,1)$

We have $f(0,0) = 0, f(\pm(1,1)) = 1$

Let us classify the critical points (for fun)
$Hf(0,0) = \begin{pmatrix} 0 & 4 \\ 4 & -4 \end{pmatrix} = 4A, A = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}$

$f_A(x) = \det\begin{pmatrix} x & -1 \\ -1 & x+1 \end{pmatrix} = x^2 + x - 1$

The eigenvalues of A are $\lambda = \frac{-1 \pm \sqrt{1+4}}{2}$

So A, hence also $Hf(0,0)$ havs one positive and one negative eigenvalue.

So f has a saddle point at $(0,0)$

$Hf(\pm(1,1)) = \begin{pmatrix} -12 & 4 \\ 4 & -4 \end{pmatrix} = 4B$
where $B = \begin{pmatrix} -3 & 1 \\ 1 & -1 \end{pmatrix}$
$\det B^{(1)} = -3, \det B^{(2)} = 2$

So B, hence also $Hf(\pm(1,1))$, is negative definite. So f has a local max at $(x,y) = \pm(1,1)$ with $f(\pm(1,1)) = 1$.

We need to consider $\partial\bar{B}(0,2) = S(0,2) = \{(x,y) | x^2 + y^2 = 4\}$

$f(x,y) = 4xy - x^4 - 2y^2$

Let $\alpha(t) = (2\cos t, 2\sin t)$
and consider

$$g(t) = f(\alpha(t)) = 16\cos t \sin t - 16\cos^4 t - 8\sin^2 t$$

We need to find the max and min values of $g(t)$ for $0 \leq t \leq 2\pi$
Note that $g(t) = 8\cos 2t - 4(1 + \cos 2t)^2 - 4(1 - \cos 2t) = 8\sin 2t - 4 - 8\cos 2t - 4\cos^2 2t\, 4 + 4\cos 2t$
$= 8\sin 2t - 8 - 4\cos 2t - 4\cos^2 2t$
$= 4(2\sin\theta - 2 - \cos\theta - \cos^2\theta)$
where $\theta = 2t$
$= 4h(\theta)$

$h'(\theta) = 2\cos\theta + \sin\theta + 2\sin\theta\cos\theta$
$h'(\theta) = 0 \Leftrightarrow 2\cos\theta\,(1 + \sin\theta) + \sin\theta = 0$
$\Leftrightarrow 2\cos\theta = -\dfrac{\sin\theta}{1 + \sin\theta}$

we give up

Verify that $g(t)$ is $4(2\sin\theta - 2 - \cos\theta - \cos^2\theta) \leq 0$ for all θ.

chap7

2019年7月3日 17:44

chap7

Chapter 7. Introduction to Integrals

7.1 Remark: In this chapter we give an informal introduction to integration. We might include some additional chapters later in which we provide a rigorous theoretical treatment of integration. For now, we explore some computational aspects of integration.

7.2 Theorem: *(Fubini's Theorem) When $D = \{x \in \mathbf{R} | a \le x \le b\}$ and $f : D \subseteq \mathbf{R} \to \mathbf{R}$ is continuous, the integral of f on D is written as*

$$\int_D f\, dL = \int_D f(x)\, dL = \int_{x=a}^{b} f(x)\, dx.$$

When $D = \{(x, y) \in \mathbf{R}^2 | a \le x \le b\,,\ g(x) \le y \le h(x)\}$ and $f : D \subseteq \mathbf{R}^2 \to \mathbf{R}$ is continuous, the integral of f on D is given by

$$\int_D f\, dA = \iint_D f(x,y)\, dA = \iint_D f(x,y)\, dx\, dy = \int_{x=a}^{b}\left(\int_{y=g(x)}^{h(x)} f(x,y)\, dy\right) dx.$$

When $D = \{(x, y) \in \mathbf{R}^2 | c \le y \le d\,,\ k(y) \le x \le l(x)\}$ and $f : D \subseteq \mathbf{R}^2 \to \mathbf{R}$ is continuous, the integral of f on D is given by

$$\int_D f\, dA = \iint_D f(x,y)\, dA = \iint_D f(x,y)\, dx\, dy = \int_{y=c}^{d}\left(\int_{x=k(y)}^{l(y)} f(x,y)\, dx\right) dy.$$

More generally, when $D \subseteq \mathbf{R}^2$ is a union $D = \bigcup_{i=1}^{n} D_i$ of sets $D_i \subseteq \mathbf{R}^2$ which only overlap along their boundaries, with each set D_i of one of the above two forms, the integral of f on D is

$$\int_D f\, dA = \sum_{i=1}^{n} \int_{D_i} f\, dA.$$

When $D = \{(x, y, z) \in \mathbf{R}^3 | a \le x \le b\,,\ g(x) \le y \le h(x)\,,\ k(x,y) \le z \le l(x,y)\}$ and $f : D \subseteq \mathbf{R}^3 \to \mathbf{R}$ is continuous, the integral of f on D is given by

$$\int_D f\, dV = \iiint_D f(x,y,z)\, dV = \iiint_D f(x,y,z)\, dx\, dy\, dz$$
$$= \int_{x=a}^{b}\left(\int_{y=g(x)}^{h(x)}\left(\int_{z=k(x,y)}^{l(x,y)} f(x,y,z)\, dz\right) dy\right) dx.$$

There are similar formulas in the case that the roles of x, y and z are permuted. More generally, when $D \subseteq \mathbf{R}^3$ is a union $D = \bigcup_{i=1}^{n} D_i$ of sets $D_i \subseteq \mathbf{R}^3$ which only overlap along their boundaries, with each set D_i of the above form or of a similar form with x, y and z permuted, the integral of f on D is

$$\int_D f\, dA = \sum_{i=1}^{n} \int_{D_i} f\, dA.$$

Proof: We may provide a proof later.

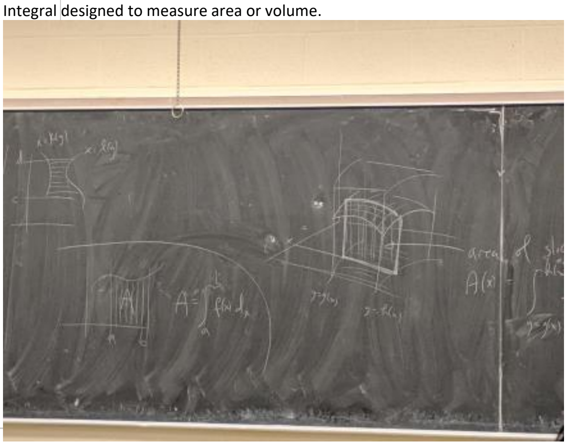
1

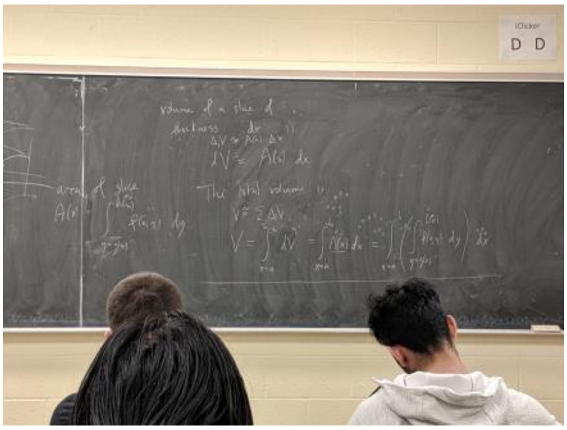
$D_i^o \cap D_j^o = \emptyset$ for all $i \ne j$,

Ex. Let D be the triangle with vertices $(0,-1), (2,-1), (2,3)$

Find $\int_D 2xy\, dA$

Integral designed to measure area or volume.





7.3 Note: When $D \subseteq \mathbf{R}^2$, the integral of the constant function 1 on D measures the **area** of the region D and, when $f : D \subseteq \mathbf{R}^2 \to \mathbf{R}$, the integral of f on D measures the **signed volume** of the region between the graph of f and the region D and, in the case that D represents the shape of a flat object and the function $f : D \subseteq D \to \mathbf{R}$ represents its **density** (or the **charge density**), the integral of f on D measures the total **mass** (or **charge**) of the object.

When $D \subseteq \mathbf{R}^3$, the integral of the constant function 1 on D measures the **volume** of the region D and, when D represents the shape of a solid object and $f : D \subseteq \mathbf{R}^3 \to \mathbf{R}$ represents its **density** (or **charge density**), the integral of f on D measures the total **mass** (or **charge**) of the object.

7.4 Exercise: Let D be the triangle in $\mathbf{R}^2$ with vertices at $(0,-1)$, $(2,1)$ and $(2,3)$. Find $\int_D 2xy\, dA$.

