Selected Recitation Notes for Multivariable Calculus

Notes by Zhengning Hu
zhy8b@umsystem.edu

1 January 17 **2**

2 January 18 **6**

3 January 19 **8**

4 January 20 **11**

5 January 25 **15**

6 January 27 **18**

7 February 1 **20**

8 February 3 **22**

9 February 8 **25**

10 February 10 **28**

11 February 17 **31**

12 February 22 **34**

13 February 24 **37**

14 March 8 **41**

15 March 10 **44**

16 March 17 **48**

17 March 22 & 23 **51**

18 April 5 **56**

19 April 6 **59**

20 April 12 **61**

21 April 14 **63**

22 April 21 **67**

23 April 26 **70**

24 April 28 **75**

25 May 3 **78**

This is the recitation note I've prepared for MATH 4900 - Multivariable Calculus in University of Missouri – Columbia, Spring 2023, taught by Professor Carlo Morpurgo.

1 January 17

In this first class, we will introduce some notations and basic objects in multivariable calculus. The main objects we will consider throughout this semester are

- n-dimensional Euclidean space $\mathbb{R}^n$, and
- Real-valued functions defined on subset of $\mathbb{R}^n$.

Notations The Euclidean space $\mathbb{R}^n$ (usually $n > 1$) consists of

$$\mathbb{R}^n = \{\boldsymbol{x} = (x_1, \cdots, x_n) \mid x_i \in \mathbb{R}\}$$

where the element $\boldsymbol{x} = (x_1, \cdots, x_n)$ is called a **vector** or a **point** in $\mathbb{R}^n$ and $x_i \in \mathbb{R}$ is called the i-th coordinate of $\boldsymbol{x}$. We can also write

$$\boldsymbol{x} = \sum_{i=1}^{n} x_i e_i$$

where $\{e_1, \cdots, e_n\}$ is called the standard orthogonal basis for $\mathbb{R}^n$ and

$$e_i = (0, \cdots, 1, \cdots, 0)$$

with a single 1 on the ith coordinate. In particular, elements in $\mathbb{R}$ are referred to as **scalars**.

Analogous to $\mathbb{R}$, there are also algebraic operations on $\mathbb{R}^n$. The first thing we can do is to extend addition on $\mathbb{R}$ to $\mathbb{R}^n$ componentwisely: suppose $\boldsymbol{x} = (x_1, \cdots, x_n), \boldsymbol{y} = (y_1, \cdots, y_n) \in \mathbb{R}^n$,

$$\boldsymbol{x} + \boldsymbol{y} = (x_1 + y_1, \cdots, x_n + y_n) \in \mathbb{R}^n.$$

We call $\underline{0} = (0, \cdots, 0) \in \mathbb{R}^n$ the zero vector in $\mathbb{R}^n$.

For multiplication, first notice that we can define the scalar multiplication. For $a \in \mathbb{R}$ and $\boldsymbol{x} = (x_1, \cdots, x_n) \in \mathbb{R}^n$, define

$$a\boldsymbol{x} = (ax_1, \cdots, ax_n) \in \mathbb{R}^n$$

which works like stretching or contracting a vector by the factor of $a \in \mathbb{R}$.

With this in mind, it is natural to define the notion of lines and line segments.

- Line passing through $\mathbf{0}$ with direction $\boldsymbol{b} \in \mathbb{R} \setminus \mathbf{0}$ is the set $\{t\boldsymbol{b} \mid t \in \mathbb{R}\}$;
- Line passing though $\boldsymbol{a}$ with direction $\boldsymbol{b} \in \mathbb{R} \setminus \mathbf{0}$ is the set $\{\boldsymbol{a} + t\boldsymbol{b} \mid t \in \mathbb{R}\}$;
- Line segment from $\boldsymbol{a}$ to $\boldsymbol{b}$ is $L(\boldsymbol{a}, \boldsymbol{b}) = \{\boldsymbol{a} + t(\boldsymbol{b} - \boldsymbol{a}) \mid t \in [0,1]\}$ with starting point $\boldsymbol{a}$ corresponding to the value of parameter $t = 0$ and end point $\boldsymbol{b}$ associated to the value of parameter $t = 1$.

Different from $\mathbb{R}$, there are infinitely many ways to link from a point $\boldsymbol{a}$ to another $\boldsymbol{b}$ other than the line segment $L(\boldsymbol{a}, \boldsymbol{b})$. Take n continuous functions $x_i : [0,1] \to \mathbb{R}$. A path Γ in $\mathbb{R}^n$ is defined by

$$\Gamma = \{(x_1(t), \cdots, x_n(t)) \mid t \in [0,1]\}$$

parametrized by $t \in [0,1]$.

Moreover, using the scalar multiplication, we can also defined the parallelogram. The parallelogram spanned by $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$ is defined by

$$P(\boldsymbol{a}, \boldsymbol{b}) = \{\lambda\boldsymbol{a} + \mu\boldsymbol{b} \mid \lambda, \mu \in [0,1]\}.$$

We can also define the **inner/dot product** of two vectors in $\mathbb{R}^n$. Use the above notation, define

$$\boldsymbol{x} \cdot \boldsymbol{y} = x_1y_1 + \cdots + x_ny_n = \sum_{i=1}^{n} x_iy_i \in \mathbb{R}.$$

It is actually a function $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ given by $(\boldsymbol{x}, \boldsymbol{y}) \mapsto \boldsymbol{x} \cdot \boldsymbol{y}$ which associates each pair of vectors a real number. The inner product satisfies the following properties:

- (Positivity) $\boldsymbol{x} \cdot \boldsymbol{x} \geq 0$ with equation holds if and only if $\boldsymbol{x} = \mathbf{0}$;
- (Symmetry) $\boldsymbol{x} \cdot \boldsymbol{y} = \boldsymbol{y} \cdot \boldsymbol{x}$;
- (Bilinearity) $(\alpha\boldsymbol{x} + \beta\boldsymbol{y}) \cdot \boldsymbol{z} = \alpha\boldsymbol{x} \cdot \boldsymbol{z} + \beta\boldsymbol{y} \cdot \boldsymbol{z}$. Note that this is only the bilinearity on the first component, but by symmetry we get the bilinearity on the second component.

In particular, we have $\mathbf{0} \cdot \boldsymbol{x} = \mathbf{0}$, $\alpha(\boldsymbol{x} \cdot \boldsymbol{y}) = (\alpha\boldsymbol{x}) \cdot \boldsymbol{y}$ and $(\boldsymbol{x} + \boldsymbol{y}) \cdot \boldsymbol{z} = \boldsymbol{x} \cdot \boldsymbol{z} + \boldsymbol{y} \cdot \boldsymbol{z}$.

Remark 1.1. You may want to ask if there is a reasonable division on $\mathbb{R}^n$ or in other words it is possible to define the reciprocal of a nonzero vector. In general the answer is no. But for special n, for example $n = 2$ we get $\mathbb{R}^2 \cong \mathbb{C}$ we can define division. Other examples with well-defined division like $n = 4$ which is the quaternion $\mathbb{H}$ (loses commutativity for multiplication) and $n = 8$ called octonion (loses both commutativity and associativity for multiplication). This is due to Frobenius.

Just as absolute value gives a way to measure the distance to the origin, there are multiple ways to define **norm** on $\mathbb{R}^n$. First of all, we have the nonnegative inner product

$$\boldsymbol{x} \cdot \boldsymbol{x} = \sum x_i^2 = \sum |x_i|^2$$

which gives the definition of **Euclidean norm** of a vector $\boldsymbol{x}$ in $\mathbb{R}^n$:

$$||\boldsymbol{x}|| = \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2} = \sqrt{\boldsymbol{x} \cdot \boldsymbol{x}}.$$

It is also called l^2-norm. For $p \geq 1$ there are more general norms, called l^p-norm:

$$||\boldsymbol{x}||_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}.$$

For example, the l^1-norm means

$$||\boldsymbol{x}||_1 = \sum_{i=1}^{n} |x_i|$$

and the sup norm or l^∞-norm means

$$||\boldsymbol{x}||_\infty = \max_{1 \leq i \leq n} \{|x_i|\}.$$

Any function $|| \cdot || : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ satisfying the following properties is called a norm on $\mathbb{R}^n$:

- $||\boldsymbol{x}|| \geq 0$ with equality holds if and only if $\boldsymbol{x} = \boldsymbol{0}$;
- $||\lambda \boldsymbol{x}|| = |\lambda| ||\boldsymbol{x}||$ for any scalar $\lambda \in \mathbb{R}$;
- (Triangle inequality) $||\boldsymbol{x} + \boldsymbol{y}|| \leq ||\boldsymbol{x}|| + ||\boldsymbol{y}||$.

One can check that the l^p-norm ($p \geq 1$) defined above are indeed norms, meaning that they all satisfy the above three conditions. In particular, the trick for proving the triangle inequality for the Euclidean norm is the Cauchy-Schwartz inequality which we will prove on Thursday.

Cauchy-Schwartz inequality For any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, with $|| \cdot ||$ the Euclidean norm on $\mathbb{R}^n$ we have

$$|\boldsymbol{x} \cdot \boldsymbol{y}| \leq ||\boldsymbol{x}|| ||\boldsymbol{y}||$$

with equality holds if and only if $\boldsymbol{x} = \boldsymbol{0}$ or $\boldsymbol{y} = 0$ or $\boldsymbol{x} = \lambda \boldsymbol{y}$ for some $\lambda \in \mathbb{R}$. We provide two different proofs for this result.

Proof. The key point is

$$||\boldsymbol{a} + \boldsymbol{b}||^2 = (\boldsymbol{a} + \boldsymbol{b}) \cdot (\boldsymbol{a} + \boldsymbol{b}) = ||\boldsymbol{a}||^2 + 2\boldsymbol{a} \cdot \boldsymbol{b} + ||\boldsymbol{b}||^2.$$

Thus for any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, and $t \in \mathbb{R}$

$$0 \leq ||\boldsymbol{x} - t\boldsymbol{y}||^2 = ||\boldsymbol{x}||^2 - 2t\boldsymbol{x} \cdot \boldsymbol{y} + t^2||\boldsymbol{y}||^2.$$

The inequality holds obviously if $\boldsymbol{y} = \mathbf{0}$. Now we assume $\boldsymbol{y} \neq \mathbf{0}$. Then the RHS of the above inequality is a quadratic form in terms of t with positive leading coefficient. Thus we denote

$$\phi(t) = ||\boldsymbol{y}||^2 t^2 - 2t\boldsymbol{x} \cdot \boldsymbol{y} + ||\boldsymbol{x}||^2.$$

Since $\phi(t) \geq 0$ it implies the discriminant

$$\Delta = (2\boldsymbol{x} \cdot \boldsymbol{y})^2 - 4||\boldsymbol{y}||^2||\boldsymbol{x}||^2 \leq 0$$

and thus $|\boldsymbol{x} \cdot \boldsymbol{y}| \leq ||\boldsymbol{x}||||\boldsymbol{y}||$. In particular, if $|\boldsymbol{x} \cdot \boldsymbol{y}| = ||\boldsymbol{x}||||\boldsymbol{y}||$ meaning $\Delta = 0$ thus the quadratic form $\phi(t)$ admits a double root $t^* = \frac{\boldsymbol{x} \cdot \boldsymbol{y}}{||\boldsymbol{y}||^2}$. Therefore $||\boldsymbol{x} - t^*\boldsymbol{y}||^2 = 0$ if and only if $\boldsymbol{x} = t^*\boldsymbol{y}$. Conversely if $\boldsymbol{x} = t\boldsymbol{y}$ then $|\boldsymbol{x} \cdot \boldsymbol{y}| = ||\boldsymbol{x}||||\boldsymbol{y}||$ as desired. □

Here is an alternate proof:

Proof. The inequality is invariant under dilation and in particular, for nonzero $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ to prove the inequality is equivalent to showing

$$\left| \frac{\boldsymbol{x}}{||\boldsymbol{x}||} \cdot \frac{\boldsymbol{y}}{||\boldsymbol{y}||} \right| \leq 1$$

where $\frac{\boldsymbol{x}}{||\boldsymbol{x}||} = 1$ is a unit vector. Thus we can reduce to show the inequality for unit vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$. It is clear that

$$0 \leq ||\boldsymbol{x} - \boldsymbol{y}||^2 = ||\boldsymbol{x}||^2 - 2\boldsymbol{x} \cdot \boldsymbol{y} + ||\boldsymbol{y}||^2 = 2(1 - \boldsymbol{x} \cdot \boldsymbol{y})$$

which implies $\boldsymbol{x} \cdot \boldsymbol{y} \leq 1$. Likewise if we consider $||\boldsymbol{x} + \boldsymbol{y}||^2 \geq 0$ we get $\boldsymbol{x} \cdot \boldsymbol{y} \geq -1$. Therefore $|\boldsymbol{x} \cdot \boldsymbol{y}| \leq 1$. □

Now the triangle inequality can be easily proved.

Proof. In order to show $||\boldsymbol{x} + \boldsymbol{y}|| \leq ||\boldsymbol{x}|| + ||\boldsymbol{y}||$ we will prove

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

Then

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

shows the triangle inequality of Euclidean norm. □

In general the triangle inequality for l^p-norm is essentially the Minkowski inequality.

2 January 18

Before talking about norms, let's first take a look at inner products on $\mathbb{R}^n$ since the most commonly used Euclidean norm does come from the inner product we defined above. We have defined the inner product of two vectors before. In general, an inner product is a function

$$\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$$

satisfying positivity, symmetry and bilinearity.

Let's look at some examples on the real plane $\mathbb{R}^2$. Let $\boldsymbol{x} = (x_1, x_2)$ and $\boldsymbol{y} = (y_1, y_2)$:

Example 2.1.

- Consider $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 2x_1y_1 + 3x_2y_2$. One can check this forms an inner product on $\mathbb{R}^2$.
- Consider $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 2x_1 + 3y_2$. This is not an inner product because

$$\langle (1,0), (2,0) \rangle = 2 \neq 4 = \langle (2,0), (1,0) \rangle.$$

- Consider $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = x_1y_2 + x_2y_1$. This is not an inner product on $\mathbb{R}^2$ because if we take $\boldsymbol{x} = \boldsymbol{y} = (1,0)$, we have

$$\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0$$

but $\boldsymbol{x}$ is not the zero vector $\mathbf{0}$ in $\mathbb{R}^2$.

Remark 2.2. Every inner product induces a norm on $\mathbb{R}^n$. But note that not every norm comes from some inner product and there are inner products that do not come from norms. A norm is induced by an inner product if and only if it satisfies the parallelogram law: $||\boldsymbol{x} + \boldsymbol{y}||^2 + ||\boldsymbol{x} - \boldsymbol{y}||^2 = 2(||\boldsymbol{x}||^2 + ||\boldsymbol{y}||^2)$.

Indeed one can check that the norm induced by some inner product does satisfy the parallelogram law:

$$\begin{aligned} LHS &= \langle \boldsymbol{x} + \boldsymbol{y}, \boldsymbol{x} + \boldsymbol{y} \rangle + \langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y} \rangle \\ &= 2\langle \boldsymbol{x}, \boldsymbol{x} \rangle + 2\langle \boldsymbol{y}, \boldsymbol{y} \rangle \\ &= 2(||\boldsymbol{x}||^2 + ||\boldsymbol{y}||^2) = RHS. \end{aligned}$$

Now let's focus on l^p-norms first for $1 \leq p \leq \infty$. Recall that for a vector $\boldsymbol{x} \in \mathbb{R}^n$ the l^p-norm of $\boldsymbol{x} = (x_1, \cdots, x_n)$ is defined by

$$||\boldsymbol{x}||_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}$$

with the sup norm or l^∞-norm given by

$$||\boldsymbol{x}||_\infty = \max_{1 \leq i \leq n} \{|x_i|\}.$$

These are related by the following where we will only prove the case for $\mathbb{R}^2$ and you will see the proof for the general case soon.

Proposition 2.3. *For any* $\boldsymbol{x} = (x_1, x_2) \in \mathbb{R}^2$,

$$\lim_{p\to\infty} ||\boldsymbol{x}||_p = ||\boldsymbol{x}||_\infty.$$

Proof. Without loss of generality we may assume $|x_1| \geq |x_2|$. Then

$$||\boldsymbol{x}||_p = (|x_1|^p + |x_2|^p)^{1/p} = |x_1| \left(1 + \frac{|x_2|^p}{|x_1|^p}\right)^{1/p} \to |x_1|. \qquad \square$$

A norm $||\cdot||$ on $\mathbb{R}^n$ defines a metric d (or distance) on $\mathbb{R}^n$ by setting $d(\boldsymbol{x}, \boldsymbol{y}) = ||\boldsymbol{x} - \boldsymbol{y}||$. This is not only for $\mathbb{R}^n$ but works for all k-vector space where $(k, |\cdot|)$ is a valued field, *i.e.,* a field with valuation. Once a metric is defined we get open subsets and closed sets of $\mathbb{R}^n$ with respect to the norm $||\cdot||$. Though different norms provide different metric but they may determine the "same" open sets of $\mathbb{R}^n$ thus the same topology on $\mathbb{R}^n$.

Therefore it is essential to figure out the underlying relation among all the norms we have, in terms of whether they define the same topology on $\mathbb{R}^n$ (although we didn't get to that far but we will cover some necessary topology facts later).

Definition 2.4. Consider the space $\mathbb{R}^n$. Two norms $||\cdot||_{(1)}$ and $||\cdot||_{(2)}$ are **equivalent** if there are positive constants $A, B \in \mathbb{R}$ such that for any $\boldsymbol{x} \in \mathbb{R}^n$,

$$A||\boldsymbol{x}||_{(2)} \leq ||\boldsymbol{x}||_{(1)} \leq B||\boldsymbol{x}||_{(2)}.$$

The following proposition shows that all the l^p-norms on $\mathbb{R}^n$ are equivalent, and thus they induce the same topology on $\mathbb{R}^n$.

Proposition 2.5. *For any* $\boldsymbol{x} \in \mathbb{R}^n$,

(1) $||\boldsymbol{x}||_\infty \leq ||\boldsymbol{x}||_p \leq \sqrt[p]{n}||\boldsymbol{x}||_\infty$.

(2) $||\boldsymbol{x}|| \leq ||\boldsymbol{x}||_1 \leq \sqrt{n}||\boldsymbol{x}||$.

(3) In general, for any p, q *with* $1 \leq p \leq q \leq \infty$, *we have*

$$||\boldsymbol{x}||_q \leq ||\boldsymbol{x}||_p \leq n^r ||\boldsymbol{x}||_q$$

where

$$r = \begin{cases} \frac{1}{p} & \text{if } p < q = \infty, \\ \frac{1}{p} - \frac{1}{q} & \text{if } p \leq q < \infty, \\ 0 & \text{if } p = q = \infty. \end{cases}$$

Proof. Part (1) can be proved by the following easy inequality

$$\max\{|x_i|\} \leq \sqrt[p]{|x_1|^p + \cdots + |x_n|^p} \leq \sqrt[p]{n \max\{|x_i|\}^p}.$$

For (2), first notice that

$$||\boldsymbol{x}||_1 = \sum_{i=1}^{n} |x_i| = \sum_{i=1}^{n} |x_i| \cdot 1 \le \left(\sum_{i=1}^{n} |x_i|^2 \right)^{1/2} \left(\sum_{i=1}^{n} 1^2 \right)^{1/2} \le \sqrt{n}||\boldsymbol{x}||.$$

For $||\boldsymbol{x}|| \le ||\boldsymbol{x}||_1$, it is easy to see that

$$||\boldsymbol{x}||_1^2 = \left(\sum_{i=1}^{n} |x_i| \right)^2 \ge \sum_{i=1}^{n} |x_i|^2 = ||\boldsymbol{x}||^2.$$

To prove the general case for $1 \le p \le q \le \infty$, it remains to show the case for $1 \le p \le q < \infty$. To prove the first inequality, note that we can reduce to the case $||\boldsymbol{x}||_p = 1$. Then $|x_i| \le 1$ implies that $|x_i|^q \le |x_i|^p$ since $p \le q$. The second inequality is just the power mean inequality: for $p < q < \infty$

$$\left(\frac{|x_1|^p + \cdots + |x_n|^p}{n} \right)^{1/p} \le \left(\frac{|x_1|^q + \cdots + |x_n|^q}{n} \right)^{1/q}$$

which can be shown by consider the concave function $f(x) = x^{p/q}$ and apply the Jensen's inequality. More precisely, we get from Jensen's that

$$\left(\frac{|x_1|^q + \cdots + |x_n|^q}{n} \right)^{p/q} \ge \frac{|x_1|^{q\frac{p}{q}} + \cdots + |x_n|^{q\frac{p}{q}}}{n}$$

which is exactly the inequality we want. □

The fact is that all norms on a finite-dimensional vector space over $\mathbb{R}$ (actually over a complete valued field) are equivalent. In particular, all norms on $\mathbb{R}^n$ define the same topology on $\mathbb{R}^n$.

3 January 19

This is the note after the lecture delivered on January 17.

A quick note: if $\boldsymbol{x} \cdot \boldsymbol{y} = 0$ then we have the Pythagorean Theorem:

$$||\boldsymbol{x} + \boldsymbol{y}||^2 = ||\boldsymbol{x}||^2 + ||\boldsymbol{y}||^2.$$

If $\boldsymbol{x} \cdot \boldsymbol{y} = 0$ we say $\boldsymbol{x}, \boldsymbol{y}$ are orthogonal or perpendicular; if $\boldsymbol{x} = t\boldsymbol{y}$ for some scalar $t \in \mathbb{R}$ we say $\boldsymbol{x}, \boldsymbol{y}$ are parallel. We call $\{\boldsymbol{e}_i\}$ a set of orthogonal unit vectors (forms a orthonormal system). Moreover we define the angle between $\boldsymbol{x}, \boldsymbol{y} \ne \boldsymbol{0}$ as $\theta \in [0, \pi]$ such that

$$\cos\theta = \frac{\boldsymbol{x} \cdot \boldsymbol{y}}{||\boldsymbol{x}||||\boldsymbol{y}||}.$$

In addition, we define the hyperplane through a point $\boldsymbol{a}$ perpendicular to a vector $\boldsymbol{b}$ to be

$$\Pi_{\boldsymbol{b}}(\boldsymbol{a}) = \{\boldsymbol{x} \in \mathbb{R}^n \mid (\boldsymbol{x} - \boldsymbol{a}) \cdot \boldsymbol{b} = \boldsymbol{0}\}.$$

Therefore given $\boldsymbol{b} \neq \boldsymbol{0}$ the equation $\boldsymbol{x} \cdot \boldsymbol{b} = \lambda$ for some $\lambda \in \mathbb{R}$ define a set of points $\boldsymbol{x}$ in a hyperplane perpendicular to $\boldsymbol{b}$. For instance if $b_1 \neq 0$ then

$$\lambda = \frac{\boldsymbol{e}_1 \cdot \boldsymbol{b}}{b_1} \lambda = \boldsymbol{a} \cdot \boldsymbol{b}$$

where $\boldsymbol{a} = \lambda \boldsymbol{e}_1 / b_1$.

Next let's talk about linear transformations.

Definition 3.1. A linear transformation is a function $A : \mathbb{R}^n \to \mathbb{R}^m$ such that for any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$,

(1) $A(\boldsymbol{x} + \boldsymbol{y}) = A(\boldsymbol{x}) + A(\boldsymbol{y})$;
(2) $A(\lambda \boldsymbol{x}) = \lambda A(\boldsymbol{x})$.

We will use $A\boldsymbol{x}$ as $A(\boldsymbol{x})$ and write $L(\mathbb{R}^n, \mathbb{R}^m)$ as the set of all linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$.

Linear transformation can be identified with matrices $A \in \mathbb{R}^{m \times n}$ (we will use this notation $\mathbb{R}^{m \times n}$ for the set of all real $n \times m$ matrices),

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}.$$

For any linear transformation $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ there is a unique matrix $A \in \mathbb{R}^{m \times n}$ such that

$$A\boldsymbol{x} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

under this interpretation $\boldsymbol{x}$ is represented by column vector as the RHS as above.

Moreover, if we write $A = [\boldsymbol{a}_1 | \cdots | \boldsymbol{a}_n]$ as column vectors, then $\boldsymbol{a}_k = A\boldsymbol{e}_k$. We denote by A^T the transpose of A and with this notation $\boldsymbol{x} \cdot \boldsymbol{y} = \boldsymbol{x}^T \boldsymbol{y} = [\boldsymbol{x}]^T [\boldsymbol{y}]$.

For linear transformations, we can also define norms on them.

Definition 3.2. For $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ or equivalently, $A \in \mathbb{R}^{m \times n}$, the Frobenius norm of A is

$$||A||_2 = \left(\sum_{j=1}^{n} \sum_{k=1}^{m} |a_{jk}|^2 \right)^{1/2}.$$

This is a norm in the space of matrices $\mathbb{R}^{m \times n}$.

Proposition 3.3. *If $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ then*

$$||A\boldsymbol{x}|| \le ||A||_2 ||\boldsymbol{x}||$$

for any $\boldsymbol{x} \in \mathbb{R}^n$.

Proof. Let $\boldsymbol{x} = \sum_{k=1}^n x_k \boldsymbol{e}_k$, we have

$$\begin{aligned} ||A\boldsymbol{x}|| &= ||\sum_k x_k A\boldsymbol{e}_k|| \le \sum_k |x_k| ||A\boldsymbol{e}_k|| \\ &\le \left(\sum_{k=1}^n |x_k|^2\right)^{1/2} \left(\sum_{k=1}^n ||A\boldsymbol{e}_k||\right)^2 \\ &= \left(\sum_{k=1}^n |x_k|^2\right)^{1/2} \left(\sum_{k=1}^n \sum_{j=1}^m |a_{jk}|^2\right)^{1/2}. \end{aligned}$$

□

Here is an example:

Example 3.4. Consider $A : \mathbb{R}^2 \to \mathbb{R}$ given by $A(x_1, x_2) = 3x_1 + x_2$. Then $A = [3, 1]$ with $||A||_2 = \sqrt{10}$. And

$$|3x_1 + x_2| \le 3|x_1| + |x_2| \le \sqrt{3^2 + 1}\sqrt{x_1^2 + x_2^2}$$

with equality holds for some $\boldsymbol{x}$, for instance $x_1 = 3, x_2 = 1$.

Thus we can ask: if $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ then there is a constant C such that $||A\boldsymbol{x}|| \le C||\boldsymbol{x}||$ for any $\boldsymbol{x} \in \mathbb{R}^n$, what is the smallest C? This leads to the following definition:

Definition 3.5. The operator norm of $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ is

$$||A|| = \sup_{\boldsymbol{x} \neq \boldsymbol{0}} \frac{||A\boldsymbol{x}||}{||\boldsymbol{x}||} < \infty$$

(it is bounded since $||A||_2$ is an upper bound) the smallest C such that $||A\boldsymbol{x}|| \le C||\boldsymbol{x}||$ holds and $||A\boldsymbol{x}|| \le ||A|| ||\boldsymbol{x}||$.

Note that we always have $||A|| \le ||A||_2$ but in general $||A|| < ||A||_2$.

Proposition 3.6. *If $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ then*

$$||A|| = \sup_{||\boldsymbol{x}|| \le 1} ||A\boldsymbol{x}|| = \sup_{||\boldsymbol{x}|| = 1} ||A\boldsymbol{x}||.$$

Proof. The key point is that

$$\frac{||A\boldsymbol{x}||}{||\boldsymbol{x}||} = \left|\left|\frac{1}{||\boldsymbol{x}||}A\boldsymbol{x}\right|\right| = \left|\left|A\frac{\boldsymbol{x}}{||\boldsymbol{x}||}\right|\right|.$$

Since $||\frac{\boldsymbol{x}}{||\boldsymbol{x}||}|| = 1$

$$\{||A\boldsymbol{x}||/||\boldsymbol{x}|| \mid \boldsymbol{x} \neq \boldsymbol{0}\} = \{||A\boldsymbol{x}^*|| \mid ||\boldsymbol{x}^*|| = 1\} \subseteq \{||A\boldsymbol{x}|| \mid ||\boldsymbol{x}|| \leq 1\}.$$

It implies

$$||A|| = \sup_{||\boldsymbol{x}^*||=1} ||A\boldsymbol{x}^*|| \leq \sup_{||\boldsymbol{x}||\leq 1} ||A\boldsymbol{x}|| \leq \sup_{||\boldsymbol{x}||\leq 1} \frac{||A\boldsymbol{x}||}{||\boldsymbol{x}||} \leq ||A||.$$

□

We will see some examples of when $||A|| = ||A||_2$ and $||A|| < ||A||_2$ in tomorrow's recitation.

4 January 20

Starting with the Euclidean norm on $\mathbb{R}^n$, we can define the metric $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ by $d(\boldsymbol{x}, \boldsymbol{y}) = ||\boldsymbol{x} - \boldsymbol{y}||$. It admits a nice geometric interpretation: $d(\boldsymbol{x}, \boldsymbol{y})$ is the length of the line segment from $\boldsymbol{x}$ to $\boldsymbol{y}$. Therefore we could define the so-called Euclidean ball centered at $\boldsymbol{a} \in \mathbb{R}$ with radius $R \in \mathbb{R}$ to be

$$B_R(\boldsymbol{a}) := \{\boldsymbol{x} \in \mathbb{R}^n \mid ||\boldsymbol{x} - \boldsymbol{a}|| < R\}.$$

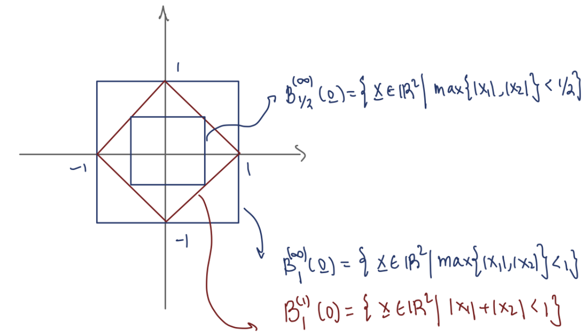
And in general, we can define the p-balls as follows:

Definition 4.1. We define the (open) p-ball of $\mathbb{R}^n$ centered at $\boldsymbol{a} \in \mathbb{R}^n$ with radius R by

$$B_R^{(p)}(\boldsymbol{a}) = \{\boldsymbol{x} \in \mathbb{R}^n \mid ||\boldsymbol{x} - \boldsymbol{a}||_p < R\}.$$

A quick exercise: Draw the corresponding p-balls for various $1 \leq p \leq \infty$. For example what does a unit p-ball look like when $p = 1, 2, \infty$? How do we relate them to their equivalence inequalities?

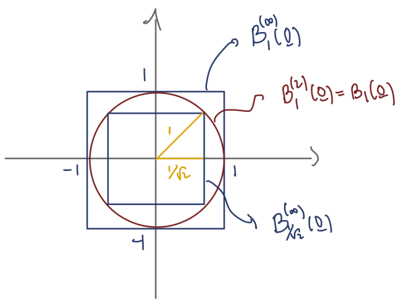
It is actually quite clear if we draw these balls with respect to different l^p-norms on $\mathbb{R}^2$:



They are related by the equivalence: $\|x\|_\infty \le \|\underline{x}\|_1 \le 2\|x\|_\infty$

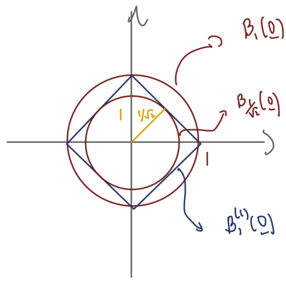
if $\underline{x} \in B^{(\infty)}_{1/2}(\underline{0}) \Rightarrow \underline{x} \in B^{(1)}_1(\underline{0})$

if $\underline{x} \in B^{(1)}_1(0) \Rightarrow \underline{x} \in B^{(\infty)}_1(0)$ (Both can be seen from the graph)



Reveals the relation:

$\|\underline{x}\|_\infty \le \|\underline{x}\|_2 \le \sqrt{2}\,\|\underline{x}\|_\infty$



Reveals the relation:

$\|\underline{x}\| \le \|\underline{x}\|_1 \le \sqrt{2}\,\|\underline{x}\|$

Now consider the space of linear operators $L(\mathbb{R}^n, \mathbb{R}^m)$. A linear operator $A : \mathbb{R}^n \to \mathbb{R}^m$ is a function satisfying

- $A(\boldsymbol{x} + \boldsymbol{y}) = A(\boldsymbol{x}) + A(\boldsymbol{y})$;
- $A(\lambda \boldsymbol{x}) = \lambda A(\boldsymbol{x})$,

for any $\boldsymbol{x}, \boldsymbol{x} \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$. A function $L \in L(\mathbb{R}^n, \mathbb{R}^m)$ is linear if and only if there exists a unique $(m \times n)$-matrix A such that $L(\boldsymbol{x}) = A\boldsymbol{x}$ for any $\boldsymbol{x} \in \mathbb{R}^n$. Therefore we can also use the notation A for the matrix to represent the corresponding linear operator.

Just like we put different norms on $\mathbb{R}^n$, we can also define norms for linear operators. The norms of linear operators/matrices we have seen so far are the followings:

Given a linear operator, or an $(m \times n)$-matrix $A = [a_{ij}]$,

- First of all the matrix A can be viewed as a mn-dimensional vector for which we can take either sup norm or standard Euclidean norm. Explicitly, we have

$$||[a_{ij}]||_\infty = \max_{i,j} |a_{ij}|, \quad ||[a_{ij}]||_2 = \sqrt{\sum a_{ij}^2}.$$

The latter is called the *Frobenius norm* on $L(\mathbb{R}^n, \mathbb{R}^m)$. Though these two norms may not be better choices, they do provide some useful information. Let us fix some notation first:

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^n a_{1j} x_j \\ \vdots \\ \sum_{j=1}^n a_{mj} x_j \end{bmatrix}$$

and write $A = [A_1, \cdots, A_m]^T$ where A_i are the ith row vector of A.
It can be checked (check yourself) that for any $\boldsymbol{x} \in \mathbb{R}^n$,

$$||A\boldsymbol{x}|| \leq n\sqrt{n}||A||_\infty ||\boldsymbol{x}||$$

which implies that for any nonzero $\boldsymbol{x} \in \mathbb{R}^n$, the quotient $||A\boldsymbol{x}||/||\boldsymbol{x}||$ is bounded from above by a constant which only depends on the matrix A and n. Therefore it is natural to consider the set

$$\mathcal{S} := \left\{ \frac{||A\boldsymbol{x}||}{||\boldsymbol{x}|||} \,\middle|\, \boldsymbol{x} \in \mathbb{R}^n \setminus \mathbf{0} \right\}$$

which is nonempty and bounded from above (by $n\sqrt{n}||A||_\infty$ for instance although it is not optimal). Thus by the completeness axiom it is legit to define the *operator norm*

$$||A|| := \sup_{\boldsymbol{x} \neq \mathbf{0}} \frac{||A\boldsymbol{x}||}{||\boldsymbol{x}||}$$

and we proved in class that it satisfies

$$||A|| \leq ||A||_2.$$

Remark 4.2. Recall that we apply the Cauchy-Schwartz inequality to show

$$\begin{aligned} ||A\boldsymbol{x}||^2 &= (A_1 \cdot \boldsymbol{x})^2 + \cdots + (A_m \cdot \boldsymbol{x})^2 \\ &\leq ||A_1||^2 ||\boldsymbol{x}||^2 + \cdots + ||A_m||^2 ||\boldsymbol{x}||^2 \\ &\leq ||A||_2^2 ||\boldsymbol{x}||^2. \end{aligned}$$

Note that this also proves the set $\mathcal{S}$ is bounded from above. Here is another simple application of the Cauchy-Schwartz inequality. Consider the expression $(2x + 3y + 4z)$. How to find its maximum over (x, y, z) on the unit sphere?