7.5 Exercise: Find the volume of the region in $\mathbf{R}^3$ which lies above the paraboloid $z = x^2 + y^2$ and below the plane $z = 2x$.

7.6 Exercise: Find the mass of the tetrahedron with vertices at $(0,0,0)$, $(2,0,0)$, $(2,2,0)$ and $(2,2,2)$ given that the density is given by $\rho(x,y,z) = 2xy(3-z)$.

7.7 Definition: Let U and V be open sets in $\mathbf{R}^n$, let $C = \overline{U}$ and $D = \overline{V}$. An **orientation preserving change of coordinates map** from C to D is a continuous map $g : C \to D$ such that the map $g : U \to V$ is invertible and $\mathcal{C}^1$ with $\det\big(Dg(a)\big) > 0$ for all $a \in U$, and an **orientation reversing change of coordinates map** from C to D is a continuous map $g : C \to D$ such that the map $g : U \to V$ is invertible and $\mathcal{C}^1$ with $\det\big(Dg(a)\big) < 0$ for all $a \in U$.

7.8 Example: Three important orientation preserving change of coordinates maps are the **polar coordinates map** in $\mathbf{R}^2$, which is given by

$$(x,y) = g(r,\theta) = \big(r\cos\theta\,,\, r\sin\theta\big) \text{ with } \det Dg(r,\theta) = r\,,$$

the **cylindrical coordinates map** in $\mathbf{R}^3$, which is given by

$$(x,y,z) = g(r,\theta,z) = \big(r\cos\theta\,,\, r\sin\theta\,,\, z\big) \text{ with } \det Dg(r,\theta,z) = r,$$

and the **spherical coordinates map** in $\mathbf{R}^3$, which is given by

$$(x,y,z) = g(r,\phi,\theta) = \big(r\sin\phi\cos\theta\,,\, r\sin\phi\sin\theta\,,\, r\cos\phi\big) \text{ with } \det Dg(r,\phi,\theta) = r^2\sin\phi.$$

2

Usually smooth bijective map

7.9 Theorem: *(Change of Variables) When $D = [a,b] \subseteq \mathbf{R}$, and $g : C \subseteq \mathbf{R} \to D \subseteq \mathbf{R}$ is a change of variables map from C to D given by $x = g(u)$ with inverse $u = h(x)$, and $f : D \subseteq \mathbf{R} \to \mathbf{R}$ is continuous, we have*

$$\int_{x=a}^{b} f(x)\,dx = \int_D f(x)dx = \int_C f\big(g(u)\big)\Big|\det Dg(u)\Big|\,du = \int_{u=h(a)}^{h(b)} f\big(g(u)\big)g'(u)\,du.$$

When $f : D \subseteq \mathbf{R}^2 \to \mathbf{R}$ is continuous and $g : C \subseteq \mathbf{R}^2 \to D \subseteq \mathbf{R}^2$ is a change of variables map from C to D given by $(x,y) = g(u,v)$, we have

$$\iint_D f(x,y)\,dx\,dx = \iint_C f\big(g(u,v)\big)\Big|\det Dg(u,v)\Big|\,du\,dv.$$

When $f : D \subseteq \mathbf{R}^3 \to \mathbf{R}$ is continuous and $g : C \subseteq \mathbf{R}^3 \to D \subseteq \mathbf{R}^3$ is a change of variables map from C to D given by $(x,y,z) = g(u,v,w)$, we have

$$\iiint_D f(x,y,z)\,dx\,dy\,dz = \iiint_C f\big(g(u,v,w)\big)\Big|\det Dg(u,v,w)\Big|\,du\,dv\,dw.$$

Proof: We may provide a proof in a later chapter.

7.10 Exercise: Find the area inside the cardioid $r = 2 + 2\cos\theta$.

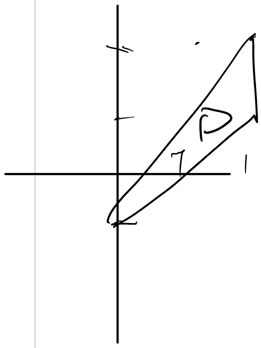
7.11 Exercise: Find the volume of the region under the graph of $z = e^{-(x^2+y^2)}$.

7.12 Exercise: Find the volume of the region which lies inside the sphere $x^2+y^2+z^2 = 4$ and inside the cylinder $x^2 - 2x + y^2 = 0$.

7.13 Exercise: Find the mass of the ball $x^2+y^2+z^2 \le 4$ given that the density is given by $\rho(x,y,z) = 1 - \frac{1}{2}\sqrt{x^2+y^2+z^2}$.

7.14 Definition: Let $n = 2$ or 3, let $\alpha : [a,b] \subseteq \mathbf{R} \to \mathbf{R}^n$ be continuous on $[a,b]$ and $\mathcal{C}^1$ in (a,b), let C be the curve in $\mathbf{R}^n$ which is given parametrically by $(x,y) = \alpha(t)$ or by $(x,y,z) = \alpha(t)$ for $a \le t \le b$, and let $f : C \subseteq \mathbf{R}^n \to \mathbf{R}$ be continuous on $C = \text{Range}(\alpha)$. Then we write $dL = \big|\alpha'(t)\big|\,dt$ and we define the (curve) **integral of f on C** to be

7.4 Solution1:



Note that

D = $\{(x,y)\,|\,0 \le x \le 2, x-1 \le y \le 2x-1\}$

So $\int_D f = \int_{x=0}^{2}\int_{y=x-1}^{2x-1} 2xy\,dy\,dx$

$= \int_{x=0}^{2} \left[xy^2\right]_{y=x-1}^{2x-1} dx$

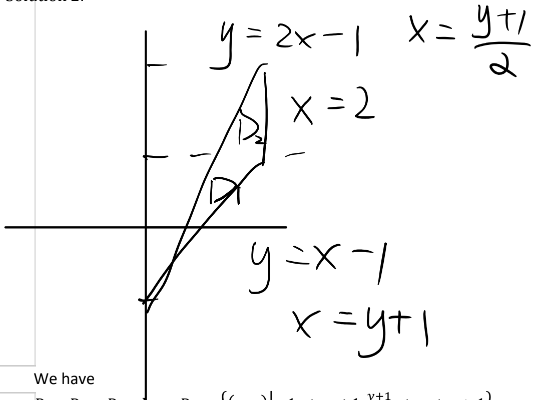
$= \int_{x=0}^{2} x(2x-1)^2 - x(x-1)^2\,dx$

$= \int_{x=0}^{2} 3x^2 - 2x^2\,dx$

$= \left[\frac{3}{4}x^4 - \frac{2}{3}x^3\right]_{x=0}^{2}$

$= 12 - \frac{16}{3} = \frac{20}{3}$

Solution 2:



We have

$D = D_1 \cup D_2$ where $D_1 = \left\{(x,y)\,\middle|\,-1 \le y \le 1, \frac{y+1}{2} \le x \le y+1\right\}$

$D_2 = \left\{(x,y)\,\middle|\,1 \le y \le 3, \frac{y+1}{2} \le x \le 2\right\}$

So we have

$\int_D f = \int_{D_1} f + \int_{D_2} f$

$= \int_{y=-1}^{1}\int_{x=\frac{y+1}{2}}^{y+1} 2xy\,dx\,dy + \int_{y=1}^{3}\left(\int_{x=\frac{y+1}{2}}^{2} 2xy\,dx\right)dy$

$= \int_{y=-1}^{1} \left[x^2y\right]_{x=\frac{y+1}{2}}^{y+1} dy + \int_{y=1}^{3} \left[x^2y\right]_{x=\frac{(y+1)}{2}}^{2} dy$

$= \int_{y=-1}^{1} \frac{3}{4}(y+1)^2 y\,dy + \int_{y=1}^{3} 4y - \frac{1}{4}(y+1)^2 y\,dy$

$= \cdots$

7.5 Solution:

7.14 Definition: Let $n = 2$ or 3, let $\alpha : [a, b] \subseteq \mathbf{R} \to \mathbf{R}^n$ be continuous on $[a, b]$ and $\mathcal{C}^1$ in (a, b), let C be the curve in $\mathbf{R}^n$ which is given parametrically by $(x, y) = \alpha(t)$ or by $(x, y, z) = \alpha(t)$ for $a \le t \le b$, and let $f : C \subseteq \mathbf{R}^n \to \mathbf{R}$ be continuous on $C = \text{Range}(\alpha)$. Then we write $dL = |\alpha'(t)|\, dt$ and we define the (curve) **integral of** f **on** C to be

$$\int_\alpha f\, dL = \int_C f\, dL = \int_{t=a}^{b} f\big(\alpha(t)\big)\big|\alpha'(t)\big|\, dt.$$

When C is a union $C = \bigcup_{k=1}^{m} C_k$ of curves C_k as above, we define $\int_C f dA = \sum_{k=1}^{m} \int_{C_k} f\, dA$.