By the Cauchy-Schwartz inequality we get

$$\begin{aligned}
2x+3y+4z &= (2,3,4)\cdot(x,y,z) \\
&\le ||(2,3,4)||||(x,y,z)|| \\
&= \sqrt{29}\sqrt{x^2+y^2+z^2} \\
&= \sqrt{29}
\end{aligned}$$

with equality holds when $(x,y,z)=\lambda(2,3,4)$ for some scalar λ. Therefore we get the linear expression $(2x+3y+4z)$ attains its maximum $\sqrt{29}$ at the point

$$\left(\frac{\pm 2}{\sqrt{29}}, \frac{\pm 3}{\sqrt{29}}, \frac{\pm 4}{\sqrt{29}}\right)$$

on the unit sphere. (This gives a special case of $||A|| = ||A||_2$ when $m=1$.)

Question: When is $||A|| = ||A||_2$?

a) First consider the lower dimensional case. Let $m=1$ or $n=1$. If $n=1$, *i.e.,* the domain is $\mathbb{R}$ and A is an m-dimensional column vector. Thus

$$||A|| = \sup_{x=1} ||Ax|| = ||A||_2.$$

If $m=1$, then $A=[a_1,\cdots,a_n]$ can be viewed as an n-dimensional row vector and the operator is precisely taking dot product. Take $\boldsymbol{x} = (x_1,\cdots,x_n)$ with

$$x_i = \frac{a_i}{\sqrt{\sum a_i^2}},$$

then

$$||A\boldsymbol{x}|| = \sqrt{\sum a_i^2} = ||A||_2$$

meaning that an upper bound $||A||_2$ is attained thus $||A|| = ||A||_2$.

b) The equality immediately fails to hold when $m=n=2$. Consider the diagonal matrix

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}.$$

Let's first find its operator norm (though we will know the operator norm of a symmetric matrix is the max eigenvalue). It suffices to find the sup of $x^2+(2y)^2$ on the unit circle $x^2+y^2=1$. Since

$$x^2+4y^2 = 1+3y^2$$

and $y\le 1$, then the expression $||A\boldsymbol{x}||^2 = 1+3y^2$ attains its maximum 4 at $(x,y)=(0,1)$. Therefore

$$||A|| = 2 < \sqrt{5} = ||A||_2.$$

Facts: If A is a (square) symmetric matrix (thus all eigenvalues are real), then its operator norm $||A||$ is the maximal absolute eigenvalue. If A is neither a symmetric matrix nor a square matrix consider the positive semi-definite symmetric matrix A^TA (thus all eigenvalues are nonnegative reals), then the operator norm $||A||$ is the square root of maximal eigenvalue of A^TA.

- In general, we can even define the (p,q)-norm of linear operators in $L(\mathbb{R}^n, \mathbb{R}^m)$ where the domain $\mathbb{R}^n$ equipped with p-norm and the codomain $\mathbb{R}^m$ equipped with q-norm. Define
$$||A||_{p,q} = \sup_{\boldsymbol{x}\neq\boldsymbol{0}} \frac{||A\boldsymbol{x}||_q}{||\boldsymbol{x}||_p}$$
for $A \in L(\mathbb{R}^n, \mathbb{R}^m)$.
  a) If $p = q = 2$, then $(2,2)$-norm is just the operator norm.
  b) If $p = q = 1$, recall $||\boldsymbol{x}||_1 = \sum |x_i|$. Since
$$\begin{aligned} ||A\boldsymbol{x}||_1 &= \sum_{i=1}^m \left| \sum_{j=1}^n a_{ij}x_j \right| \\ &\leq \sum_{i=1}^m \sum_{j=1}^n |a_{ij}||x_j| \\ &= \sum_{j=1}^n \sum_{i=1}^m |a_{ij}||x_j| \\ &\leq \max_i \{||col_i(A)||_1\} ||\boldsymbol{x}||_1 \end{aligned}$$
  Then $||A\boldsymbol{x}||_1/||\boldsymbol{x}||_1$ is bounded from above by the maximum absolute column sum. Check if it can be attained by taking special $\boldsymbol{x} \in \mathbb{R}^n$.
  c) Similarly, we can work out the case for $p = q = \infty$. Since
$$||A\boldsymbol{x}||_\infty = \max_i \left\{ \left| \sum_{j=1}^n a_{ij}x_j \right| \right\} \leq \max_i \left\{ \sum_{j=1}^n |a_{ij}| \right\} ||\boldsymbol{x}||_\infty$$
  it means that $||A\boldsymbol{x}||_\infty/||\boldsymbol{x}||_\infty$ is bounded from above by the maximum absolute row sum. Check if it can be attained by taking special $\boldsymbol{x} \in \mathbb{R}^n$.

5 January 25

The goal for the following two recitations are about topology on $\mathbb{R}^n$. I will prove some (nontrivial) propositions but we will mainly focus on examples.

Recall that a set $E \subseteq \mathbb{R}^n$ is **open** if for any point $\boldsymbol{x} \in E$ there is an $r > 0$ such that $B(\boldsymbol{x}, r) \subseteq E$. This is equivalent to say every point in E is an interior point to E.

Proposition 5.1. *Open balls are open.*

Proof. Given a ball centered at $\boldsymbol{a} \in \mathbb{R}^n$ with radius $r > 0$

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

For any point $\boldsymbol{b} \in B(\boldsymbol{a}, r)$, take

$$r_0 = \frac{r - ||\boldsymbol{b} - \boldsymbol{a}||}{2} > 0.$$

We claim that $B(\boldsymbol{b}, r_0) \subseteq B(\boldsymbol{a}, r)$. For any $\boldsymbol{y} \in B(\boldsymbol{b}, r_0)$, by triangle inequality we have

$$||\boldsymbol{y} - \boldsymbol{a}|| \leq ||\boldsymbol{y} - \boldsymbol{b}|| + ||\boldsymbol{b} - \boldsymbol{a}|| < \frac{r + ||\boldsymbol{b} - \boldsymbol{a}||}{2} < \frac{r + r}{2} = r$$

which implies that $\boldsymbol{y} \in B(\boldsymbol{a}, r)$. Therefore $B(\boldsymbol{b}, r_0) \subseteq B(\boldsymbol{a}, r)$ and it shows that any ball is open. $\square$

Before looking at some examples let's prove some of the basic properties:

Proposition 5.2. *Given $E \subseteq \mathbb{R}^n$:*

(3) $\boldsymbol{x} \in \partial E$ if and only if for any $r > 0$, we have $B(\boldsymbol{x}, r) \cap E \neq \emptyset$ and $B(\boldsymbol{x}, r) \cap E^c \neq \emptyset$.

(9) $\partial E = \partial E^c$.

(11) E is closed if and only if $\partial E \subseteq E$ if and only if $E' \subseteq E$ if and only if $\{$limit points for $E\} \subseteq E$.

Proof. (3) Assume $\boldsymbol{x}$ is a boundary point of E which by definition means $\boldsymbol{x}$ is neither an interior point of E nor an interior point of E^c (we call this exterior point of E). If there exists some $r > 0$ such that $B(\boldsymbol{x}, r) \cap E = \emptyset$ then $B(\boldsymbol{x}, r) \subseteq E^c$ meaning that $\boldsymbol{x}$ is an interior point of E^c, a contradiction. If there exists $r > 0$ such that $B(\boldsymbol{x}, r) \cap E^c = \emptyset$ then $B(\boldsymbol{x}, r) \subseteq E$ which means $\boldsymbol{x}$ is an interior point of E, again a contradiction. The other direction follows similarly.

(9) Using (3) and $(E^c)^c = E$.

(11) Start with E is closed. For any $\boldsymbol{x} \in \partial E$, the point $\boldsymbol{x}$ is not an interior point of E^c. Since E^c is open, we must have $\boldsymbol{x} \notin E^c$ and thus $\boldsymbol{x} \in E$.
Assume $\partial E \subseteq E$ show E is closed. In other words, we need to prove E^c is open. For any $\boldsymbol{x} \in E^c$, we have $\boldsymbol{x} \notin \partial E$. Then $\boldsymbol{x}$ has to be an interior point of E^c which shows E^c is open.
Now assume E is closed. $\boldsymbol{x} \in E'$ means that $\boldsymbol{x}$ is a cluster point of E (Recall any ball centered at $\boldsymbol{x}$ contains a point in E other than $\boldsymbol{x}$ itself). If there is $\boldsymbol{x} \in E'$ but not in E, we have $\boldsymbol{x} \in E^c$. Since E^c is open there exists $r > 0$ such that $B(\boldsymbol{x}, r) \subseteq E^c$ which implies that $B(\boldsymbol{x}, r) \cap E = \emptyset$ contradicting $\boldsymbol{x} \in E'$.

Next assume $E' \subseteq E$. We claim that $\partial E \subseteq E$. By contradiction assume there is $\boldsymbol{x} \in \partial E$ which is not in E. Then $\boldsymbol{x} \in E^c$. Since $\boldsymbol{x}$ is a boundary point of E then for any $r > 0$ we must have $B(\boldsymbol{x}, r) \cap E \neq \emptyset$ which implies that $B(\boldsymbol{x}, r) \setminus \{\boldsymbol{x}\} \cap E \neq \emptyset$ because $\boldsymbol{x} \notin E$. It then implies by definition of cluster points, $\boldsymbol{x} \in E' \subseteq E$ which contradicts the assumption that $\boldsymbol{x} \notin E$.
Since set of limit points contains set of cluster points, thus $\{\text{limit points for } E\} \subseteq E$ implies $E' \subseteq E$.
It remains to show the following: if E is closed then E contains all its limit points. Since E^c is open, then for any $\boldsymbol{x} \in E^c$ there exists $r > 0$ such that $B(\boldsymbol{x}, r) \cap E = \emptyset$. Thus for any sequence $\{\boldsymbol{x}_n\} \subseteq E$ we have $||\boldsymbol{x}_n - \boldsymbol{x}|| \geq r > 0$ for all $n \in \mathbb{N}$. Then $\boldsymbol{x}$ cannot be a limit point of E. Therefore any point in E^c cannot be limit point, so $\{\text{limit points for } E\} \subseteq E$. □

Let's start with some easy examples:

Example 5.3.
- Consider $E = (-1, 1] \cup \{3\} \subset \mathbb{R}$. Then $E^\circ = (-1, 1)$, $E' = [-1, 1]$ and $\partial E = \{-1, 1, 3\}$. Note that 3 is an isolated point (so it is a limit point) but it is not a cluster point.
- Consider $E = [0, 1] \cup \mathbb{Q} \subset \mathbb{R}$. Then $E^\circ = \emptyset$ because the density of irrationals in $\mathbb{R}$; $E' = [0, 1] = \partial E$ also by the density statements.
- Consider $E = \mathbb{Q} \times \mathbb{R} \subset \mathbb{R}^2$. Similarly as in the second example, $E^\circ = \emptyset$ and $E' = \partial E = \mathbb{R}^2$.
- Consider $E = (-1, 1) \times [-2, 2]$. Then $E^\circ = (-1, 1) \times (-2, 2)$, $E' = [-1, 1] \times [-2, 2]$ and $\partial E = (\{-1, 1\} \times [-2, 2]) \cup ([-1, 1] \times \{-2, 2\})$.
- Consider $E = \{(x, y) \mid x = y\} \subset \mathbb{R}^2$, namely the diagonal in $\mathbb{R}^2$. Then $E^\circ = \emptyset$ and $E' = E = \partial E$. Note that later we will see that E is the graph of the continuous function $f(x) = x$ and thus E is closed. Indeed we have $\overline{E} = \partial E \cup E^\circ = E$.
- Consider
$$E = \{(x, y) \mid x^2 + y^2 < 1\} \setminus \left(\left(\mathbb{Q} \cap \left[-\frac{1}{2}, \frac{1}{2}\right]\right) \times \{0\}\right) \subset \mathbb{R}^2.$$
Then $E^\circ = \{(x, y) \mid x^2 + y^2 < 1\} \setminus ([-1/2, 1/2] \times \{0\})$, $E' = \{(x, y) \mid x^2 + y^2 \leq 1\} = \overline{E}$, and $\partial E = \{(x, y) \mid x^2 + y^2 = 1\} \cup ([-1/2, 1/2] \times \{0\})$.

One last example, consider
$$E = \{(x, y) \mid x^8 - y^7 + 2xy \leq 1\} \subset \mathbb{R}^2.$$

- First E is closed. It will be easier to use continuity which we haven't yet covered so make sure to go back and look at this example again. We can consider the continuous function $f(x, y) = x^8 - y^7 + 2xy : \mathbb{R}^2 \to \mathbb{R}$. Then $E = f^{-1}((-\infty, 1])$ is closed in $\mathbb{R}^2$ since $(-\infty, 1]$ is closed in $\mathbb{R}$. Or use the fact that continuous function maps convergent sequence to convergent sequence.

- $E^\circ = \{(x,y) \mid x^8 - y^7 + 2xy < 1\}$. It suffices to show this set on the RHS is the largest open subset in E. We can show for any (x_0, y_0) satisfying $x_0^8 - y_0^7 + 2x_0y_0 = 1$, for any $r > 0$ the ball $B((x_0, y_0), r)$ is not contained in E. In other words, find $(x, y) \in B((x_0, y_0), r)$ such that $(x, y) \notin E$.
- Moreover E is not bounded. Pick a sequence $\{(0, n)\} \subset E$.

6 January 27

Taking interior/closure/boundary does not always commute with taking intersection or union. For instance we have already given an example showing that only one-sided inclusion $A^\circ \cup B^\circ \subseteq (A \cup B)^\circ$ holds. Another possible example can be $A = \mathbb{Q}$ and $B = \mathbb{R} \setminus \mathbb{Q}$ and thus $A^\circ = B^\circ = \emptyset$ but $(A \cup B)^\circ = \mathbb{R}$. In addition, we also have the following examples for taking closure and boundary:

- We ONLY have $\overline{A \cap B} \subseteq \overline{A} \cap \overline{B}$. We can take the same example as above, say $A = \mathbb{Q}$ and $B = \mathbb{R} \setminus \mathbb{Q}$. Since both A and B are dense in $\mathbb{R}$ then $\overline{A} = \overline{B} = \mathbb{R}$. But $A \cap B = \emptyset$ therefore we only get $\overline{A \cap B} = \emptyset \subsetneqq \mathbb{R} = \overline{A} \cap \overline{B}$.
- We ONLY have $\partial(A \cup B) \subseteq \partial A \cup \partial B$. An easy example for the strict inclusion is $A = (1, 2)$ and $B = [2, 3]$ where the boundary point 2 of A will turn to an interior point after taking the union. Namely, $\partial A = \{1, 2\}$ and $\partial B = \{2, 3\}$ but $A \cup B = (1, 3]$ and thus $\partial(A \cup B) = \{1, 3\} \subsetneqq \{1, 2, 3\} = \partial A \cup \partial B$.
- We ONLY have $\partial(A \cap B) \subseteq \partial A \cup \partial B$. The above example also shows that $\partial(A \cap B) \subsetneqq \partial A \cup \partial B$ since $A \cap B = \emptyset$.

Remark 6.1. One might want to ask if there is a relation between $\partial(A \cap B)$ and $\partial A \cap \partial B$ the answer is no, and either one of them can be big and small. Here is a small example. We can take two overlapping open disks in $\mathbb{R}^2$, called A and B. Then $\partial(A \cap B)$ is the boundary of the overlapping area, but $\partial A \cap \partial B$ is just two separated points. But if we take disk A tangent to B, both open, then $\partial(A \cap B) = \emptyset$ but $\partial A \cap \partial B$ now is a singleton.

Now consider the following question: Given a nonempty set $E \subseteq \mathbb{R}^n$ together with a point $\boldsymbol{a} \in E$, how to canonically define the distance between point $\boldsymbol{a}$ and the set E? If $\boldsymbol{a} \in E$ then it natural to get the distance being defined by 0. Moreover since the norm is always nonnegative, we can define the distance between $\boldsymbol{a}$ and E to be

$$\rho(\boldsymbol{a}, E) = \inf_{\boldsymbol{x} \in E} ||\boldsymbol{x} - \boldsymbol{a}||,$$

i.e., the greatest lower bound of norms from $\boldsymbol{a}$ to any point in E. One might ask when will $\rho(\boldsymbol{a}, E) > 0$ happen, is $\boldsymbol{a} \notin E$ sufficient to imply $\rho > 0$? The answer is no, for example one can take E be the open half plane in $\mathbb{R}^2$ and $\boldsymbol{a} = \boldsymbol{0}$ be the origin.

But if we add an extra condition on the set E, we can have strictly positivity

Example 6.2. **FACT**: If $E \subseteq \mathbb{R}^n$ is closed and $\boldsymbol{a} \notin E$, then $\rho(\boldsymbol{a}, E) > 0$.

We will prove this statement in two different ways. First of all, using contradiction. If $\inf_{\boldsymbol{x}\in E} ||\boldsymbol{x}-\boldsymbol{a}|| = 0$, then there exists a sequence $\{\boldsymbol{x}_n\} \subset E$ such that $||\boldsymbol{x}_n - \boldsymbol{a}|| \to \boldsymbol{0}$, *i.e.*, $\{\boldsymbol{x}_n\} \to \boldsymbol{a}$ eventually. Since E is closed the limit point $\boldsymbol{a} \in E$ which is a contradiction.

Alternatively, since E is closed and $\boldsymbol{a} \notin E$, E^c is open and $\boldsymbol{a}$ is an interior point of E^c. Thus there is $r > 0$ such that $B(\boldsymbol{a}, r) \cap E = \emptyset$. By the definition of ρ, we have for any $\boldsymbol{x} \in E$, $||\boldsymbol{x} - \boldsymbol{a}|| \geq r > 0$ thus $\rho(\boldsymbol{a}, E) \geq r > 0$.

Now let's look at some explicit examples:

Example 6.3.
- $\rho(0, [1, \infty)) = 1$.
  Since the $\min_{x\in[1,\infty)} |x|$ is attained at $x = 1$ thus it is also the inf.
- $\rho(0, (1, \infty)) = 1$.
  Since it is easy to show $\inf_{x\in(1,\infty)} |x| = \inf(1, \infty) = 1$.

The notion of distance can also characterize cluster point.

Example 6.4. FACT: $\boldsymbol{a} \in E'$ if and only if $\rho(\boldsymbol{a}, E \setminus \{\boldsymbol{a}\}) = 0$.

If $\boldsymbol{a} \in E'$, then by definition $(B(\boldsymbol{a}, r) \setminus \{\boldsymbol{a}\}) \cap E \neq \emptyset$ for all $r > 0$. It means that for any $r > 0$, there exists a point $\boldsymbol{x}_r \in E$ but $\boldsymbol{a} \neq \boldsymbol{x}_r$ with $||\boldsymbol{x}_r - \boldsymbol{a}|| < r$. Since $r > 0$ is arbitrary, we must have $\rho(\boldsymbol{a}, E) = \inf_{\boldsymbol{x}\in E} ||\boldsymbol{x} - \boldsymbol{a}|| = 0$.

Conversely, if $\rho(\boldsymbol{a}, E \setminus \{\boldsymbol{a}\}) = 0$, then for any $r > 0$, r cannot be a lower bound of the set

$$\{||\boldsymbol{x} - \boldsymbol{a}|| \mid \boldsymbol{x} \in E \setminus \{\boldsymbol{a}\}\}.$$

It means that for any positive $r > 0$ there exists some point $\boldsymbol{x}_r \in E \setminus \{\boldsymbol{a}\}$ such that $||\boldsymbol{x}_r - \boldsymbol{a}|| < r$. Thus $(B(\boldsymbol{a}, r) \setminus \{\boldsymbol{a}\}) \cap E \neq \emptyset$ for any $r > 0$. Therefore $\boldsymbol{a} \in E'$.

Example 6.5. $\rho(\sqrt{2}, \mathbb{Q}) = 0$.

Since $\mathbb{Q}$ is dense in $\mathbb{R}$ or $\sqrt{2} \in \mathbb{Q}' = \mathbb{R}$ thus $\rho(\sqrt{2}, \mathbb{Q}) = 0$.

Remark 6.6.
- Can we generalize the above definition to a distance function defined on $(\mathbb{R}^n, ||\cdot||)$, if we define similarly

$$\rho(A, B) := \inf_{\boldsymbol{a}\in A, \boldsymbol{b}\in B} ||\boldsymbol{a} - \boldsymbol{b}||$$

  for $A, B \in \mathbb{R}^n$?
- Use the definition of $\rho(A, B)$ as above, if A, B are both closed sets in $\mathbb{R}^n$ and disjoint $A \cap B = \emptyset$, do we have $\rho(A, B) > 0$? Counterexamples?
  (1) Take

$$A = \{(x, y) \in \mathbb{R}^2 \mid y \leq 0\}, \quad B = \{(x, y) \in \mathbb{R}^2 \mid y \geq e^x\}.$$

  (2) Take $A = \mathbb{N}$ then $A^\circ = \emptyset, \partial A = \mathbb{N} \subseteq A$ thus A is closed. And

$$B = \left\{n + \frac{1}{2n} : n \in \mathbb{N}\right\}$$

  then $\partial B = B$. But $\rho(A, B) = \inf_{n\in\mathbb{N}} |1/(2n)| = 0$.
- How can we modify the condition in above to make $\rho(A, B) > 0$?

7 February 1

For limits of functions, since we have the proposition showed in class, we only need to consider the limit of each coordinate function f_i for $\boldsymbol{f} = (f_1, \cdots, f_m) : A \to \mathbb{R}^m$ where $A \subseteq \mathbb{R}^n$. Thus throughout today's recitation we will only consider real-valued multivariable function $f : A \to \mathbb{R}$ with domain $A \subseteq \mathbb{R}^n$.

Given $A \subseteq \mathbb{R}^n$ and $\boldsymbol{x}_0 \in A'$. We say that the function $f : A \to \mathbb{R}$ has limit l as $\boldsymbol{x} \to \boldsymbol{x}_0$ if for any $\epsilon > 0$ there is $\delta > 0$ such that for any $\boldsymbol{x} \in A \setminus \{\boldsymbol{x}_0\}$ with $||\boldsymbol{x} - \boldsymbol{x}_0|| < \delta$ we have $|f(\boldsymbol{x}) - l| < \epsilon$.

Let's start with two examples using only the definition:

Example 7.1. (1) Prove the following limit of a polynomial function

$$\lim_{(x,y)\to(2,1)} (x^2 + y^3 - 3) = 2.$$

For any $\epsilon > 0$. Since

$$\begin{aligned} |x^2 + y^3 - 5| &\leq |x^2 - 4| + |y^3 - 1| \\ &\leq |x - 2|(|x| + 2) + |y - 1|(|y|^2 + |y| + 1). \end{aligned}$$

Now we restrict $||(x, y) - (2, 1)|| = \sqrt{(x - 2)^2 + (y - 1)^2} < 1$. Then we have $|x - 2| < 1$ and $|y - 1| < 1$ which imply $|x| < 3$ and $|y| < 2$. And thus

$$\begin{aligned} |x^2 + y^3 - 5| &< 5|x - 2| + 7|y - 1| \\ &\leq \sqrt{74}||(x, y) - (2, 1)|| \end{aligned}$$

which means we can take $\delta = \min\{1, \epsilon/\sqrt{74}\}$.

(2) Prove the following limit of a rational function

$$\lim_{(x,y)\to(1,2)} \frac{x}{1 + y} = \frac{1}{3}.$$

This can be proved similarly as above. For any $\epsilon > 0$. Since

$$\left|\frac{x}{1 + y} - 3\right| = \frac{|3x - 1 - y|}{3|1 + y|} \leq \frac{3|x - 1| + |y - 2|}{3|1 + y|}.$$

So if $||(x, y) - (1, 2)|| < 1$ then $|y - 2| < 1$. Thus

$$3 = |2 - y + y + 1| \leq |y - 2| + |y + 1| < 1 + |y + 1| \implies |y + 1| > 2.$$

Then

$$\left|\frac{x}{1 + y} - 3\right| \leq \frac{1}{2}|x - 1| + \frac{1}{6}|y - 2| \leq \frac{\sqrt{10}}{6}||(x, y) - (1, 2)||.$$

Therefore we only need to take $\delta = \min\{1, 6\epsilon/\sqrt{10}\}$.

Indeed, we have the following facts:

- If $P(\boldsymbol{x})$ is a polynomial function on $\mathbb{R}^n$, then for any $\boldsymbol{x}_0 \in \mathbb{R}^n$ we have $\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} P(\boldsymbol{x}) = P(\boldsymbol{x}_0)$.
- If $P(\boldsymbol{x})$ and $Q(\boldsymbol{x})$ are two polynomials on $\mathbb{R}^n$, then for any $\boldsymbol{x}_0 \in \mathbb{R}^n$ with $Q(\boldsymbol{x}_0) \neq 0$ we have
$$\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} \frac{P(\boldsymbol{x})}{Q(\boldsymbol{x})} = \frac{P(\boldsymbol{x}_0)}{Q(\boldsymbol{x}_0)}.$$
But if $Q(\boldsymbol{x}_0) = 0$ then the limit may not exist in general.

For example to the second item above: for any $m, k \in \mathbb{N}$ consider $f(x,y) = \frac{x^m}{y^k}$ as $(x,y) \to (0,0)$. There are infinitely many ways to approach the origin $(0,0)$, for example along the parametrized curve

$$\begin{cases} x(t) = \alpha t^k \\ y(t) = \beta t^m \end{cases}$$

for $-1 \leq t \leq 1$ and α, β nonzero constants. Or even discretely we can also take sequences:

$$(x_n, y_n) = \left(\frac{1}{n^k}, \frac{1}{n^m}\right), \quad (a_n, b_n) = \left(\frac{2}{n^k}, \frac{1}{n^m}\right)$$

both approaching to $(0,0)$ as $n \to \infty$. But $f(x_n, y_n) \to 1$ while $f(a_n, b_n) = 2^m$.

Now let's look at some other examples

Example 7.2.

- Consider
$$\lim_{(x,y)\to(0,0)} \frac{x^3 + y^3}{x^2 + y^4}.$$
Let (x,y) approach to $(0,0)$ along the y-axis with $y \to 0^+$. Then $x = 0$ and thus
$$\lim_{y\to 0^+} \frac{y^3}{y^4} = \infty.$$
But if (x,y) approaches to $(0,0)$ along the x-axis, then
$$\lim_{x\to 0} \frac{x^3}{x^2} = \lim_{x\to 0} x = 0.$$
Therefore the limit does not exist.
- Consider the following limit
$$\lim_{(x,y)\to(0,0)} \frac{\sin(x^3 y)}{\sqrt{x^8 + y^6}}.$$

First of all we can write

$$\frac{\sin(x^3y)}{\sqrt{x^8+y^6}} = \frac{\sin(x^3y)}{x^3y}\frac{x^3y}{(x^8+y^6)^{1/2}}.$$

For the limit of $\sin(x^3y)/x^3y$, it is simply $\sin(z)/z$ composed with $z \mapsto x^3y$. Thus by composition theorem and $\lim_{z\to 0}\frac{\sin z}{z} = 1$ we have

$$\lim_{(x,y)\to(0,0)} \frac{\sin(x^3y)}{x^3y} = 1.$$

For the other limit, we use the trick used in class:

$$x \le (x^8+y^6)^{1/8}, \quad y \le (x^8+y^6)^{1/6}.$$

Thus

$$\left|\frac{x^3y}{(x^8+y^6)^{1/2}}\right| \le \frac{(x^8+y^6)^{\frac{3}{8}+\frac{1}{6}}}{(x^8+y^6)^{1/2}} = (x^8+y^6)^{\frac{1}{24}} \to 0$$

as $(x,y)\to(0,0)$. It implies that

$$\lim_{(x,y)\to(0,0)} \frac{\sin(x^3y)}{\sqrt{x^8+y^6}} = 0.$$

8 February 3

We are going to see more examples of limits of functions today, but before starting I want to make a quick note on the composition of limits. The composition of limits tells us that given $A \subseteq \mathbb{R}^n$ and $B \subseteq \mathbb{R}^p$ and consider the composition of functions f and g:

$$B \xrightarrow{\boldsymbol{g}} g(B) \subseteq A \xrightarrow{\boldsymbol{f}} \mathbb{R}^m.$$

Let $\boldsymbol{x}_0 \in A'$ and $\boldsymbol{t} \in B'$. If $\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{l}$ and $\lim_{\boldsymbol{t}\to\boldsymbol{t}_0} \boldsymbol{g}(\boldsymbol{t}) = \boldsymbol{x}_0$, then $\lim_{\boldsymbol{t}\to\boldsymbol{t}_0} \boldsymbol{f}(\boldsymbol{g}(\boldsymbol{t})) = \boldsymbol{l}$. One special case for this theorem is to take the outer function $\boldsymbol{g}$ be simply the projection map. In other words, we consider for $m < n$, the composition

$$f : \mathbb{R}^n \longrightarrow \mathbb{R}^m \xrightarrow{g} \mathbb{R}$$

where the first map is the projection map given by

$$\begin{aligned} p : \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ (x_1, \cdots, x_m, \cdots, x_n) &\longmapsto (x_1, \cdots, x_m). \end{aligned}$$

Under this special situation, take for example $\boldsymbol{x}_0 \in \mathbb{R}^n$ and $p(\boldsymbol{x}_0) \in \mathbb{R}^m$ be the point projected from $\boldsymbol{x}_0$. Then the composition of limits means that in order to figure whether

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

it suffices to check whether

$$\lim_{\boldsymbol{x}'\to\boldsymbol{x}'_0} g(\boldsymbol{x}') = l$$

where we write $\boldsymbol{x}' = p(\boldsymbol{x})$ and $\boldsymbol{x}'_0 = p(\boldsymbol{x}_0)$, *i.e.*, check the limit in lower dimensional vector space $\mathbb{R}^m$. To see this, we only need to prove

$$\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} p(\boldsymbol{x}) = p(\boldsymbol{x}_0) = \boldsymbol{x}'_0.$$

Indeed this is true because projection maps are continuous. But it is also easy to see because for any $\epsilon > 0$, just take $\delta = \epsilon$, then for any $\boldsymbol{x} \in \mathbb{R}^n$ with $||\boldsymbol{x} - \boldsymbol{x}_0|| < \delta$, after projection to lower dimension subspace, we always have

$$||p(\boldsymbol{x}) - p(\boldsymbol{x}_0)|| \leq ||\boldsymbol{x} - \boldsymbol{x}_0|| < \delta = \epsilon.$$

Example 8.1. Suppose $f : \mathbb{R}^3 \to \mathbb{R}$ is defined by $f(x, y, z) = x^2 + y^4$. To study the limit of $f(x, y, z)$ it suffices to study the limit of $g(x, y) = f(x, y, 0)$ with domain in $\mathbb{R}^2$.

Now let's continue with more examples about limits of functions. The first one is a slight generalization of the example we have seen last time:

Example 8.2. Consider the following limit

$$\lim_{(x,y)\to(0,0)} \frac{x^3|y|^a}{x^4 + y^2}.$$

Find α so that the limit exists.

We will use the same trick as

$$x \leq (x^4 + y^2)^{1/4}, \quad y \leq (x^4 + y^2)^{1/2}$$

to write

$$0 \leq \frac{x^3|y|^a}{x^4 + y^2} \leq \frac{(x^4 + y^2)^{3/4}(x^4 + y^2)^{a/2}}{x^4 + y^2} = (x^4 + y^2)^{\frac{3}{4}+\frac{a}{2}-1}$$

which approaches to 0 as $(x, y) \to (0, 0)$ if the exponent $\frac{3}{4} + \frac{a}{2} - 1 > 0$. Thus if $a > \frac{1}{2}$ the limit exists and equals to zero.

For $a \leq \frac{1}{2}$, let us consider two cases

- If $0 < a \leq \frac{1}{2}$, then $\frac{1}{a} > 2$. If approach to the origin along the x-axis, the limit equals 0; if take path to the origin along $y = x^{1/a}$ then

$$\lim_{(x,y)\to(0,0), y=x^{1/a}} \frac{x^3|y|^a}{x^4+y^2} = \lim_{x\to 0} \frac{x^4}{x^4+x^{2/a}} = \lim_{x\to 0} \frac{1}{1+x^{\frac{2}{a}-4}} = 1 \neq 0$$

  since $\frac{2}{a} > 4$. Thus the limit cannot exist.
- If $a \leq 0$ we consider

$$\lim_{(x,y)\to(0,0)} \frac{x^3}{|y|^a(x^4+y^2)}$$

  where $a \geq 0$. Then along the y-axis the limit equal 0 while along the line $y = x$, the limit is ∞. Thus the limit does not exist.

Some other examples:

Example 8.3. (1) Consider the limit

$$\lim_{(x,y)\to(0,1)} \frac{x^{\frac{3}{2}}\sin(xy)}{x^2+(y-1)^2}.$$

We can write

$$\frac{x^{\frac{3}{2}}\sin(xy)}{x^2+(y-1)^2} = \frac{\sin(xy)}{xy}\frac{x^{\frac{5}{2}}y}{x^2+(y-1)^2}.$$

Since $xy \to 0$ as $(x,y) \to (0,1)$ the first term has limit 1. Moreover we have

$$\left|\frac{x^{\frac{5}{2}}y}{x^2+(y-1)^2}\right| \leq \frac{|x^{\frac{5}{2}}y|}{x^2} = \sqrt{|x|}|y| \to 0$$

as $(x,y) \to (0,1)$ thus the limit equals 0.