Let D be the closure of a bounded open set U in $\mathbf{R}^2$, let $\sigma : D \subseteq \mathbf{R}^2 \to \mathbf{R}^3$ be continuous in D and $\mathcal{C}^1$ in U, let S be the surface in $\mathbf{R}^3$ which is given parametrically by $(x, y, z) = \sigma(s, t)$, and let $f : S \subseteq \mathbf{R}^3 \to \mathbf{R}$ be continuous on S. Then we write $dA = |\sigma_s \times \sigma_t|\, ds\, dt$ and we define the (surface) **integral of** f **on** S to be

$$\iint_\sigma f\, dA = \iint_S f\, dA = \iint_D f\big(\sigma(s, t)\big)|\sigma_s \times \sigma_t|\, ds\, dt.$$

where $\sigma_s = \left(\frac{\partial x}{\partial s}(s,t), \frac{\partial y}{\partial s}(s,t), \frac{\partial z}{\partial s}(s,t)\right)^T$ and $\sigma_t = \left(\frac{\partial x}{\partial s}(s,t), \frac{\partial y}{\partial s}(s,t), \frac{\partial z}{\partial s}(s,t)\right)^T$.

When S is a union $S = \bigcup_{k=1}^{m} S_k$ of surfaces S_k as above, we define $\int_S f dA = \sum_{k=1}^{m} \int_{S_k} f\, dA$.

7.15 Note: When C is a curve in $\mathbf{R}^n$ with $n = 2$ or 3, which is given by $(x, y) = \alpha(t)$ or by $(x, y, z) = \alpha(t)$ for $a \le t \le b$, the integral of the constant function 1 on C measures the **length** (or **arclength**) of the curve C, and in the case that C represents the shape of a physical object and the function $f : C \subseteq \mathbf{R}^2 \to \mathbf{R}$ represents its density (or charge density), the integral of f on C measures the total **mass** (or **charge**) of the object.

When S is a surface in $\mathbf{R}^3$ which is given by $(x, y, z) = \sigma(s, t)$ for $(s, t) \in D \subseteq \mathbf{R}^2$, the integral of the constant function 1 on S measures the **area** (or **surface-area**) of the surface S, and in the case that S represents the shape of a physical object and the function $f : S \to \mathbf{R}$ represents its **density** (or **charge density**), the integral of f on S measures the total **mass** (or **charge**) of the surface.

7.16 Exercise: Find the arclength of the helix $\alpha(t) = (t, \cos t, \sin t)$ for $0 \le t \le 2\pi$.

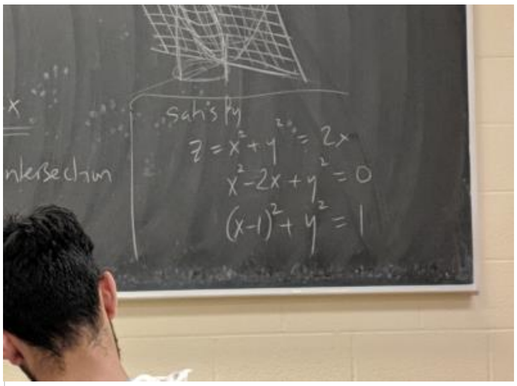
7.17 Exercise: Find the surface area of the torus given by

$$(x, y, z) = \sigma(\theta, \phi) = \Big((2 + \cos\phi)\cos\theta\,, (2 + \cos\phi)\sin\theta\,, \sin\phi\Big)$$

for $0 \le \theta \le 2\pi$ and $0 \le \phi \le 2\pi$.

7.18 Exercise: Find the mass of the hollow sphere $x^2 + y^2 + z^2 = 1$ when the density (mass per unit area) is given by $\rho(x, y, z) = 3 - z$.

7.19 Exercise: Find the mass of the curve of intersection of the parabolic sheet $z = x^2$ with the paraboloid $z = 2 - x^2 - 2y^2$, when the density (mass per unit length) is given by $\rho(x, y, z) = |xy|$.



Points on the curve of intersection satisfies $z = x^2 + y^2 = 2x$

$x^2 - 2x + y^2 = 0$

$(x-1)^2 + y^2 = 1$

The intersection

It lies above the circle

The given solid D is the set

$D = \{(x, y, z) | (x-1)^2 + y^2 \le 1, x^2 + y^2 \le z \le 2x\}$

$= \{(x, y, z) | 0 \le x \le 2, -\sqrt{2x - x^2} \le y \le \sqrt{2x - x^2}, x^2 + y^2 \le z \le 2x\}$

The volume is

$V = \iiint_D 1\, dV = \int_{x=0}^{2} \int_{y=-\sqrt{2x-x^2}}^{\sqrt{2x-x^2}} \int_{z=x^2+y^2}^{2x} 1\, dz\, dy\, dx$

$= \int_{x=0}^{2} \int_{y=-\sqrt{2x-x^2}}^{\sqrt{2x-x^2}} (2x - x^2 - y^2)\, dy\, dx$

$= \int_{x=0}^{2} \left[(2x - x^2)y = \frac{1}{3}y^3\right]_{y=-\sqrt{2x-x^2}}^{\sqrt{2x-x^2}} dx$

$= \int_{x=0}^{2} \frac{2}{3}(2x - x^2)^{\frac{3}{2}} + \frac{2}{3}(2x - x^2)^{\frac{3}{2}}\, dx$

$= \int_{x=0}^{2} \frac{4}{3}(2x - x^2)^{\frac{3}{2}}\, dx$

$= \int_{x=0}^{2} \frac{4}{3}(1 - (x-1)^2)^{\frac{3}{2}}\, dx$

Let $u = x - 1$

$du = dx$

$= \int_{u=-1}^{1} \frac{4}{3}(1 - u^2)^{\frac{3}{2}}\, du$

$= \int_{\theta=-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{4}{3}(\cos^3\theta)\cos\theta\, d\theta$

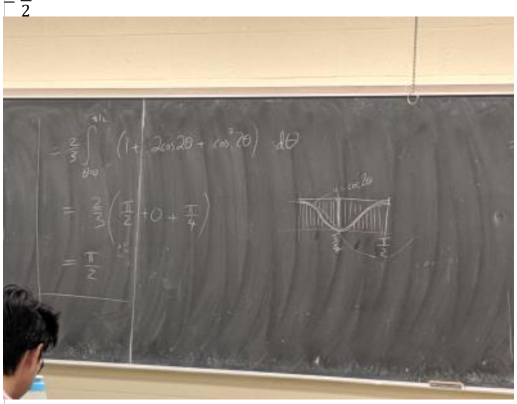
$= \int_{\theta=0}^{\frac{\pi}{2}} \frac{8}{3}(\cos^4\theta)\, d\theta$ By Symmetry

$= \int_{\theta=0}^{\frac{\pi}{2}} \frac{8}{3}\left(\frac{1 + \cos 2\theta}{2}\right)^2 d\theta$

$= \frac{2}{3}\int_{\theta=0}^{\frac{\pi}{2}} (1 + 2\cos 2\theta + \cos^2 2\theta)\, d\theta$

$= \frac{2}{3}\left(\frac{\pi}{2} + 0 + \frac{\pi}{4}\right)$

$= \frac{\pi}{2}$

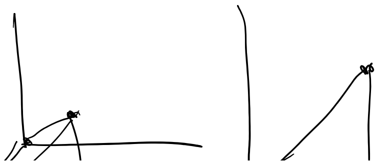


7.6 Exercise

Note that the given tetrahedron is the set $D = \{(x, y, z) | 0 \le x \le 2, 0 \le y \le x, 0 \le z \le y\}$

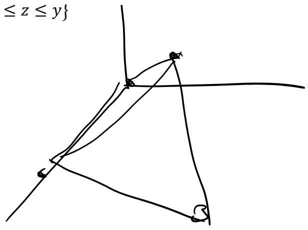
So the mass is

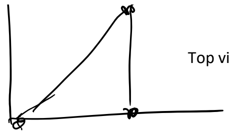
$M = \int^{2} \int^{x} \int^{y} 2xy(3 - z)\, dz\, dy\, dx$



Top view

Note that the given tetrahedron is the set $D = \{(x, y, z) | 0 \le x \le 2, 0 \le y \le x, 0 \le z \le y\}$