(2) Consider the limit

$$\lim_{(x,y)\to(0,0)} \frac{x^2e^{-1/y}}{x^2+y^2}$$

on $A = \{(x,y) \in \mathbb{R}^2 \mid y > 0\}$.
Since

$$0 \leq \left|\frac{x^2e^{-1/y}}{x^2+y^2}\right| \leq \frac{x^2e^{-1/y}}{x^2} = e^{-1/y} \to 0$$

as $y \to 0^+$. Thus the limit equals 0.

(3) Consider the limit

$$\lim_{(x,y,z)\to\mathbf{0}} \frac{xy^2z^3}{(x^2+|y|^3+z^4)^{7/8}}.$$

Since

$$\left|\frac{xy^2z^3}{(x^2+|y|^3+z^4)^{7/8}}\right| \le \frac{1}{2}\frac{(xy^2)^2}{|y|^{21/8}} + \frac{1}{2}\frac{z^6}{|z|^{7/2}} = \frac{1}{2}x^2|y|^{11/8} + \frac{1}{2}|z|^{5/2} \to 0$$

as $(x, y, z) \to \mathbf{0}$.

(4) Consider the limit

$$\lim_{(x,y)\to(0,0)} \frac{x\sin(y/\sqrt{x})}{\sqrt{x^2+y^2}}.$$

Again by estimation

$$0 \le \left|\frac{x\sin(y/\sqrt{x})}{\sqrt{x^2+y^2}}\right| \le \left|\frac{\sqrt{x}y}{\sqrt{x^2+y^2}}\right| \le \frac{|\sqrt{x}y|}{y} = \sqrt{x} \to 0$$

as $(x, y) \to (0, 0)$. Note that we are using $\sin(t) \le t$ for all real $t \ge 0$ for the second inequality.

9 February 8

Analogous to one variable functions, we can also define continuous, uniformly continuous, Hölder continuous functions.

Given real-valued function $f : E \to \mathbb{R}^m$ with $E \subseteq \mathbb{R}^n$. We say f is continuous at $\boldsymbol{a} \in E$ if for any $\epsilon > 0$ there exists $\delta > 0$ such that $|f(\boldsymbol{x}) - f(\boldsymbol{a})| < \epsilon$ whenever $\boldsymbol{x} \in E$ with $||\boldsymbol{x} - \boldsymbol{a}|| < \delta$. In particular, if $\boldsymbol{a}$ is a cluster point of E we can also use limit to characterize the continuity at $\boldsymbol{a}$. Moreover continuity can also be characterized sequentially and topologically. For example, sometimes to show a function is not continuous at a certain point $\boldsymbol{a}$ it suffices to find a sequence $\{\boldsymbol{x}_n\} \subset E$ with limit $\boldsymbol{a}$ but $f(\boldsymbol{x})$ does not converge to $f(\boldsymbol{a})$.

Example 9.1. Consider a real-valued function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

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

It is not hard to show that $f(x, y)$ is not continuous at $(0, 0)$. Since by taking a sequence

$$(x_n, y_n) = \left(\frac{1}{n}, \frac{1}{\sqrt{n}}\right) \to (0, 0)$$

as $n \to \infty$, we have

$$f(x_n, y_n) = 1 \neq f(0, 0) = 0.$$

From this choice of path we see that in order to have continuity of f at $(0, 0)$ that the speed of y^2 approaching to 0 must be faster than that of x. That is to say, if we restrict the domain of f to some subset of $\mathbb{R}^2$ which excludes such paths it is still possible to have f being continuous at $(0,0)$. Based on this fact, we consider the same function f on $D \subset \mathbb{R}^n$ defined by

$$D = \{(x,y) \in \mathbb{R}^2 \mid |y| \leq x \leq 1\}.$$

It is easy to see that $(0,0) \in D$. But now the function $f : D \to \mathbb{R}$ is actually continuous at $(0,0)$. This is because on D,

$$\left|\frac{y^2}{x} - 0\right| = \frac{|y|}{x}|y| \leq |y| \to 0$$

as $(x,y) \to (0,0)$.

Given a function $\boldsymbol{f} : E \to \mathbb{R}^m$ with $E \subseteq \mathbb{R}^n$. We say $\boldsymbol{f}$ is uniformly continuous on E if for any $\epsilon > 0$ there exists $\delta > 0$ such that $||\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{y})|| < \epsilon$ whenever $\boldsymbol{x}, \boldsymbol{y} \in E$ with $||\boldsymbol{x} - \boldsymbol{y}|| < \delta$. There is also a sequential characterization of uniform continuity. Moreover if $\boldsymbol{f}$ is uniformly continuous on E then it must be continuous on E. But the converse may not be true.

Example 9.2. (1) Consider a function $f : D \to \mathbb{R}$ given by $f(x,y) = \frac{1}{x+y}$ with $D = \{(x,y) \in \mathbb{R}^2 \mid x, y \in [0,1], (x,y) \neq (0,0)\}$.
The function f is of course continuous on D but it is not uniformly continuous on D. We consider two sequences,

$$(x_n, y_n) = \left(\frac{1}{n}, 0\right), \quad (x_n', y_n') = \left(\frac{1}{n+1}, 0\right).$$

Then

$$||(x_n, y_n) - (x_n', y_n')|| = \frac{1}{n(n+1)} \to 0$$

but $|f(x_n, y_n) - f(x_n', y_n')| = 1 \neq 0$.

(2) Similarly, consider $f : \mathbb{R}^2 \to \mathbb{R}$ given by $f(x,y) = x^2 + y^2$. Then f is continuous on $\mathbb{R}^2$. But if we take

$$(x_n, y_n) = (n, 0), \quad (x_n', y_n') = \left(n - \frac{1}{n}, 0\right)$$

then $||(x_n, y_n) - (x_n', y_n')|| = \frac{1}{n} \to 0$ but $|f(x_n, y_n) - f(x_n', y_n')| = 2 - \frac{1}{n^2} \to 2 \neq 0$.

Given a function $\boldsymbol{f} : E \to \mathbb{R}^m$ with $E \subseteq \mathbb{R}^n$. We say $\boldsymbol{f}$ is Hölderian of order $0 < \alpha \leq 1$ if there is a nonnegative real constant $C \in \mathbb{R}$ such that $||f(\boldsymbol{x}) - f(\boldsymbol{y})|| \leq C||\boldsymbol{x} - \boldsymbol{y}||^\alpha$ whenever $\boldsymbol{x}, \boldsymbol{y} \in E$. If $\boldsymbol{f}$ is Hölder continuous then it is also uniformly continuous. But again the converse may not be true.

Example 9.3. Consider a one variable example $f : [0, 1/2] \to \mathbb{R}$ given by

$$f(x) = \begin{cases} 0 & \text{if } x = 0; \\ \frac{1}{\ln x} & \text{if } x \in (0, 1/2]. \end{cases}$$

Then f is uniformly continuous since $\lim_{x\to 0} f(x) = 0$ thus it is continuous on $[0, 1/2]$ thus uniformly continuous. If f is Hölderian then we must have for any $y \in (0, 1/2]$,

$$|f(0) - f(y)| = |1/\ln y| \leq C|y|^\alpha$$

for some nonnegative constant $C \in \mathbb{R}$. Then it implies

$$C|y|^\alpha |\ln y| \geq 1.$$

But one can check by L'hopital that $\lim_{y\to 0^+} c|y|^\alpha |\ln y| = 0$. Thus f is not Hölderian.

Continuous functions have lots of properties such as IVT, Extreme value theorem and etc. Let's take look at one important example which is a generalization of estimation of positive definite quadratic form: If Q is a positive definite matrix *i.e.,* it is symmetric and $Q(\boldsymbol{x}) = \boldsymbol{x}^T Q\boldsymbol{x} > 0$ for any $\boldsymbol{x} \in \mathbb{R}^n$. Then for any nonzero $\boldsymbol{x}$ we have inequalities

$$\lambda_{\text{min}}||\boldsymbol{x}||^2 \leq Q(\boldsymbol{x}) \leq \lambda_{\text{max}}||\boldsymbol{x}||^2$$

where $\lambda_{\text{min}}, \lambda_{\text{max}}$ are minimal and maximal eigenvalues respectively (they are both positive).

Example 9.4.

- A homogeneous polynomial of degree d in $\mathbb{R}^n$ is a degree d polynomial $P(\boldsymbol{x})$ such that $P(t\boldsymbol{x}) = t^d P(\boldsymbol{x})$ for any nonzero scalar $t \in \mathbb{R}$. For example, we consider a degree 5 homogeneous polynomial in $\mathbb{R}^3$:

  $$P(x, y, z) = x^3y^2 + yz^4.$$

  This polynomial function P is continuous on $\mathbb{R}^3$. Thus it attains its extrema on the unit sphere $S_1(\mathbf{0})$ since $S_1(\mathbf{0})$ is closed and bounded thus compact (for closedness we have the unit sphere $S_1(\mathbf{0})$ is the preimage $g^{-1}(\{1\})$ where $g : \mathbb{R}^3 \to \mathbb{R}$ by taking Euclidean norm is a continuous function). Let M be the max and m be the min of P on $S_1(\mathbf{0})$ and $C = \max\{|M|, |m|\}$ then for any $(x, y, z) \in S_1(\mathbf{0})$ we have

  $$|P(x, y, z)| \leq C.$$

  Now take any nonzero $\boldsymbol{x} \in \mathbb{R}^3$ we have $\frac{\boldsymbol{x}}{||\boldsymbol{x}||} \in S_1(\mathbf{0})$, together with the fact that P is homogeneous,

  $$|P(\boldsymbol{x})| = ||\boldsymbol{x}||^5 |P(\boldsymbol{x}/||\boldsymbol{x}||)| \leq C||\boldsymbol{x}||^5.$$

- In general, we define a homogeneous function $f(\boldsymbol{x})$ on $\mathbb{R}^n$ of degree $a > 0$ being a function such that $f(t\boldsymbol{x}) = t^a f(\boldsymbol{x})$ holds for any nonzero $t \in \mathbb{R}$ and $\boldsymbol{x} \in \mathbb{R}^n$. For example linear operators are homogeneous functions of degree 1, the function
$$f(x, y, z) = (x^2y + yz^2 + z^3)^{1/5}$$
is homogeneous of degree 3/5 since we can check
$$f(tx, ty, tz) = t^{3/5} f(x, y, z).$$
Likewise, using the homogeneity together with the extreme value theorem on unit sphere we can derive the following result. If f is nonzero homogeneous of degree $a > 0$ then there exists a positive $C \in \mathbb{R}$ such that
$$|f(\boldsymbol{x})| \leq C||\boldsymbol{x}||^a$$
for any $\boldsymbol{x} \in \mathbb{R}^n$.
- Moreover, just as the positive definite quadratic forms, if we add extra condition that $f(\boldsymbol{x}) \geq 0$ and $f(\boldsymbol{x}) = 0$ iff $\boldsymbol{x} = \boldsymbol{0}$ (we actually only need $f(\boldsymbol{x}) = 0$ iff $\boldsymbol{x} = \boldsymbol{0}$ since f has IVP), then there exist positive $C_1, C_2 \in \mathbb{R}$ such that
$$C_1||\boldsymbol{x}||^a \leq |f(\boldsymbol{x})| \leq C_2||\boldsymbol{x}||^a$$
for all $\boldsymbol{x} \in \mathbb{R}^n$. This is because on the unit sphere $S_1(\boldsymbol{0})$(compact) the continuous homogeneous function f attains its max C_2 and min C_1, both positive. Then for any nonzero $\boldsymbol{x} \in \mathbb{R}^n$ we have
$$C_1||\boldsymbol{x}||^a \leq f(\boldsymbol{x}) = ||\boldsymbol{x}||^a f(\boldsymbol{x}/||\boldsymbol{x}||) \leq C_2||\boldsymbol{x}||^a.$$

10 February 10

The differentiability in multivariable calculus is different from that in one-variable. Roughly speaking, given a function $f : A \to R$, where $A \subset \mathbb{R}^n$ and $\boldsymbol{a} \in A^\circ$, when considering a multivariable function in terms of only one coordinate x_k, we get the so-called kth partial derivative when the limit
$$\frac{\partial f}{\partial x_k}(\boldsymbol{a}) := \lim_{t \to 0} \frac{f(\boldsymbol{a} + t\boldsymbol{e}_k) - f(\boldsymbol{a})}{t}$$
exists; but $\boldsymbol{x}$ can also move towards $\boldsymbol{a}$ along some specific direction, in this case we call the directional derivative of f at $\boldsymbol{a}$ in the direction $\boldsymbol{v}$ (with $||\boldsymbol{v}|| = 1$ as a unit vector) if the limit
$$D_{\boldsymbol{v}} f(\boldsymbol{a}) := \lim_{t \to 0} \frac{f(\boldsymbol{a} + t\boldsymbol{v}) - f(\boldsymbol{a})}{t}$$

exists. Thus one can easily see that partial derivative is just a special type of directional derivative along the coordinate direction:

$$\frac{\partial f}{\partial x_k}(\boldsymbol{a}) = D_{\boldsymbol{e}_k} f(\boldsymbol{a}).$$

Unlike that differentiability implies continuity in one-variable case, we do not have this implication in the multivariable setting:

Example 10.1. Consider the following two examples:

(1) Consider the function

$$f(x,y) = \begin{cases} 0 & \text{if } |y| \geq x^2 \text{ or } y = 0; \\ 1 & \text{otherwise.} \end{cases}$$

Let's first check that all directional derivatives at the origin $(0,0)$ exist and equal 0. Let $\boldsymbol{v} = (a,b)$ be any unit vector representing the direction, then

a) If $\boldsymbol{v}$ is along one of the axis directions, namely $(\pm 1, 0)$ or $(0, \pm 1)$, then for any $t \neq 0$, we have

$$\frac{f(ta,tb)}{t} = 0.$$

b) Otherwise, we have $\boldsymbol{v} = (a,b)$ where neither a nor b is 0. In this case, we take $t_0 = |b/a^2|$, then for any $t \neq 0$ with $|t| \leq t_0 = |b/a^2|$ we have

$$(ta)^2 = t(ta^2) \leq |tb|.$$

It then implies that for sufficiently small t, we again have

$$\frac{f(ta+tb)}{t} = 0.$$

Therefore $D_{\boldsymbol{v}} f(0,0) = 0$ for all unit vector $\boldsymbol{v}$. But the function f is not continuous at $(0,0)$ because if we let (x,y) approach to $(0,0)$ along the parabola $y = \frac{1}{2}x^2$, we have

$$\lim_{x\to 0} f(x, x^2/2) = 1 \neq f(0,0) = 0$$

since $|y| = x^2/2 < x^2$ when $x \neq 0$. This shows that f is not continuous at the origin but all directional derivative exist.

(2) Consider the function

$$f(x,y) = \begin{cases} \frac{xy^2}{x^2+y^4} & \text{if } (x,y) \neq (0,0); \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

Too see all directional derivatives at the origin exist, take any unit vector $\boldsymbol{v} = (a, b)$, then the directional derivative can be computed by

$$D_{\boldsymbol{v}} f(0,0) = \lim_{t\to 0} \frac{f(ta,tb) - f(0,0)}{t} = \lim_{t\to 0} \frac{ab^2}{a^2 + b^4t^2} = \begin{cases} b^2/a & \text{if } a \neq 0; \\ 0 & \text{if } a = 0. \end{cases}$$

But the function f is not continuous at $(0,0)$ since if approaching to $(0,0)$ along $x = y^2$ we have

$$\lim_{y\to 0} \frac{y^2y^2}{y^4 + y^4} = \frac{1}{2} \neq f(0,0) = 0.$$

Also, even if the partial derivatives exist it does not imply all directional derivatives exist. Here is a simple example:

Example 10.2. Consider the function

$$f(x,y) = \begin{cases} \sqrt{x} & \text{if } x = y; \\ 0 & \text{otherwise.} \end{cases}$$

It is easy to check that

$$\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = 0.$$

But if we take the direction $\boldsymbol{v}$ to be along $y = x$, then

$$D_{\boldsymbol{v}} f(0,0) = \lim_{t\to 0} \frac{f(t\boldsymbol{v})}{t} = \lim_{t\to 0} \frac{1}{\sqrt[4]{2}\sqrt{t}}$$

does not exists as a finite limit.

We say the function $\boldsymbol{f} : E \to \mathbb{R}^m$ where $E \subseteq \mathbb{R}^n$ is totally differentiable at $\boldsymbol{a} \in E^\circ$ if there exists a linear operator $L : \mathbb{R}^n \times \mathbb{R}^m$ interpreted as an $m \times n$ matrix, such that in $\mathbb{R}^m$ the limit

$$\lim_{\boldsymbol{h}\to\boldsymbol{0}} \frac{\boldsymbol{f}(\boldsymbol{a}+\boldsymbol{h}) - \boldsymbol{f}(\boldsymbol{a}) - L\boldsymbol{h}}{||\boldsymbol{h}||} = \boldsymbol{0}$$

exists. In this case we say the function $\boldsymbol{f}$ is differentiable at $\boldsymbol{a}$ and write $D\boldsymbol{f}(\boldsymbol{a}) = L$ the derivative of $\boldsymbol{f}$ at $\boldsymbol{a}$. Likewise, there is also a componentwise characterization of differentiability by looking at only the coordinate functions. The standard way to check if a function is totally differentiable is two-fold:

- check if all partial derivatives exist;
- if so, check whether the Jacobian matrix provides the local linearization.

Consider the following two examples first:

Example 10.3. (1) Consider the function given by

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

First check partial derivatives:

$$\frac{\partial f}{\partial x}(0,0) = \lim_{t\to 0} \frac{f(t,0)}{t} = 0 = \frac{\partial f}{\partial y}(0,0).$$

But

$$\lim_{(x,y)\to(0,0)} \frac{f(x,y) - f(0,0) - 0}{\sqrt{x^2+y^2}} \neq 0$$

since one can take the path approaching to $(0,0)$ to be $y = x$. Therefore the function is not differentiable at the origin.

(2) How to modify the function to make it totally differentiable? Make the total degree of the numerator greater than 1. For example consider

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

Similarly we get all partials exist and equal to zero. Moreover the limit

$$\lim_{(x,y)\to(0,0)} \frac{|xy|^{2/3}}{\sqrt{x^2+y^2}} = 0$$

since

$$\frac{|xy|^{2/3}}{\sqrt{x^2+y^2}} \leq (x^2+y^2)^{\frac{2}{3}-\frac{1}{2}} = (x^2+y^2)^{\frac{1}{6}} \to 0$$

as $(x,y) \to (0,0)$. Thus such function f is totally differentiable at $(0,0)$.

11 February 17

Before discussing more examples regarding the partial/total derivatives, let's first consider an example about continuous extension which we fail to cover before:

Example 11.1. Discuss the uniform continuity of function

$$f(x,y) = \frac{1}{1 + x(y-1)}$$

on the open balls $B_1 = B((0,1),1)$ and $B_2 = B((0,1),\sqrt{2})$.

The points where the function $f(x,y)$ is not defined at is along the curves $x(y-1) = -1$. Notice that the ball $B((0,1),1)$ is away from the singularities so let's try proving it is Lipschitz continuous on $B((0,1),1)$ first. For any $(x,y) \in B_1$ we have

$$2|x(y-1)| \leq x^2 + (y-1)^2 < 1$$

and thus $\frac{1}{2} < 1 + x(y-1) < \frac{3}{2}$. Now take any points $(x_1,y_1),(x_2,y_2) \in B_1$, we have

$$\begin{aligned}
|f(x_1,y_1) - f(x_2,y_2)| &= \frac{|x_2y_2 - x_2 - x_1y_1 + x_1|}{(1+x_1(y_1-1))(1+x_2(y_2-1))} \\
&\leq \frac{|x_2||y_2-y_1| + |y_1-1||x_1-x_2|}{(1+x_1(y_1-1))(1+x_2(y_2-1))} \\
&\leq 4|y_1-y_2| + 4|x_1-x_2| \\
&\leq 4\sqrt{2}||(x_1,y_1)-(x_2,y_2)||
\end{aligned}$$

which shows that the function is indeed Lipschitz on B_1 thus it has to be uniformly continuous on B_1.

In the larger ball $B_2 = B((0,1),\sqrt{2})$ we can check easily that $(1,0)$ is a boundary point of B_2, therefore by taking sequence

$$(x_n,y_n) = \left(1 - \frac{1}{n}, \frac{1}{n}\right) \to (1,0)$$

in B_2 one can show using the sequential characterization that the function cannot be continuously extended to the boundary point $(1,0)$. That is to say the function is not uniformly continuous on $B_2 = B((0,1),\sqrt{2})$.

Let's consider two more examples:

Example 11.2. (3) Discuss the differentiability of the function

$$f(x,y) = (1+x)\sqrt{|y|}$$

at $(-1,0)$.
Since $f(x,0) = f(-1,y) = 0$ for any $x,y \in \mathbb{R}$, it is easy to check that

$$\frac{\partial f}{\partial x}(-1,0) = \frac{\partial f}{\partial y}(-1,0) = 0.$$

Moreover, the limit

$$\lim_{(x,y)\to(-1,0)} \frac{f(x-1,y) - f(-1,0) - 0}{\sqrt{x^2+y^2}} = \lim_{(x,y)\to(0,0)} \frac{x\sqrt{|y|}}{\sqrt{x^2+y^2}} = 0$$

since

$$0 \leq \left| \frac{x\sqrt{|y|}}{\sqrt{x^2+y^2}} \right| \leq \frac{|x|\sqrt{|y|}}{|x|} = \sqrt{|y|} \to 0$$

as $(x,y) \to (0,0)$. Thus the function f is differentiable at $(-1,0)$.

(4) Discuss the differentiability of the function

$$f(x,y) = \begin{cases} x^2y^3\cos\dfrac{1}{x^4+y^4} + \sin(x+y^2) & \text{if } (x,y) \neq (0,0); \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$$

at $(0,0)$.
First of all all partials exist:

$$\lim_{t\to 0}\frac{f(t,0)-f(0,0)}{t} = \lim_{t\to 0}\frac{\sin t}{t} = 1 = \frac{\partial f}{\partial x}(0,0),$$
$$\lim_{t\to 0}\frac{f(0,t)-f(0,0)}{t} = \lim_{t\to 0}\frac{\sin t^2}{t} = 0 = \frac{\partial f}{\partial y}(0,0).$$

Moreover we can check that

$$\lim_{(x,y)\to(0,0)}\frac{f(x,y)-f(0,0)-x}{\sqrt{x^2+y^2}}$$
$$= \lim_{(x,y)\to(0,0)}\frac{x^2y^3\cos\frac{1}{x^4+y^4}+\sin(x+y^2)-x}{\sqrt{x^2+y^2}} = 0$$

since

$$\frac{|\sin(x+y^2)-x|}{\sqrt{x^2+y^2}} \leq \frac{|\sin x\cos y^2 - x|}{|x|} + \frac{|\cos x \sin y^2|}{|y|} \to 0$$

as $(x,y)\to(0,0)$. Thus the function is differentiable at $(0,0)$.

We also have the Chain Rule for composition of multivariable functions, which says that given $\boldsymbol{g} : A \to \mathbb{R}^p$ where $A \subseteq \mathbb{R}^n$, and $\boldsymbol{f} : B \to \mathbb{R}^m$ with $B \subseteq \mathbb{R}^p$ are two functions with $\boldsymbol{g}(A) \subseteq B$:

$$A \xrightarrow{\boldsymbol{g}} \boldsymbol{g}(A) \subseteq B \xrightarrow{\boldsymbol{f}} \mathbb{R}^m,$$

if $\boldsymbol{g}$ is differentiable at $\boldsymbol{a} \in A^\circ$ and $\boldsymbol{f}$ is differentiable at $\boldsymbol{b} = \boldsymbol{g}(\boldsymbol{a})$ then the composite function $\boldsymbol{f}\circ\boldsymbol{g}$ is differentiable at $\boldsymbol{a}$ with

$$D(\boldsymbol{f}\circ\boldsymbol{g})(\boldsymbol{a}) = D\boldsymbol{f}(\boldsymbol{b})D\boldsymbol{g}(\boldsymbol{a}).$$

In the matrix form, with $\boldsymbol{g} = (g_1,\cdots,g_p)$ and coordinate on $\mathbb{R}^n$ being $\boldsymbol{x} = (x_1,\cdots,x_n)$; $\boldsymbol{f} = (f_1,\cdots,f_m)$ and coordinate on $\mathbb{R}^p$ being $\boldsymbol{y} = (y_1,\cdots,y_p)$, we can write

$$D(\boldsymbol{f}\circ\boldsymbol{g})(\boldsymbol{a}) = \begin{bmatrix} \dfrac{\partial f_1}{\partial y_1}(\boldsymbol{b}) & \cdots & \dfrac{\partial f_1}{\partial y_p}(\boldsymbol{b}) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial y_1}(\boldsymbol{b}) & \cdots & \dfrac{\partial f_m}{\partial y_p}(\boldsymbol{b}) \end{bmatrix} \begin{bmatrix} \dfrac{\partial g_1}{\partial x_1}(\boldsymbol{a}) & \cdots & \dfrac{\partial g_1}{\partial x_n}(\boldsymbol{a}) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial g_p}{\partial x_1}(\boldsymbol{a}) & \cdots & \dfrac{\partial g_p}{\partial x_n}(\boldsymbol{a}) \end{bmatrix}.$$

A direct application using Chain Rule:

Example 11.3. Compute all partial derivatives at $(0,0)$ of $\boldsymbol{F} = \boldsymbol{f} \circ \boldsymbol{g}$ where

$$\begin{aligned}\boldsymbol{f}(y_1, y_2) &= ((y_1+1)\sqrt{|y_2|}, y_1^2 + y_2),\\ \boldsymbol{g}(x_1, x_2) &= (x_1 - 1, x_1 + 2x_2).\end{aligned}$$

For simplicity, let us denote $f_1(\boldsymbol{y}) = (y_1+1)\sqrt{|y_2|}$, $f_2(\boldsymbol{y}) = y_1^2 + y_2$ and $g_1(\boldsymbol{x}) = x_1 - 1$, $g_2(\boldsymbol{x}) = x_1 + 2x_2$, and $\boldsymbol{F} = (F_1, F_2)$.

First of all $\boldsymbol{g}$ is differentiable at $(0,0)$ and $\boldsymbol{f}$ is differentiable at $\boldsymbol{g}(0,0) = (-1,0)$ (see above). Componentwisely we have

$$\begin{aligned}\frac{\partial F_1}{\partial x_1}(\mathbf{0}) &= \frac{\partial f_1}{\partial y_1}(-1,0)\frac{\partial g_1}{\partial x_1}(0,0) + \frac{\partial f_1}{\partial y_2}(-1,0)\frac{\partial g_2}{\partial x_1}(0,0) = 0,\\ \frac{\partial F_1}{\partial x_2}(\mathbf{0}) &= \frac{\partial f_1}{\partial y_1}(-1,0)\frac{\partial g_1}{\partial x_2}(0,0) + \frac{\partial f_1}{\partial y_2}(-1,0)\frac{\partial g_2}{\partial x_2}(0,0) = 0,\\ \frac{\partial F_2}{\partial x_1}(\mathbf{0}) &= \frac{\partial f_2}{\partial y_1}(-1,0)\frac{\partial g_1}{\partial x_1}(0,0) + \frac{\partial f_2}{\partial y_2}(-1,0)\frac{\partial g_2}{\partial x_1}(0,0) = -1,\\ \frac{\partial F_2}{\partial x_2}(\mathbf{0}) &= \frac{\partial f_2}{\partial y_1}(-1,0)\frac{\partial g_1}{\partial x_2}(0,0) + \frac{\partial f_2}{\partial y_2}(-1,0)\frac{\partial g_2}{\partial x_2}(0,0) = 2.\end{aligned}$$

12 February 22

Here is a not-so-direct application of the Chain Rule:

Example 12.1. Consider a function $f : \mathbb{R}^2 \to \mathbb{R}$ which can be written as $f(x,y) = \varphi(x^2+y^2)$ for some function $\varphi : \mathbb{R}_{\geq 0} \to \mathbb{R}$, *i.e.*, sometimes we call such function a radial function, meaning that the restriction of f to any sphere of some fixed radius centered at the origin is a constant. Then

$$x\frac{\partial f}{\partial y} - y\frac{\partial f}{\partial x} = 0. \tag{$\star$}$$

This is not hard to see by applying the Chain Rule to $f(x,y) = \varphi(x^2+y^2)$:

$$\begin{aligned}\frac{\partial f}{\partial x}(x,y) &= \varphi'(x^2+y^2)(2x),\\ \frac{\partial f}{\partial y}(x,y) &= \varphi'(x^2+y^2)(2y)\end{aligned}$$

which imply the equality as desired.

Conversely, every function f satisfying $(\star)$ can be written as $\varphi(x^2+y^2)$ for some function $\varphi : \mathbb{R} \to \mathbb{R}$. In other words, every function satisfying $(\star)$ takes constant value on each circle centered at the origin. By using the polar coordinates, define

$$F(r,\theta) = f(r\cos\theta, r\sin\theta).$$

It is then equivalent to show that the function $F(r, \theta)$ only depends on r. So it is suffices to show that

$$\frac{\partial F}{\partial \theta}(r, \theta) = 0.$$

By the Chain Rule again, we have

$$\frac{\partial F}{\partial \theta}(r, \theta) = \frac{\partial f}{\partial x}(r\cos\theta, r\sin\theta)(-r\sin\theta) + \frac{\partial f}{\partial y}(r\cos\theta, r\sin\theta)(r\cos\theta) = 0.$$

Then it shows that f is constant on circles centered at the origin. Thus

$$f(x, y) = f(\sqrt{x^2 + y^2}, 0).$$

Therefore we can take $\varphi(r) = f(\sqrt{r}, 0)$.

Given a vector-valued function $\boldsymbol{f} : E \to \mathbb{R}^m$ with $E \subseteq \mathbb{R}^n$ we say that $\boldsymbol{f}$ is Lipschitz on E if there is a constant $M > 0$ such that for any $\boldsymbol{x}, \boldsymbol{y} \in E$,

$$||\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{y})|| \leq M||\boldsymbol{x} - \boldsymbol{y}||.$$

One of the corollaries as a direct application of Mean Value Theorem is to prove a certain differentiable function is Lipschitz on an open and *convex* set if all partials/total derivative is bounded over the set.

Remark 12.2. Note that the condition of convexity is required since to apply the Mean Value Theorem we need to guarantee that the line segments are contained in the set. A simple example one can consider is take a set which is not convex, say

$$E = \{(x, y) \in \mathbb{R}^2 \mid ||(x, y)|| \leq 1\} \setminus \{y = 0\}$$

which is a unit disk with the x-axis removed. Consider the sign function $f(x, y) = \operatorname{sgn}(y)$ on E. Then it is easy to check that

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

for any $(x, y) \in E$ but f is obviously not uniformly continuous on E since one can take sequences

$$(x_n, y_n) = \left(0, \frac{1}{n}\right), \quad (x'_n, y'_n) = \left(0, \frac{-1}{n}\right)$$

but $|f(x_n, y_n) - f(x'_n, y'_n)| = 2$. Thus f is not Lipschitz on E.

Consider the following example first:

Example 12.3. Given a function $f(x,y,z) = x^2y + z^3$ defined over any bounded set $K \subseteq \mathbb{R}^3$. Then f is Lipschitz on K.

There are two ways to see this. Firstly, since K is bounded we have $K \subseteq B(\mathbf{0}, \delta)$ for some $\delta > 0$. Then for any $(x,y,z), (a,b,c) \in K$,

$$\begin{aligned} |f(x,y,z) - f(a,b,c)| &= |x^2y + z^3 - a^2b - c^3| \\ &\leq y|x+a||x-a| + a^2|y-b| + |z-c||z^2+zc+c^2| \\ &\leq (2\delta^2 + \delta^2 + 3\delta^2)||(x,y,z) - (a,b,c)|| \end{aligned}$$

thus by definition f is Lipschitz on K.

Alternatively, consider the function f on $B(\mathbf{0}, \delta)$ as above. It is easy to compute the partials

$$\frac{\partial f}{\partial x}(x,y,z) = 2xy, \quad \frac{\partial f}{\partial y}(x,y,z) = x^2, \quad \frac{\partial f}{\partial z}(x,y,z) = 3z^2$$

all of which are bounded on $B(\mathbf{0}, \delta)$. Moreover the function f itself is differentiable on $B(\mathbf{0}, \delta)$ thus by the corollary in the lecture notes (right after the definition of Lipschitz), the function f is Lipschitz on $B(\mathbf{0}, \delta)$ and thus Lipschitz on K as well.

But this function f is not globally Lipschitz, *i.e.,* it is not Lipschitz over $\mathbb{R}^3$. One can use the sequential characterization, by taking sequences

$$(x_n, y_n, z_n) = (n, 1, 0)\,, \quad (x'_n, y'_n, z'_n) = \left(n + \frac{1}{n}, 1, 0\right)$$

but

$$|f(x_n, y_n, z_n) - f(x'_n, y'_n, z'_n)| = \left(n + \frac{1}{n}\right)^2 - n^2 = 2 + \frac{1}{n^2} \to 2 \neq 0.$$

Therefore it is not uniformly continuous on $\mathbb{R}^3$ thus it is not Lipschitz on $\mathbb{R}^3$.

For a not so obvious example, let us consider

Example 12.4. Consider the following function over different domains

$$f(x,y) = \begin{cases} x^2 e^{-y/x} & \text{if } x \neq 0; \\ 0 & \text{if } x = 0. \end{cases}$$

(1) Over the compact and convex set in the first quadrant

$$D_1 = \{(x, y \in \mathbb{R}^2 \mid 0 \leq y \leq x \leq 1)\}.$$

First check f is C^1 over D_1. Since if $x = 0$ then $y = 0$, we only need to check the existence and continuity of partials at $(0,0)$. For nonzero x, we have

$$\begin{aligned} \frac{\partial f}{\partial x}(x,y) &= (2x+y)e^{-y/x}, \\ \frac{\partial f}{\partial y}(x,y) &= -xe^{-y/x}. \end{aligned}$$

Moreover,

$$\frac{\partial f}{\partial x}(0,0) = \lim_{s\to 0}\frac{f(s,0)-f(0,0)}{s} = \lim_{s\to 0}\frac{s^2}{s} = 0,$$
$$\frac{\partial f}{\partial y}(0,0) = \lim_{t\to 0}\frac{f(0,t)-f(0,0)}{t} = 0.$$

And it can be checked that both partials are continuous at $(0,0)$ since

$$\lim_{(x,y)\to(0,0)}(2x+y)e^{-y/x} = \lim_{(x,y)\to(0,0)} xe^{-y/x} = 0$$

due to the fact that $-y/x < 0$ and thus $e^{-y/x} < 1$. By the corollary, with D_1 is compact, the function f is Lipschitz on D_1.