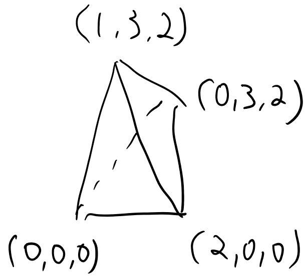




Top view

So the mass is

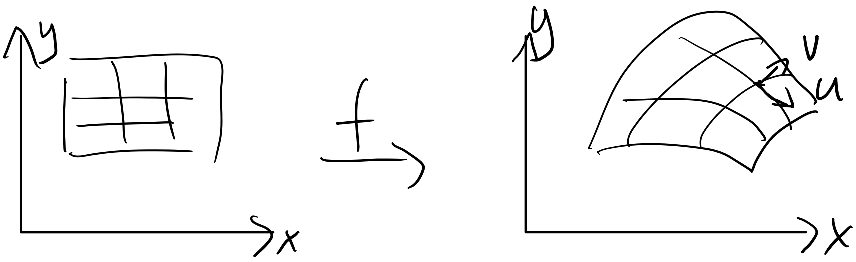
$$M = \int_{x=0}^{2}\int_{y=0}^{x}\int_{z=0}^{y} 2xy(3-z)\,dz\,dy\,dx$$
$$= \int_{x=0}^{2}\int_{y=0}^{x} \left[6xyz - xyz^2\right]_{z=0}^{y} dy\,dx$$
$$= \int_{x=0}^{2}\int_{y=0}^{x} 6xy^2 - xy^3\,dy\,dx$$
$$= \int_{x=0}^{2} \left[2xy^3 - \frac{1}{4}xy^4\right]_{y=0}^{x} dx$$
$$= \int_{x=0}^{2} 2x^4 - \frac{1}{4}x^5\,dy$$
$$= \left[\frac{2}{5}x^5 - \frac{1}{24}x^6\right]_{x=0}^{2}$$
$$= \frac{64}{5} - \frac{64}{24} = \frac{152}{15}$$



$z = y$

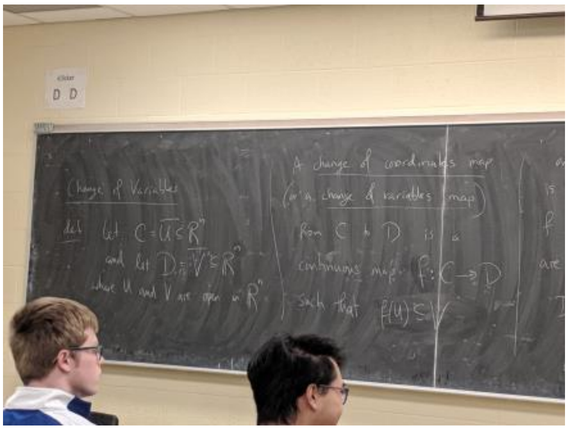
Change of variables

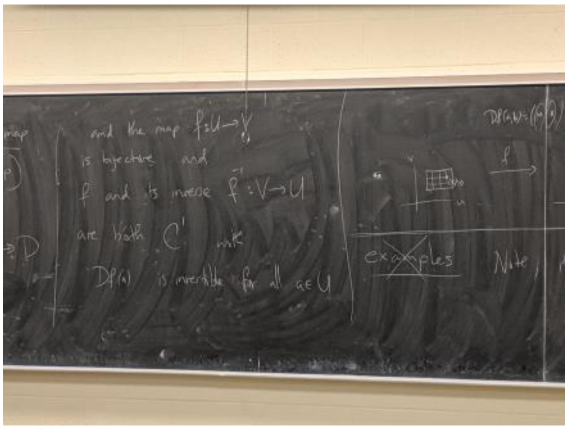
$Df(a,b) = (u,v)$



Examples:

Note : if u is connected. then either $Df(x) > 0$ for all $x \in U$ or $Df(x) < 0$ for all $x \in U$.
(because for $\phi(x) = \det Df(x)$, $\phi: U \to \mathbb{R} \setminus \{0\}$ is continuous and $\phi(u)$ is connected.
It follows that $\phi(u) \subseteq (-\infty, 0)$ or $\phi(u) \subseteq (0, \infty)$)



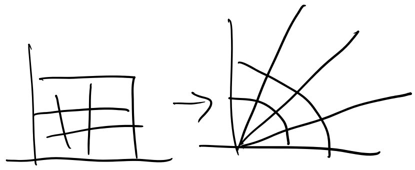


parallelogram
Trapezoid

When $\det Df(x) > 0$ for all $x \in U$, we say that f preserves orientation and when $\det Df(x) < 0$ for all $x \in U$, we say that f reverses orientation.

Examples:

When the polar coordinates map $g: \bar{U} \subseteq \mathbb{R}^2 \to \bar{V} \subseteq \mathbb{R}^2$ where $U = \{(r,\theta) | r > 0, \alpha < \theta < \alpha + 2\pi\}$ and $V = \mathbb{R}^2 = \{(r\cos\alpha, r\sin\alpha) | r \geq 0\}$ is given by $(x, y) = g(r,\theta) = (r\cos\theta, r\sin\theta)$



The polar coordinates change of variables map is given by $(x,y) = g(r,\theta) = (r\cos\theta, r\sin\theta)$

We have $Dg = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$

and $\det Dg = r\cos^2\theta + r\sin^2\theta = r > 0$
for all $(r,\theta) \in U$.

So the polar coordinates map is orientation preserving.

The cylindrical coordinates map is given by
$(x,y,z) = g(r,\theta,z) = (r\cos\theta, r\sin\theta, z)$

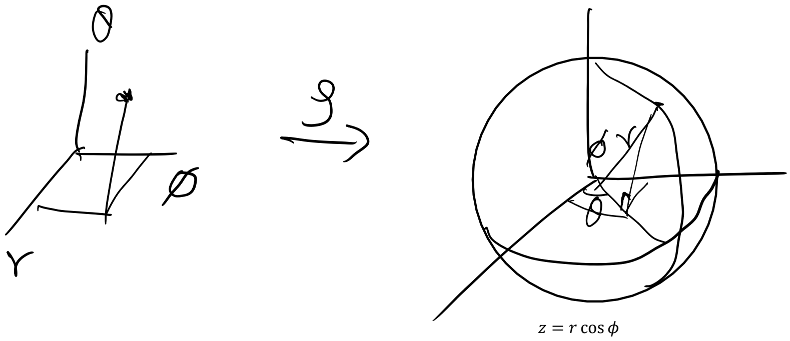
$$Dg = \begin{pmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

So that $\det Dg = r$

The spherical coordinates map

$(x,y,z) = g(r,\phi,\theta) = (r\cos\phi\, cos\theta, r\cos\phi\sin\theta, r\sin\phi)$

???



Exercise: Find $Df(r,\phi,\theta)$
and $\det Df(r,\phi,\theta)$ and simplify.

$z = r\cos\phi$
$y = r\sin\phi\sin\theta$
$x = r\sin\phi\cos\theta$

$$Dg = \begin{pmatrix} \sin\phi\cos\theta & r\cos\phi\cos\theta & -r\sin\phi\sin\theta \\ \sin\phi\sin\theta & r\cos\phi\sin\theta & r\sin\phi\cos\theta \\ \cos\phi & -r\sin\phi & 0 \end{pmatrix}$$

$\det Dg$
$= \cos\phi\left(r^2\sin\phi\cos\phi\cos^2\theta + r^2\sin\phi\cos\phi\sin^2\theta\right) + r\sin\phi\left(r\sin^2\phi\cos^2\theta + r\sin^2\phi\sin^2\theta\right)$
$= \cos\phi\left(r^2\sin\phi\cos\phi\right) + r\sin\phi\left(r\sin^2\phi\right)$
$= r^2\sin\phi$

Thm (Change of coordinates)
When $U, V \subseteq \mathbb{R}^n$ are open, $C = \bar{U}, D = \bar{V}$ and $g: C \to D$ is a change of coordinates map and $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ is continuous, we have

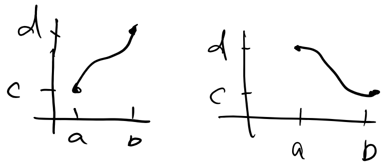
$$\int_D f = \int_C (f \circ g)\, |\det Dg|$$

For $U, V \subseteq \mathbb{R}^3, f: D = \bar{V} \subseteq \mathbb{R}^3 \to \mathbb{R}$
Writing $(x,y,z) = g(u,v,w)$
we have

$$\iiint_D f(x,y,z)\, dx\, dy\, dz = \iiint_C f(g(u,v,w)) |\det Dg(u,v,w)|\, du\, dv\, dw$$

For $U = (c,d), C = \bar{U} = [c,d] \subseteq \mathbb{R}$
$V = (a,b), D = \bar{V} = [a,b] \subseteq \mathbb{R}$
$f: D = [a,b] \to \mathbb{R}$
$x = g(u)$ and $u = u(x) = g^{-1}(x)$