(2) Now consider the domain to be

$$D_2 = [-1,0]\times[0,1]\subset\mathbb{R}^2$$

which is a closed square in the second quadrant. Notice that on D_2 the exponent $-y/x \geq 0$ in particular when $x\to 0$ the power of e can be large enough to dominate the quadratic term x^2. For example along the path given by $y=\sqrt{-x}$, we have

$$f(x,\sqrt{-x}) = x^2e^{1/\sqrt{-x}} \to \infty$$

as $x\to 0^-$. Thus the function f is not uniformly continuous over D_2 therefore it cannot be Lipschitz over D_2.

13 February 24

Throughout the notes for today, we are assuming $\boldsymbol{f}: A\to\mathbb{R}^m$ where $A\subseteq\mathbb{R}^n$ is an open set of $\mathbb{R}^n$ and $\boldsymbol{a}\in A$. All the directional vectors are taken to be unit vector. Recall that given two directional vectors $\boldsymbol{v},\boldsymbol{w}\in\mathbb{R}^n$, let us assume first that $D_{\boldsymbol{v}}\boldsymbol{f}$ exists around $\boldsymbol{a}$, we can view it as a new function defined on some neighborhood of $\boldsymbol{a}$ and define its directional derivative along the direction $\boldsymbol{w}$. If the corresponding limit indeed exists, we get the notion of second order directional derivative at $\boldsymbol{a}$, which is denoted by

$$D^2_{\boldsymbol{w},\boldsymbol{v}}\boldsymbol{f}(\boldsymbol{a}) = D_{\boldsymbol{w}}(D_{\boldsymbol{v}}\boldsymbol{f}(\boldsymbol{a})).$$

If the function is "nice" enough, we can define higher order directional derivative by repeating the same process.

Some computations about higher order directional derivatives:

Example 13.1. Consider the function

$$f(x, y) = x + y^{4/3}$$

and directional vectors

$$\boldsymbol{v} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \quad \boldsymbol{w} = \left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right).$$

By direct computation we can compute the partial derivatives and the gradient is

$$\nabla f(x, y) = \left(1, \frac{4}{3}y^{1/3}\right),$$

and both partials are continuous thus the function is totally differentiable. In particular, we have the directional derivative along $\boldsymbol{v}$ can be computed as

$$D_{\boldsymbol{v}} f(x, y) = \nabla f(x, y) \cdot \boldsymbol{v}.$$

Now $D_{\boldsymbol{v}} f(x, y) = \frac{1}{\sqrt{2}} + \frac{4}{3\sqrt{2}} y^{1/3}$, and the (partial) derivatives of $D_{\boldsymbol{v}} f$ exist as long as $y \neq 0$ since when $y = 0$:

$$\lim_{y \to 0} \frac{(D_{\boldsymbol{v}} f)(x, y) - (D_{\boldsymbol{v}} f)(x, 0)}{y} = \lim_{y \to 0} \frac{4}{3\sqrt{2} y^{2/3}} = +\infty.$$

Thus for $y \neq 0$, we get

$$\nabla D_{\boldsymbol{v}} f(x, y) = \left(0, \frac{4}{9\sqrt{2}} y^{-2/3}\right)$$

and thus the second order directional derivative when $y \neq 0$ is

$$D^2_{\boldsymbol{w},\boldsymbol{v}} f(x, y) = \nabla D_{\boldsymbol{v}} f \cdot \boldsymbol{w} = \frac{\sqrt{3}}{9\sqrt{2}} y^{-2/3},$$

and it does not exist when $y = 0$.

Third order partial derivatives with mixed orders can be computed similarly, see the example below:

Example 13.2. Consider the function $f(x, y) = x^3 y^{1/3}$. By direct computation we get that

$$\frac{\partial f}{\partial x}(x, y) = 3x^2 y^{1/3}, \quad \frac{\partial f}{\partial y}(x, y) = \frac{1}{3} x^3 y^{-2/3}.$$

And the partial in terms of y does not exist when $y = 0$ since

$$\lim_{y \to 0} \frac{f(x, y) - f(x, 0)}{y} = \lim_{y \to 0} \frac{x^3}{y^{2/3}} = \infty.$$

For higher order we can compute when $y \neq 0$ that

$$\frac{\partial^2 f}{\partial x^2}(x,y) = 6xy^{1/3}, \quad \frac{\partial^2 f}{\partial y \partial x}(x,y) = \frac{\partial^2 f}{\partial x \partial y}(x,y) = x^2 y^{-2/3}.$$

Furthermore,

$$\frac{\partial^3 f}{\partial y \partial x^2} = \frac{\partial^3 f}{\partial x \partial y \partial x} = \frac{\partial^3 f}{\partial x^2 \partial y} = 2xy^{-2/3}.$$

In particular, at the origin $(0,0)$, consider $\frac{\partial^2 f}{\partial x^2}$, we have

$$\frac{\partial^3 f}{\partial y \partial x^2}(0,0) = \lim_{t \to 0} \frac{\frac{\partial^2 f}{\partial x^2}(0,t) - \frac{\partial^2 f}{\partial x^2}(0,0)}{t} = 0$$

but other third order partials do not exist at $(0,0)$.

It is then natural to ask the question when the following equality holds

$$D^2_{\boldsymbol{w},\boldsymbol{v}} \boldsymbol{f}(\boldsymbol{a}) = D^2_{\boldsymbol{v},\boldsymbol{w}} \boldsymbol{f}(\boldsymbol{a}).$$

There are two sets of conditions that will lead to the above equality (separately):

- (Schwartz) Both $D^2_{\boldsymbol{w},\boldsymbol{v}} \boldsymbol{f}$ and $D^2_{\boldsymbol{v},\boldsymbol{w}} \boldsymbol{f}$ exist around $\boldsymbol{a}$ and are continuous at $\boldsymbol{a}$;
- The function $\boldsymbol{f}$ is C^1 (all partials exists and continuous), and $D^2_{\boldsymbol{w},\boldsymbol{v}} \boldsymbol{f}$ exists around $\boldsymbol{a}$ and continuous at $\boldsymbol{a}$.

It is worth mentioning that there is no implications between these two sets of conditions, *i.e.*, none of which is stronger than the other. Moreover if we drop some conditions we may still have the equality holds.

Example 13.3. Consider the function

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

It can be computed that

$$\frac{\partial f}{\partial x}(x,y) = \begin{cases} \dfrac{2x(y^2-1)}{(x^2+y^2)^2} & \text{if } (x,y) \neq (0,0); \\ 0 & \text{if } (x,y) = (0,0), \end{cases}$$

$$\frac{\partial f}{\partial y}(x,y) = \begin{cases} \dfrac{2y(x^2-1)}{(x^2+y^2)^2} & \text{if } (x,y) \neq (0,0); \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

and partials are not continuous at $(0,0)$ therefore the function is not of class C^1 around $(0,0)$. We can also compute

$$\frac{\partial^2 f}{\partial y \partial x}(x,y) = \begin{cases} \dfrac{4xy - 4y(x^2+y^2)}{(x^2+y^2)^2} & \text{if } (x,y) \neq (0,0); \\ 0 & \text{if } (x,y = (0,0)), \end{cases}$$

$$\frac{\partial^2 f}{\partial x \partial y}(x,y) = \begin{cases} \dfrac{4xy - 4x(x^2+y^2)}{(x^2+y^2)^2} & \text{if } (x,y) \neq (0,0); \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

and they are both not continuous at $(0,0)$ but we still have

$$\frac{\partial^2 f}{\partial y \partial x}(0,0) = \frac{\partial^2 f}{\partial x \partial y}(0,0) = 0.$$

Another standard example:

Example 13.4. Consider the function

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

The function is C^1 since

$$\frac{\partial f}{\partial x}(x,y) = \begin{cases} \dfrac{x^4y + 4x^2y^3 - y^5}{(x^2+y^2)^2} & \text{if } (x,y) \neq (0,0); \\ 0 & \text{if } (x,y) = (0,0), \end{cases}$$

$$\frac{\partial f}{\partial y}(x,y) = \begin{cases} \dfrac{x^5 - 5x^3y^2 - xy^4}{(x^2+y^2)^2} & \text{if } (x,y) \neq (0,0); \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

And it can be checked that these two partials are continuous thus the function is C^1. But

$$\frac{\partial^2 f}{\partial x \partial y}(0,0) = \lim_{x \to 0} \frac{\frac{\partial f}{\partial y}(x,0) - 0}{x} = \lim_{x \to 0} \frac{x-0}{x} = 1,$$

$$\frac{\partial^2 f}{\partial y \partial x}(0,0) = \lim_{y \to 0} \frac{\frac{\partial f}{\partial x}(0,y) - 0}{y} = \lim_{y \to 0} \frac{-y-0}{y} = -1.$$

14 March 8

Let's finish up the Taylor's expansion for multivariable functions today.

Given a vector-valued function $\boldsymbol{f} : A \to \mathbb{R}^m$ where $A \subseteq \mathbb{R}^n$ is an open subset and $\boldsymbol{a} \in A$. Then

- (Peano form) If $\boldsymbol{f}$ is p times differentiable at $\boldsymbol{a}$, then

$$\boldsymbol{f}(\boldsymbol{a}+\boldsymbol{h}) = \sum_{k=0}^{p} \frac{1}{k!}\mathrm{d}^k \boldsymbol{f}(\boldsymbol{a},\boldsymbol{h}) + o(||\boldsymbol{h}||^p)$$

  where $\mathrm{d}^k\boldsymbol{f}(\boldsymbol{a},\boldsymbol{h})$ is the k-th differential of $\boldsymbol{f}$ at $\boldsymbol{a}$;
- (Lagrange form) If $\boldsymbol{f}$ is real-valued, $\boldsymbol{f}$ is $(p+1)$ times differentiable on A, and the line segment $L(\boldsymbol{a},\boldsymbol{a}+\boldsymbol{h}) \subset A$, then there exists $\theta \in (0,1)$ such that

$$f(\boldsymbol{a}+\boldsymbol{h}) = \sum_{k=0}^{p} \frac{1}{k!}\mathrm{d}^k f(\boldsymbol{a},\boldsymbol{h}) + \frac{1}{(p+1)!}\mathrm{d}^{p+1} f(\boldsymbol{a}+\theta\boldsymbol{h},\boldsymbol{h}).$$

Recall that the p-th differential of f at $\boldsymbol{a}$ is given by

$$\begin{aligned}
\mathrm{d}^p f(\boldsymbol{a},\boldsymbol{h}) := \mathrm{d}^p f(\boldsymbol{a},\boldsymbol{h},\cdots,\boldsymbol{h}) &= \sum_{j_1,\cdots,j_p\in[n]} \frac{\partial^p f}{\partial x_{j_p}\cdots\partial x_{j_1}}(\boldsymbol{a})h_{j_1}\cdots h_{j_p} \\
&= \sum \binom{p}{k_1,\cdots,k_n} \frac{\partial^p f}{\partial x_n^{k_n}\cdots\partial x_1^{k_1}}(\boldsymbol{a})h_1^{k_1}\cdots h_n^{k_n}
\end{aligned}$$

where the last summation is taken over $k_1,\cdots,k_n \in \{0,\cdots,p\}$ with $k_1+\cdots+k_n = p$, and the combinatorial factor is called the multinomial coefficients. We rewrite the summation by grouping with respect to taking partials in terms of x_i for k_i times. Namely,

$$\binom{p}{k_1,\cdots,k_n} = \frac{p!}{k_1!\cdots k_n!}.$$

It is

- the coefficients of the expansion $(x_1+\cdots+x_n)^p$;
- and has the combinatorial meaning as the number of ways of putting p distinct balls into n distinct bins with the i-th bin consisting of exactly k_i balls.

Note that some of the $k_1,\cdots,k_n$ can be zero. It is easier to estimate the remainder using the latter sum. In other words, if the $(p+1)$st order partials are all bounded by some $M > 0$ in some small neighborhood of $\boldsymbol{a}$ (containing the line segment $L(\boldsymbol{a},\boldsymbol{a}+\boldsymbol{h})$), then

$$\left|\frac{1}{(p+1)!}\mathrm{d}^{p+1} f(\boldsymbol{a}+\theta\boldsymbol{h},\boldsymbol{h})\right| \le \frac{M}{(p+1)!}||\boldsymbol{h}||^{p+1}\sum\binom{p+1}{k_1,\cdots,k_n} = \frac{M||\boldsymbol{h}||^{p+1}}{(p+1)!}n^{p+1}.$$

Remark 14.1. (1) By the uniqueness of the Taylor's expansion, if $f(\boldsymbol{x})$ is a polynomial function in $\boldsymbol{x} = (x_1, \cdots, x_n)$ of total degree d, since by grading in terms of degrees, we can canonically write

$$f(\boldsymbol{x}) = f_0(\boldsymbol{x}) + f_1(\boldsymbol{x}) + \cdots + f_d(\boldsymbol{x})$$

where $f_j(\boldsymbol{x})$ is the homogeneous polynomial of degree j. Then

$$\mathrm{d}^j f(\mathbf{0}, \boldsymbol{x}) = j! f_j(\boldsymbol{x})$$

is the j-th differential of f at $\mathbf{0}$.

(2) With the uniqueness of the Taylor's expansion, it sometimes will simplify the computation. As the example we have seen in class to find the Taylor's expansion at $(0, 0)$ of

$$f(x, y) = (x^2 + y) \sin y$$

we can actually use the 1-dimensional Taylor's formula for $\sin y$ and multiply with the polynomial term $(x^2 + y)$.

Example 14.2. i) Similarly, to find the 5th differential of

$$f(x, y) = (x^2 + y)e^x$$

at $(0, 0)$ we can write

$$\begin{aligned} f(x, y) &= (x^2 + y) \left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots\right) \\ &= y + (x^2 + xy) + \cdots + \left(\frac{x^5}{3!} + \frac{x^4 y}{4!}\right) + \cdots \end{aligned}$$

and $\mathrm{d}^5 f(\mathbf{0}, \boldsymbol{x}) = 5! \left(\frac{x^5}{3!} + \frac{x^4 y}{4!}\right)$.

ii) Consider

$$f(x, y) = e^{xy} \sin(x + y).$$

Since

$$e^{xy} = 1 + xy + \frac{(xy)^2}{2!} + R_2(x, y)$$

and

$$\sin(x + y) = (x + y) - \frac{(x + y)^3}{3!} + R_3(x, y)$$

by the uniqueness of expansion, the third degree Taylor polynomial of $f(x, y)$ at $(0, 0)$ is

$$\begin{aligned} P_3(x, y) &= (x + y) + xy(x + y) - \frac{(x + y)^3}{3!} \\ &= (x + y) + \frac{1}{3!}(-x^3 + 3x^2 y + 3xy^2 - y^3). \end{aligned}$$

We can even conclude from the expansion that

$$\frac{\partial^3 f}{\partial x^3}(\mathbf{0}) = \frac{\partial^3 f}{\partial y^3}(\mathbf{0}) = -1, \quad \frac{\partial^3 f}{\partial x^2 \partial y}(\mathbf{0}) = \frac{\partial^3 f}{\partial x \partial y^2}(\mathbf{0}) = 1.$$

Now let us consider one estimation problem using Taylor's expansion.

Example 14.3. Consider an old example

$$f(x, y) = (x^2 + y)\sin y = (y^2) + (x^2 y) + \left(-\frac{y^4}{3!}\right) + \left(-\frac{x^2 y^3}{3!}\right) + \cdots$$

on the unit square $D_1 = [0, 1] \times [0, 1]$. For simplicity, we denote $(\theta x, \theta y) = (x_0, y_0)$ throughout.

- If using the least order of terms to estimate, there exists some $\theta \in (0, 1)$ such that

$$\begin{aligned} f(x, y) &= d^1 f((x_0, y_0), (x, y)) = \nabla f(x_0, y_0) \cdot (x, y) \\ &= (2x_0 \sin y_0)x + (\sin y_0 + (x_0^2 + y_0)\cos y_0)y. \end{aligned}$$

  Over D_1, this gives a not so good estimation

$$|f(x, y)| \le 2|x| + 3|y|.$$

- Now consider the order one estimation, we get for some $\theta \in (0, 1)$ such that

$$\begin{aligned} f(x, y) &= \frac{1}{2}\mathrm{d}^2 f((x_0, y_0), (x, y)) \\ &= \frac{1}{2}\left[(2\sin y_0)x^2 + 2(2x_0 \cos y_0)xy + (2\cos y_0 - (x_0^2 + y_0)\sin y_0)y^2\right]. \end{aligned}$$

  Thus on D_1, we have

$$\begin{aligned} f(x, y) &\le \frac{1}{2}\left[2xy^2 + 4x^2 y + (2 + (x^2 + y)y^3)\right] \\ &\le y^2 + 3x^2 y + \frac{1}{2}y^4 + \frac{1}{2}x^2 y^3. \end{aligned}$$

  If restricting on $D_2 = \{(x, y) \mid 0 \le x \le y \le 1\} \subseteq \mathbb{R}^2$, then

$$|f(x, y) - y^2| \le 3y^3 + \frac{1}{2}y^4 + \frac{1}{2}y^5 \le 4y^3.$$

- But if using $\sin y = y - \frac{y^3}{6}\cos(\theta y)$ for some $\theta \in (0, 1)$, we get

$$f(x, y) = (x^2 + y)\left(y - \frac{y^3}{6}\cos(\theta y)\right) = y^2 + x^2 y - (y^4 + x^2 y^3)\frac{\cos(\theta y)}{6}.$$

Thus on D_2 the estimation turns out to be

$$\begin{aligned} |f(x,y) - y^2| &= \left| x^2 y - (y^4 + x^2 y^3) \frac{\cos(\theta y)}{6} \right| \\ &\leq y^3 + \frac{y^4 + y^5}{6} \leq \frac{4}{3} y^3 \end{aligned}$$

which is a better estimation comparing to $4y^3$.