We have



$$\int_D f(x)\, dx = \int_C f(g(u))\, |\det Dg(u)| du$$

That is

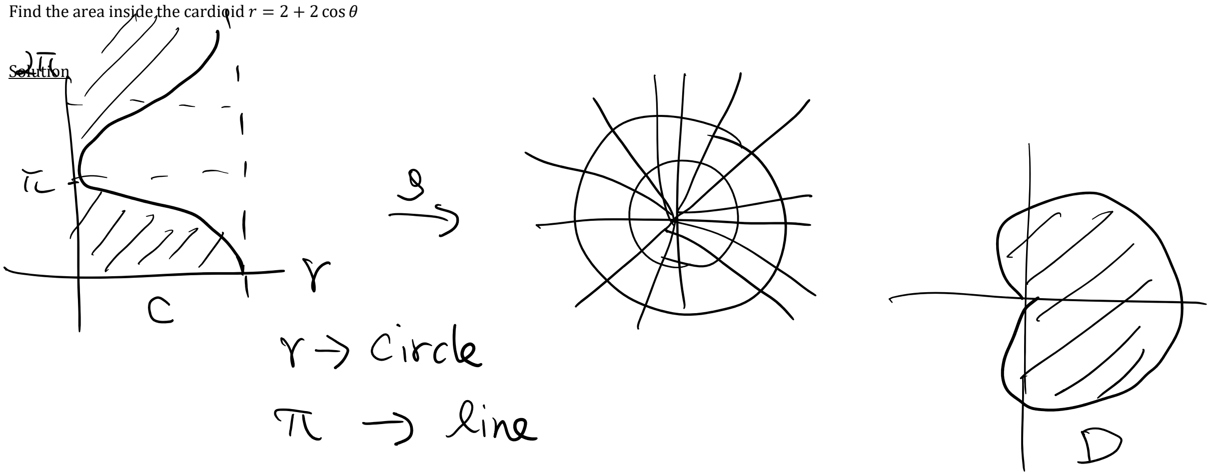
$$\int_a^b f(x)\, dx = \int_c^d f(g(u)) |g'(u)|\, du = \int_{u=u(a)}^{u(b)} f(g(u)) g'(u)\, du$$

Why take off the absolute value sign?
(when $g'(u) > 0$ for all u, $u(a) = c$ and $u(b) = d$)

Eg. Example 7.10

Find the area inside the cardioid $r = 2 + 2\cos\theta$

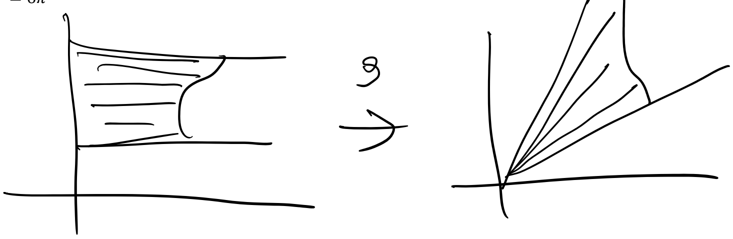
Solution

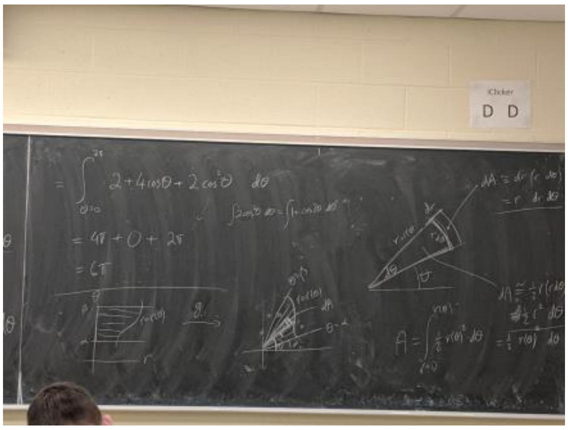


(The fact that $r = 0$ when $\theta = \pi$ implies in the image curve as θ approaches π, the corresponding point in the (x,y) plane approaches the origin $r=0$ in the direction of the ray $\theta=\pi$.

The area is $A=\iint_D 1\,dx\,dy=\iint_C 1\,\left|\det Dg(r,\theta)\right|dr\,d\theta=\iint_C r\,dr\,d\theta=\int_{\theta=0}^{2\pi}\int_{r=0}^{2+2\cos\theta} r\,dr\,d\theta=$ $\int_{\theta=0}^{2\pi}\left[\frac{1}{2}r^2\right]_{r=0}^{2+2\cos\theta}d\theta=\int_{\theta=0}^{2\pi}2+4\cos\theta+2\cos^2\theta\,d\theta$

$=4\pi+0+2\pi$
$=6\pi$



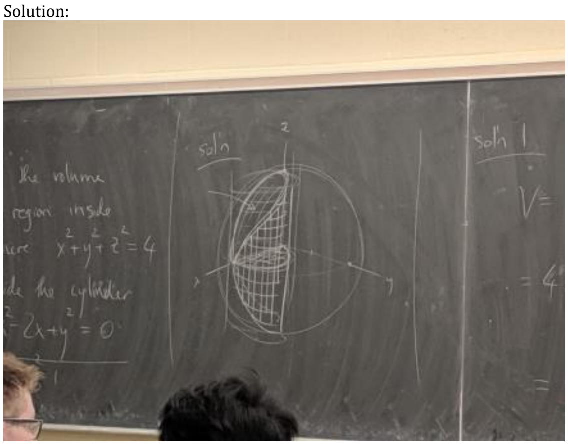


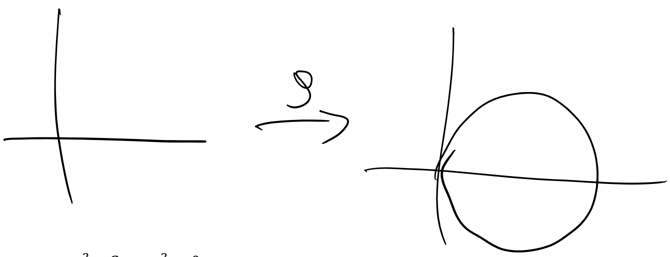
Exercise 7.12

Find the volume of the region inside the sphere $x^2+y^2+z^2=4$ and inside the cylinder $x^2-2x+y^2=0$

$D=\{(x,y,z)|x^2+y^2+z^2\le 4,(x-1)^2+y^2\le 1\}$.

Solution:





$x^2-2x+y^2=0$

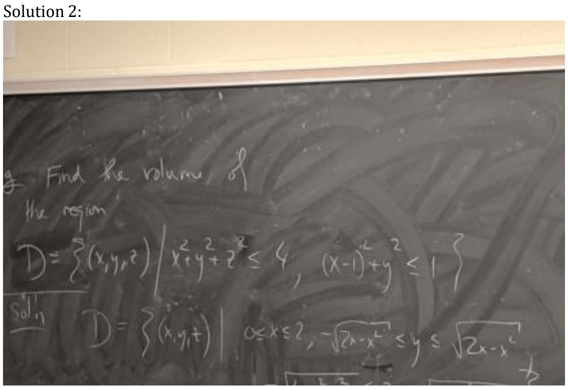
$r^2\cos^2\theta-2r\cos\theta+r^2\sin^2\theta=0$

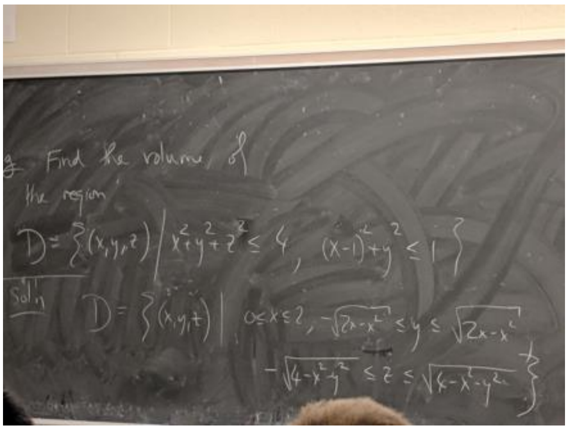
$r=0$ (not needed) or $r=2\cos\theta$

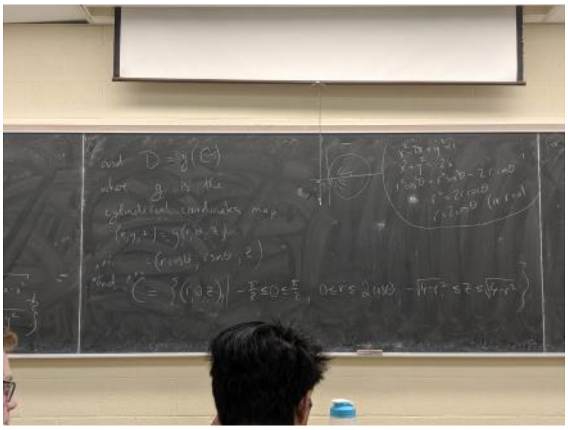
Solution 1:

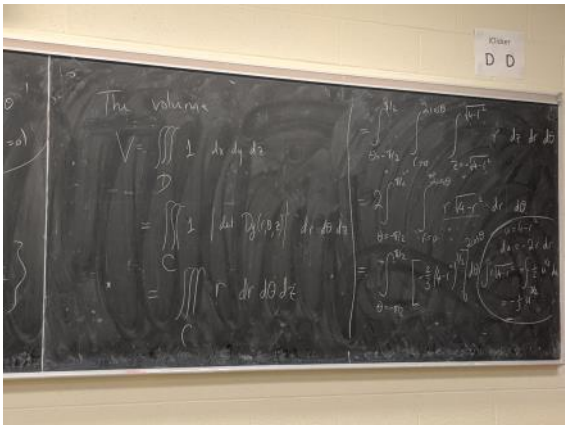
$$V=\int_D 1=4\int_{x=0}^{2}\int_{y=0}^{\sqrt{2x-x^2}}\int_{z=0}^{\sqrt{4-x^2-y^2}}1\,dz\,dy\,dx$$

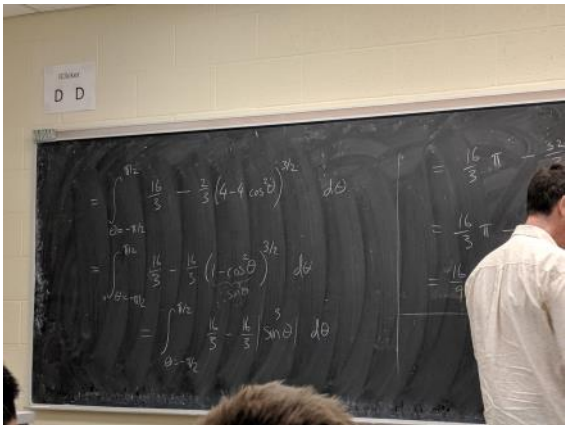
Solution 2:

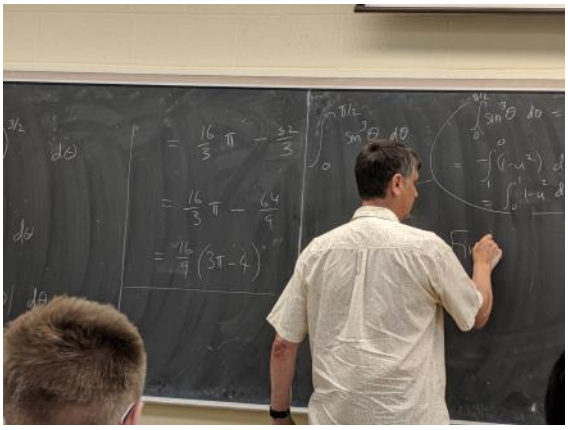


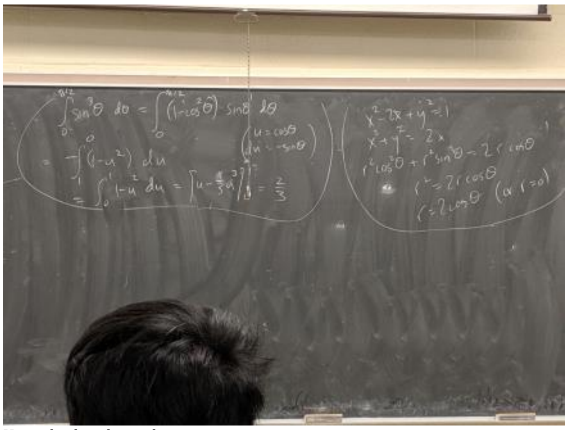

Find the volume of the region
D = {(x,y,z) | x²+y²+z² ≤ 4, (x-1)²+y² ≤ 1}
Sol'n
D = {(x,y,z) | 0≤x≤2, -√(2x-x²) ≤ y ≤ √(2x-x²)
-√(4-x²-y²) ≤ z ≤ √(4-x²-y²)}




The volume





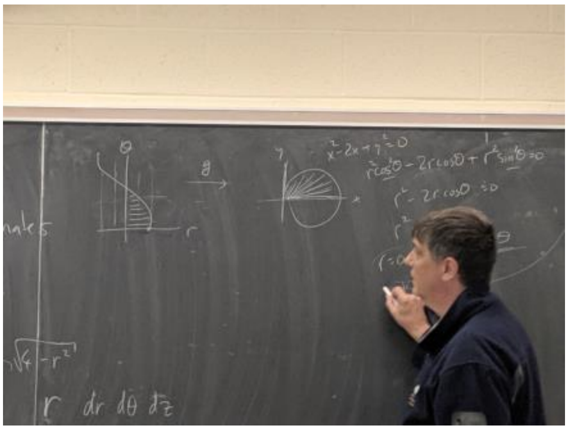


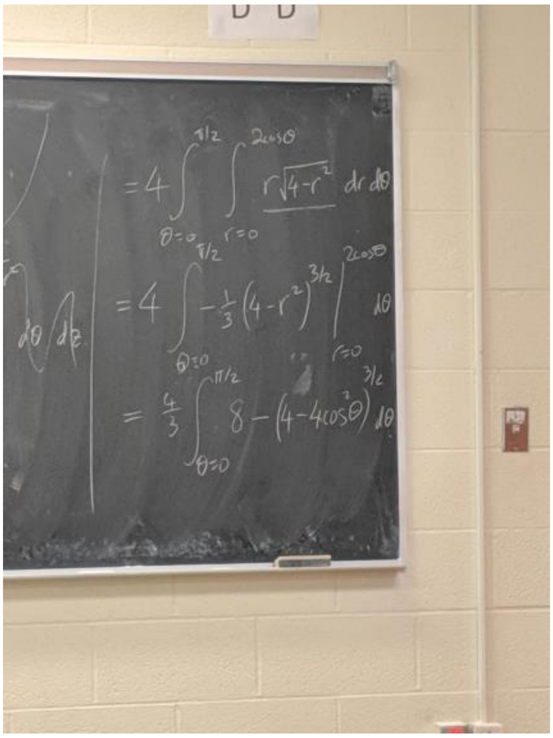
Use cylindrical coordinates

$$V = \iiint_C 1|\det Dg(r,\theta,z)|\ dr\, d\theta\, dz$$

$$V = \iiint_C r\, dr\, d\theta\, dz$$

$$= 4\int_{\theta=0}^{\frac{\pi}{2}} \int_{r=0}^{2\cos\theta} \int_{z=0}^{\sqrt{4-r^2}} r\, dz\, dr\, d\theta$$





Find the mass of the ball $x^2 + y^2 + z^2 \le 4$ where the density is given by $\rho(x, y, z) = 1 - \frac{1}{2}\sqrt{x^2 + y^2 + z^2}$

Solution:

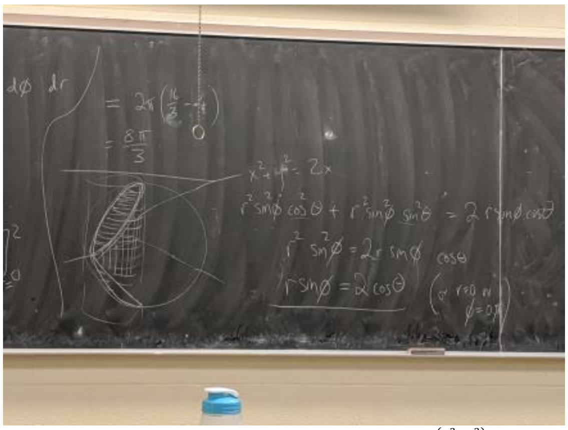
For $D = \{(x, y, z) | x^2 + y^2 + z^2 \le 4\}$ and g, the spherical coordinates map, $(x, y, z) = g(r, \phi, \theta) = (r \sin\phi \cos\theta, r \sin\phi \sin\theta, r \cos\phi)$

We have $D = g(C)$ where

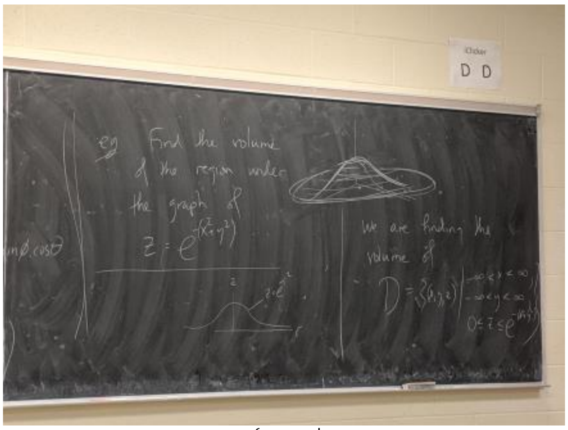
$C = \{(r, \phi, \theta) | r \le 2, 0 \le \phi \le \pi, 0 \le \theta \le 2\pi\}$

The mass is

$M = \iiint_D \left(1 - \frac{1}{2}\sqrt{x^2 + y^2 + z^2}\right) dx\, dy\, dz$
$= \iiint_C \left(1 - \frac{1}{2}r\right) |\det Dg(r, \phi, \theta)|\, dr\, d\phi\, d\theta$
$= \int_{r=0}^{2} \int_{\phi=0}^{\pi} \int_{\theta=0}^{2\pi} \left(1 - \frac{1}{2}r\right) r^2 \sin\phi\, d\theta\, d\phi\, dr$
$= \pi \int_{r=0}^{2} \int_{\phi=0}^{\pi} (2r^2 - r^3) \sin\phi\, d\phi\, dr$
$= 2\pi \int_{r=0}^{2} 2r^2 - r^3\, dr$
$= 2\pi \left[\frac{2}{3}r^3 - \frac{1}{4}r^4\right]_0^2$
$= 2\pi \left(\frac{16}{3} - \frac{12}{3}\right)$
$= \frac{8\pi}{3}$



eg. Find the volume of the region under the graph of $z = e^{-(x^2+y^2)}$



We are finding the volume of $D = \left\{(x,y,z)\middle|-\infty < x < \infty, -\infty < y < \infty, 0 \le z \le e^{-(x^2+y^2)}\right\}$

For the polar coordinates map g given

$$(x,y) = g(r,\theta) = (r\cos\theta\,, r\sin\theta)$$

We have $D = g(C)$

Where $C = \left\{(r,\theta,z)\middle|0 \le r, 0 \le \theta \le 2\pi, 0 \le z \le e^{-r^2}\right\}$

The volume is

$$\iiint_D 1\,dx\,dy\,dz = \iiint_C 1\,r\ \,dr\,d\theta\,dz$$

First one

$$\int_{x=-\infty}^{\infty}\int_{y=-\infty}^{\infty}\int_{z=0}^{e^{-(x^2+y^2)}} 1\,dz\,dy\,dx$$

On the other hand,

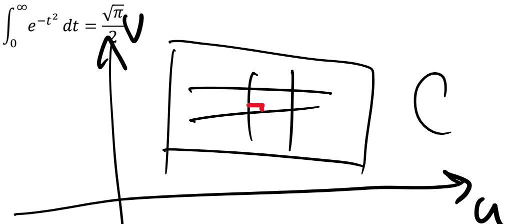
$$\iiint_C r\,dr\,d\theta\,dz$$
$$= \int_{r=0}^{\infty}\int_{\theta=0}^{2\pi}\int_{z=0}^{e^{-r^2}} r\ dz\,d\theta\,dr$$
$$= 2\pi\int_{r=0}^{\infty} re^{-r^2}\,dr$$
$$= 2\pi\left[-\frac{1}{2}e^{-r^2}\right]_{r=0}^{\infty} = 2\pi\left(0+\frac{1}{2}\right) = \pi$$

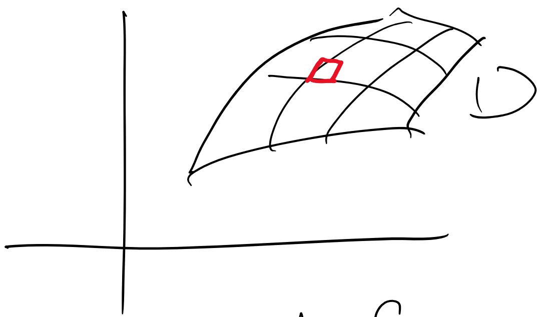
It follows that

$$4\left(\int_0^{\infty} e^{-t^2}\,dt\right)^2 = \pi$$

SO

$$\int_0^{\infty} e^{-t^2}\,dt = \frac{\sqrt{\pi}}{2}$$



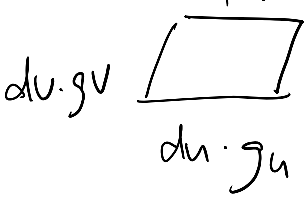


$dA = du\,dv$

$\downarrow f$

$\mathbb{R}$

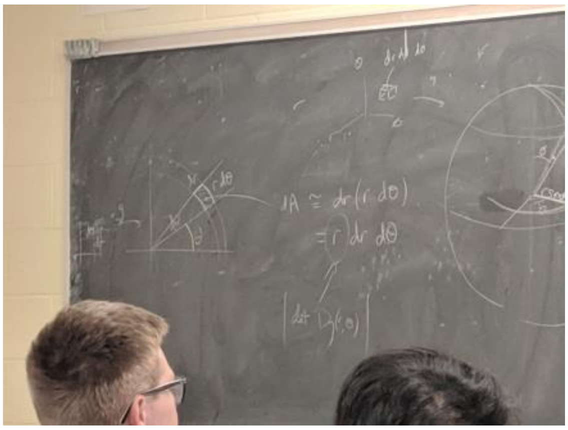
$$Dg = \begin{pmatrix} x_u & x_v \\ y_u & y_v \end{pmatrix} = \begin{pmatrix} g_u & g_v \end{pmatrix}$$

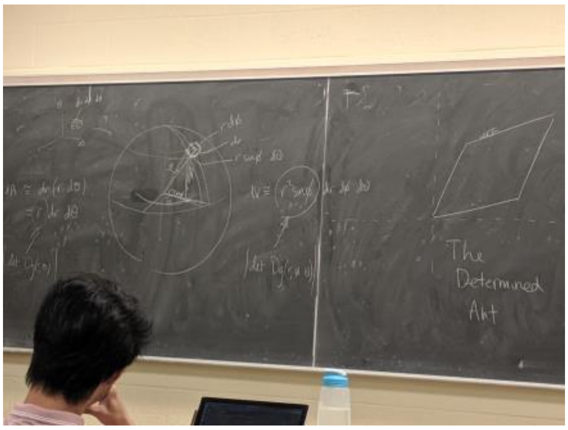


$$dA = \left|(du\cdot g_u)\times(dv\cdot g_v)\right|$$
$$= du\,dv\cdot\left|g_u\times g_v\right|$$

$= du\,dv \cdot |g_u \times g_v|$

$= |\det Dg|\, du\, dv$





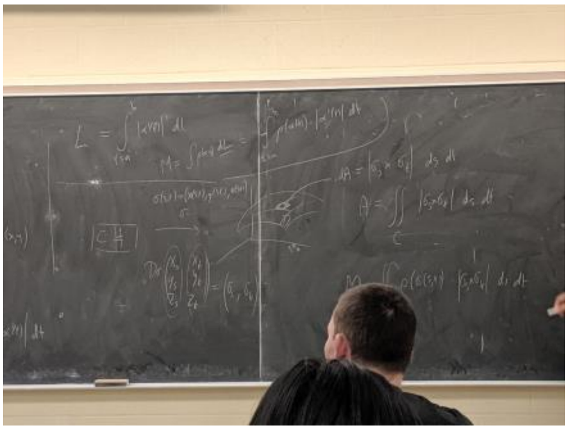
Arc Length

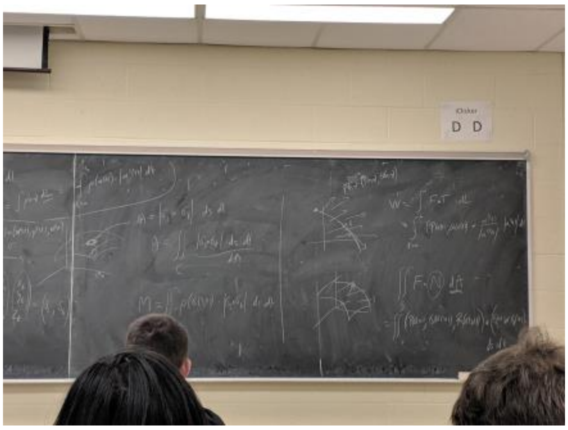
$\alpha(t) = \big(x(t), y(t), (z(t))\big)$

$\rho(x, y) =$ linear density at (x, y)

$dL = \sqrt{dx^2 + dy^2} = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt \approx \sqrt{x'(t)^2 + y'(t)^2}\, dt = |\alpha'(t)| dt$

$L = \int_{t=a}^{b} |\alpha'(t)|\, dt$





Thm (The Inverse Function Theorem)

Let $f: U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ and let $a \in U$
Suppose that f is $\mathcal{C}^1$ and $Df(a)$ is invertible
The f is locally invertible and its (local) inverse is $\mathcal{C}^1$

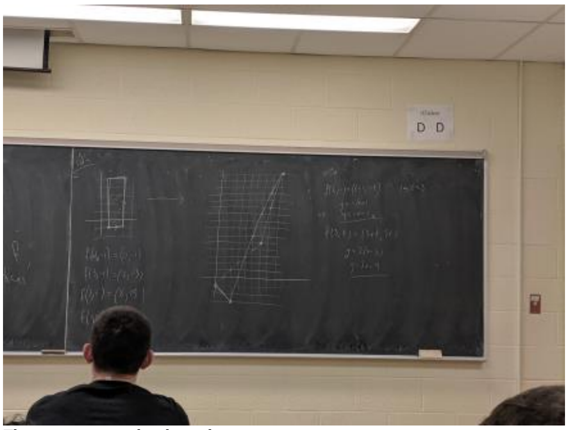
If $g = f^{-1}$, then $g(f(x)) = x$
$Dg(f(x))Df(x) = I$
$Dg(f(x)) = Df(x)^{-1}$

eg. Let $f(x,y) = (x+y, xy)$
Find the image, under f, of the rectangle with vertices $(1,-1), (3,-1), (3,5), (1,5)$

Solution:

$f(1,-1) = (0,-1)$
$f(3,-1) = (2,-3)$
$f(3,5) = (8,15)$
$f(1,5) = (6,5)$

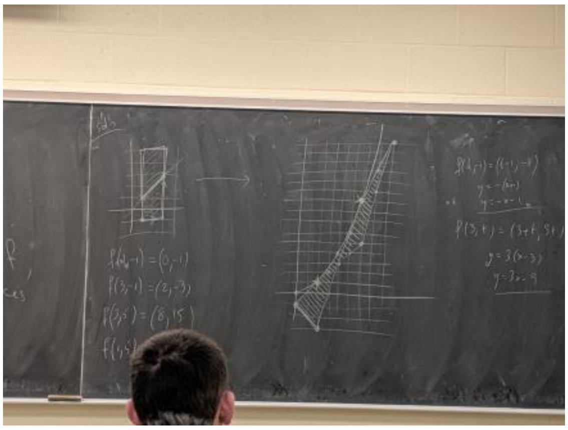
Locally invertible most of the time, but not anytime.


The image is merely a boundary

We have
$Df = \begin{pmatrix} 1 & 1 \\ y & x \end{pmatrix}$
$\det Df = x - y$

$\det Df = 0 \Leftrightarrow y = x$

$f(t,t) = (2t, t^2)$
$y = \left(\frac{1}{2}x^2\right) = \frac{1}{4}x^2$



Corollary (Parametric function Theorem)

Corollary (Implicit Function Theorem)

Let $f: U \subseteq \mathbb{R}^{n+k} \to \mathbb{R}^k$ with $p \in U$ with $f(p) = c$.

Suppose f is C^1 in U and $Df(p)$ has rank k.
Then,

Suppose f is C^1 in U and $Df(p)$ has rank k. Then the level set $f^{-1}(c)$ is locally equal to the graph of a C^1 function.

Df is a $k \times (n+k)$ matrix. Since $Df(p)$ has rank k. It follows that some $k \times k$ submatrix of $Df(p)$ is invertible.

Reorder the variables in $\mathbb{R}^{n+k}$ so that the last k columns are independent.

Write $x = (x_1, \dots, x_n), y = (y_1, \dots, y_k)$ and $z = f(x, y)$
Then $Df = \begin{pmatrix} \frac{\partial z}{\partial x} & \frac{\partial z}{\partial y} \end{pmatrix}$
where $\frac{\partial z}{\partial x} = \begin{pmatrix} \frac{\partial z_1}{\partial x_1} & \cdots & \frac{\partial z_1}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \frac{\partial z_k}{\partial x_n} \end{pmatrix}$

and $\frac{\partial z}{\partial y}(p)$ is invertible.

Define $F: U \subseteq \mathbb{R}^{n+k} \to \mathbb{R}^{n+k}$
by $F(x, y) = (w, z) = (x, f(x, y))$

Then $DF = \begin{pmatrix} I & 0 \\ \frac{\partial z}{\partial x} & \frac{\partial z}{\partial y} \end{pmatrix}$
so $DF(p)$ is invertible.
By the **inverse function theorem,** F is locally invertible.

Let G be the inverse map, and write
$(x, y) = G(w, z) = (w, g(w, z))$
Then verify that the level set $f^{-1}(c) = \{(x, y) | f(x, y) = C\}$ is the same of the graph of the function $y = h(x)$, where $h(x) = g(x, c)$.

eg. Sketch $x^2 - 2xy + 4y^2 = 12$
Solution:

$x^2 - 2xy + 4(y^2 - 3) = 0$
$\Leftrightarrow x = \dfrac{2y \pm \sqrt{4y^2 - 16(y^2 - 3)}}{2}$
$x = y \pm \sqrt{12 - 3y^2}$

Solution:
$x^2 - 2xy + 4y^2 = 12$
$\Leftrightarrow (x, y) \begin{pmatrix} 1 & -1 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 12$
diagonalize $A = \begin{pmatrix} 1 & -1 \\ -1 & 4 \end{pmatrix}$
A is positive definite

$f_A(x) = (x-1)(x-4) - 1$
$= x^2 - 5x + 3$
$\lambda = \dfrac{(5 \pm \sqrt{13})}{2}$

Solution:

Let $f(x, y) = x^2 - 2xy + 4y^2$
(we are sketching $f^{-1}(12)$)
$Df = (2x - 2y, -2x + 8y)$

For $(x, y) \in f^{-1}(12)$

$\frac{\partial f}{\partial x} = 0 \Rightarrow x = y$
$\Rightarrow 3x^2 = 12$
$x = \pm 2$
$(x, y) = (2,2), (-2, -2)$

$\frac{\partial f}{\partial y} = 0 \Rightarrow x = 4y \Rightarrow 16y^2 - 8y^2 + 4y^2 = 12$
$12y^2 = 12$
$y = \pm 1$

$\Rightarrow (x, y) = (4,1), (-4,-1)$

Inverse function theorem

Differential geometry

Lagrange Multipliers

To define the n-volume of a bounded set $A \subseteq \mathbb{R}^n$

Choose a rectangular box which contains A.
Say $R = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$
Partition each interval $[a_j, b_j]$ to get a partition P of R into subrectangles $R_{i_1, i_2, \ldots, i_n}$
We define upper and lower volume estimates

$U = \sum_{k \in T} |R_k|$
$L = \sum_{k \in S} |R_k|$
where

where $S = \{k = (k_1, \ldots, k_n) | R_k \subseteq A\}$ or $R_k \subseteq A^o$

$T = \{k \in (k_1, \ldots, k_n) | R_k \cap A \neq \emptyset\}$ or $R_k \cap \bar{A} \neq \emptyset$

We define the outer and inner volumes of A in R to be
$\overline{\text{Vol}}(A, R) = \inf\{U(R, P)(A) | \text{all partition } P \text{ of } R\}$
Vol\underbar $(A, R) = \sup\{L(R, P)(A) | \text{all partition } P \text{ of } R\}$

Verify that the two Vols do not depend on R.

Write $C^* = \text{Vol\bar}(A, R)$ for any rectangle R with $A \subseteq R$.
and $C_*(A) = \text{Vol\underbar}(A, R)$ for any R

$C^*(A)$ is called the outer Jordan content.

$C_*(A)$ is the inner Jordan content of A

We say A has a well-defined area or well-defined Jordan content
when $C^*(A) = C_*(A)$

Why not just upper limit?
Break some properties.
Volumes/Areas not properly added up

Fact
If A and B are bounded and have well-defined volume, then $C^*(A \cup B) = C^*(A) + C^*(B)$

Fact
A bounded set $A \subseteq \mathbb{R}^n$ has a well-defined Jordan content $\Leftrightarrow C^*(\partial A) = 0$

When $A \subseteq R$, a rectangle, $C^*(A) + C_*(R \setminus A) = |R|$

$R_k \subseteq A \Leftrightarrow R_k \cap A^c = \emptyset$

Term Test 4 Preparation

2019年7月27日 4:44

NOTE: The MATH 247 Term Test 4 will be held on
Monday July 29, from 12:30-1:20 in MC 4063.

It will cover Chapter 6 and Sections 7.1-7.13, and Assignment 第 6.

You will be asked to prove 1 of the following 3 theorems.
Lemma 6.1 (Iterated Limits)
Theorem 6.8 (Taylor's Theorem)
Theorem 6.17 (The Second Derivative Test)

来自 <http://www.math.uwaterloo.ca/~snew/math247-2019-S/>