15 March 10

One application of the Taylor's expansion is to give necessary/sufficient conditions for a point being local min/max.

Recall that given a real-valued function $f : A \to \mathbb{R}$ where $A \subseteq \mathbb{R}^n$, a point $\boldsymbol{a} \in A$ is called a critical point of f is when f is differentiable at $\boldsymbol{a}$ with vanishing gradient, *i.e.*, $\nabla f(\boldsymbol{a}) = \boldsymbol{0}$, or equivalently, all partials vanish at $\boldsymbol{a}$, or equivalently, the first order differential $\mathrm{d}f(\boldsymbol{a}, \boldsymbol{h}) = 0$ for all vector $\boldsymbol{h} \in \mathbb{R}^n$.

For a differentiable function, the critical points are the points which might be a local extremum or a saddle point (a critical point which is not a local extremum). Let us first recall the sufficient condition for local extrema/saddle points:

Theorem 15.1. *Suppose that f is k times differentiable at $\boldsymbol{a}$ and*

$$\mathrm{d}f(\boldsymbol{a}, \boldsymbol{h}) = \mathrm{d}^2 f(\boldsymbol{a}, \boldsymbol{h}) = \cdots = \mathrm{d}^{k-1} f(\boldsymbol{a}, \boldsymbol{h}) = 0$$

for all $\boldsymbol{h} \in \mathbb{R}^n$ and $\mathrm{d}^k f(\boldsymbol{a}, \boldsymbol{h}_0) \neq 0$ for some $\boldsymbol{h}_0 \in \mathbb{R}^n$. If $k \geq 3$ is odd then $\boldsymbol{a}$ is a saddle point. If k is even then

- $\mathrm{d}^k f(\boldsymbol{a}, \boldsymbol{h}) > 0$ *for all $\boldsymbol{h} \neq \boldsymbol{0}$ implies $\boldsymbol{a}$ is a strict min;*
- $\mathrm{d}^k f(\boldsymbol{a}, \boldsymbol{h}) < 0$ *for all $\boldsymbol{h} \neq \boldsymbol{0}$ implies $\boldsymbol{a}$ is a strict max;*
- $\mathrm{d}^k f(\boldsymbol{a}, \boldsymbol{h})$ *changes sign implies $\boldsymbol{a}$ is a saddle point.*

The situations not included are called ambiguous cases, meaning that the result is inconclusive from above theorem.

Here are some toy examples:

Example 15.2.

- The examples we have seen in class $f_1(x,y) = x^2 + y^4$ and $f_2(x,y) = x^2 - y^4$. They both have the origin $(0,0)$ as the only critical point, and both have the same degree 2 Taylor's polynomial x^2, and moreover, they both have vanishing Hessian at $(0,0)$ which leads to the ambiguous case. It is easy to see that $f_1(x,y)$ is always nonnegative, thus it has (absolute) min at $(0,0)$. But $f_2(x,y)$ has a saddle point at $(0,0)$ by considering two paths to $(0,0)$ parametrized by (t,t) and (t^3,t).

- (Monkey saddle – three up/three down) The example $f(x,y) = x^3 - 3xy^2$. Again we have the gradient $\nabla f = (3x^2 - 3y^2, -6xy)$ thus the only critical point is the origin $(0,0)$. We can check that

$$\mathrm{d}f(\mathbf{0}, \boldsymbol{h}) = \mathrm{d}^2 f(\mathbf{0}, \boldsymbol{h}) = 0$$

for any $\boldsymbol{h} \in \mathbb{R}^2$ and $\mathrm{d}^3 f(\mathbf{0}, \boldsymbol{h})$ is not constantly zero. Thus the origin is a saddle point.
- (Dog saddle) Consider the function defined by $f(x,y) = x^3y - xy^3$. The gradient can be computed as $\nabla f = (3x^2y - y^3, x^3 - 3xy^2)$ and one can check similarly that the only critical point is the origin $(0,0)$. This is also a saddle point by taking paths parametrized by $(t, t/2)$ and $(t, -t/2)$.

Now consider the following examples using the sufficient condition:

Example 15.3. (1) Consider the function

$$f(x,y) = x^2 y e^{-x^2-y^2}.$$

First find the critical points by setting all partials to be zero:

$$\begin{aligned} \frac{\partial f}{\partial x} &= 2xye^{-x^2-y^2} - x^2y2xe^{-x^2-y^2} = 2xy(1-x^2)e^{-x^2-y^2} = 0 \\ \frac{\partial f}{\partial y} &= (1-2y^2)x^2e^{-x^2-y^2} = 0 \end{aligned}$$

Thus the critical points are

(i) $(0,b)$ where $b \in \mathbb{R}$: the entire y-axis. When $b > 0$, near $(0,b)$, we have $f(x,y) \geq 0$ thus $(0,b)$ is a local min. When $b < 0$, near $(0,b)$, we have $f(x,y) \leq 0$ thus $(0,b)$ is a local max. When $b = 0$, $f(x,y)$ changes sign near the origin, thus it is a saddle point.

(ii) $(\pm 1, \pm 1/\sqrt{2})$ are local extrema, to see this one can check the second differential at those points.

(2) Consider the function

$$f(x,y) = x^4 + y^4 - (x+y)^2.$$

Firstly set partials to zero to find the critical points:

$$\begin{aligned} \frac{\partial f}{\partial x} &= 4x^3 - 2(x+y) = 0 \\ \frac{\partial f}{\partial y} &= 4y^3 - 2(x+y) = 0. \end{aligned}$$

Then the critical points are

(i) $(0,0)$. Near $(0,0)$, along $x=y$, we have $f(x,x)=2x^4-4x^2<0$; along $x=-y$ we have $f(x,-x)=2x^4>0$. Thus the origin is a saddle point.

(ii) $(1,1)$. Compute the second differential, since

$$\frac{\partial^2 f}{\partial x^2}=12x^2-2,\qquad \frac{\partial^2 f}{\partial x\partial y}=\frac{\partial^2 f}{\partial y\partial x}=-2,\qquad \frac{\partial^2 f}{\partial y^2}=12y^2-2,$$

we have

$$\begin{aligned}\mathrm{d}^2 f((1,1),(h_1,h_2))&=10h_1^2-4h_1h_2+10h_2^2\\&=2(h_1-h_2)^2+8(h_1^2+h_2^2)>0\end{aligned}$$

for any nonzero $(h_1,h_2)\neq(0,0)$. Thus $(1,1)$ is a local min.

(iii) $(-1,-1)$. This is similar to $(1,1)$.

(3) Discuss that whether $(0,0)$ is a saddle point for the function

$$f(x,y)=x^2y^5+x^4y^4.$$

First check whether $(0,0)$ is a critical point:

$$\begin{aligned}\frac{\partial f}{\partial x}(0,0)&=2xy^5+4x^3y^4=0,\\ \frac{\partial f}{\partial y}(0,0)&=5x^2y^4+4x^4y^3=0.\end{aligned}$$

Moreover $f(0,0)=0$ with

$$f(x,y)=x^2y^4(y+x^2).$$

The term x^2y^4 is nonnegative, and $y+x^2$ changes sign in some small neighborhood of the origin. Thus $(0,0)$ cannot be a local extremum, and it must be a saddle point.

The following examples show the ambiguous case.

Example 15.4. (1) First consider the function

$$f(x,y)=(y-x)^2-x^4.$$

By setting partials to be zero, we get

$$\begin{aligned}\frac{\partial f}{\partial x}&=2(x-y)-4x^3=0\\ \frac{\partial f}{\partial y}&=2(y-x)=0\end{aligned}$$

which yield $(0,0)$ being the only critical point. Then since

$$\frac{\partial^2 f}{\partial x^2} = 2 - 12x^2, \quad \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = -2, \quad \frac{\partial^2 f}{\partial y^2} = 2$$

the second differential is $\mathrm{d}^2 f(\mathbf{0}, \boldsymbol{h}) = 2h_1^2 - 4h_1h_2 + 2h_2^2 = 2(h_1 - h_2)^2 \geq 0$. Or one can check the Hessian at origin has vanishing determinant. Thus this is the ambiguous case. Along the path parametrized by (t,t) we have

$$f(t,t) = -t^4 < 0$$

near 0; along the path parametrized by $(t, 2t)$ we have

$$f(t, 2t) = t^2 - t^4 > 0$$

near 0. Together with the fact that $f(0,0) = 0$, the origin is a saddle point.

(2) Now consider the function

$$f(x,y) = (y-x)^2 + x^4 + y^5,$$

and it is easy to see the origin $(0,0)$ is a critical point. Now we want to check whether $(0,0)$ is a local extremum or a saddle point. Again, the second differential

$$\mathrm{d}^2 f(\mathbf{0}, \boldsymbol{h}) = 2(h_1 - h_2)^2 \geq 0$$

for any $\boldsymbol{h} \in \mathbb{R}^2$. Thus it also falls into an ambiguous case. We set $t = y - x$ and thus $y = x + t$ and consider

$$g(x,t) = t^2 + x^4 + (x+t)^5$$

at the origin $(0,0)$. The natural guess is that near the origin $t^2 + x^4$ is the dominant terms, thus we should have $(0,0)$ to be a local min. And we only need to consider locally when $x + t < 0$ since by taking fifth power, the term $(x+t)^5$ will have negative contribution. Note that

$$\begin{aligned} x &\leq |x| \leq (x^4 + t^2)^{1/4}, \\ t &\leq |t| \leq (x^4 + t^2)^{1/2}. \end{aligned}$$

Now we will consider (x,t) in $B(\mathbf{0}, \epsilon)$ for $\epsilon > 0$ sufficiently small. Let $z = x^4 + t^2$, then in such neighborhood z is sufficiently small, and thus

$$z^{1/5} - z^{1/4} - z^{1/2} = z^{1/5}(1 - z^{5/4} - z^{5/2}) > 0.$$

It means that

$$x + t \leq |x| + |t| \leq (x^4 + t^2)^{1/4} + (x^4 + t^2)^{1/2} < (x^4 + t^2)^{1/5}$$

which implies

$$|x+t|^5 < x^4 + t^2.$$

Therefore in a sufficiently small neighborhood of the origin,

$$t^2 + x^4 - |(x+t)^5| \geq (t^2 + x^4) - (x^4 + t^2) = 0.$$

In particular, back to the original coordinate (x, y) from the current (x, t) we only need to restrict to a smaller ball, for example $B(\mathbf{0}, \epsilon/2)$. Since when $(x, y) \in B(\mathbf{0}, \epsilon/2)$, it guarantees that $|t| = |x - y| < \epsilon$ as well. It shows that $(0, 0)$ is indeed a local min.

16 March 17

A special case for the above sufficient condition is when we have two variables, *i.e.*, $n = 2$, and $\mathrm{d}^2(\boldsymbol{a}, \boldsymbol{h}_0) \neq 0$ for some $\boldsymbol{h}_0 \in \mathbb{R}^2$, *i.e.*, $k = 2$. In this case, the sufficient condition can be reinterpreted in the following way. Let $\boldsymbol{h}_0 = (h_1, h_2)$ and $\boldsymbol{a}$ be a critical point, then the second differential can be realized as a quadratic form in terms of h_1, h_2:

$$d^2(\boldsymbol{a}, \boldsymbol{h}_0) = Q_f(\boldsymbol{a}, \boldsymbol{h}_0) = Ah_1^2 + 2Bh_1h_2 + Ch_2^2$$

where

$$A = \frac{\partial^2 f}{\partial x^2}(\boldsymbol{a}), \quad B = \frac{\partial^2 f}{\partial x \partial y}(\boldsymbol{a}) = \frac{\partial^2 f}{\partial y \partial x}(\boldsymbol{a}), \quad C = \frac{\partial^2 f}{\partial y^2}(\boldsymbol{a})$$

and the Hessian matrix of f at $\boldsymbol{a}$ is precisely given by

$$H_f(\boldsymbol{a}) = \begin{pmatrix} A & B \\ B & C \end{pmatrix}.$$

Thus

(1) If $Q_f(\boldsymbol{a}, \boldsymbol{h}_0)$ is indefinite, *i.e.*, $\det H_f(\boldsymbol{a}) < 0$ then $\boldsymbol{a}$ is a saddle point;
(2) If $Q_f(\boldsymbol{a}, \boldsymbol{h}_0)$ is positive definite, then $\boldsymbol{a}$ is a strict local min. Or one can check if $A > 0$ and $\det H_f(\boldsymbol{a}) > 0$, then it also implies $\boldsymbol{a}$ is a local min;
(3) If $Q_f(\boldsymbol{a}, \boldsymbol{h}_0)$ is negative definite, then $\boldsymbol{a}$ is a strict local max. Or one can check if $A < 0$ and $\det H_f(\boldsymbol{a}) > 0$, then it also implies $\boldsymbol{a}$ is a local max.

Note that this is nothing different from the original sufficient condition, we only use different terminology to describe the sign of $\mathrm{d}^2 f(\boldsymbol{a}, \boldsymbol{h})$. Here is an example:

Example 16.1. Consider the function

$$f(x, y) = (1 + e^y) \cos x - ye^y.$$

Find all the critical points, and determine whether they are local extrema or saddle point.

First set all partials equaling zero:

$$\frac{\partial f}{\partial x} = -(1+e^y)\sin x = 0, \quad \frac{\partial f}{\partial y} = e^y(\cos x - 1 - y) = 0.$$

Thus the critical points are (i) $(2k\pi, 0)$ and (ii) $((2k+1)\pi, -2)$. Now let us compute all the second partials:

$$\begin{aligned}
\frac{\partial^2 f}{\partial x^2} &= -(1+e^y)\cos x, \\
\frac{\partial^2 f}{\partial x \partial y} &= \frac{\partial^2 f}{\partial y \partial x} = -e^y \sin x, \\
\frac{\partial^2 f}{\partial y^2} &= e^y(\cos x - 2 - y).
\end{aligned}$$

It can be checked easily that (I am going to use the same notation A, B, C as above for simplicity):

- At $\boldsymbol{a} = (2k\pi, 0)$:
$$A = -2, \quad B = 0, \quad C = -1,$$
or in other words
$$\mathrm{d}^2 f(\boldsymbol{a}, (h,k)) = -2h^2 - k^2 < 0$$
for nonzero $(h,k) \neq \mathbf{0}$. Thus $(2k\pi, 0)$ are local max.
- At $\boldsymbol{a} = ((2k+1)\pi, -2)$:
$$A = 1 + e^{-2}, \quad B = 0, \quad C = -e^{-2},$$
and $\det H_f(\boldsymbol{a}) = -e^{-2}(1+e^{-2}) < 0$. Thus $((2k+1)\pi, -2)$ are saddle points.

Recall that

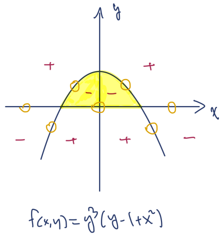
Proposition 16.2. *Let $D \subset \mathbb{R}^n$ be closed and bounded, and $f : D \to \mathbb{R}$ be differentiable on D°. Then the absolute extrema of f are attained at either a critical point of f or a boundary point of the domain D.*

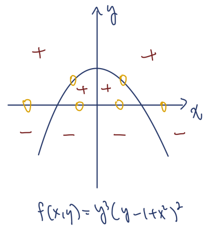
In practice, to locate absolute extrema over a compact set for a "nice" multivariable function, we first find all critical points. And find the boundary ∂D of its domain D. For example if the ambient space is $\mathbb{R}^2$, the boundary of a region D in $\mathbb{R}^2$ is just some 1-dimensional pieces, so that we can apply the method of finding extremal values in one variable calculus. Finally we compare the evaluation of f at those points and determine the greatest and least values among those.

Let us start with an easy example:

Example 16.3. Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x,y) = y^3(y - 1 + x^2).$$





Let's find the critical points first. Set

$$\frac{\partial f}{\partial x} = 2xy^3 = 0, \quad \frac{\partial f}{\partial y} = y^2(4y - 3 + 3x^2) = 0.$$

Then the critical points are (i) all the points on x-axis, *i.e.*, $y = 0$; (2) point $(0, 3/4)$. All the points on the x-axis are saddle points. But if we restrict the domain to the compact set

$$D = \{(x, y) \mid 0 \le y \le 1 - x^2\},$$

then all the points on the boundary ∂D are absolute max. Since f is continuous on the compact set D, it must attain its absolute min and max on D. We already know f attains its max on the boundary, so it attains its min at some interior point of D. And the point where f attains its min has to be a critical point, therefore $(0, 3/4)$ is an absolute min of f on D.

Now let us consider the function to be

$$f(x, y) = y^3(y - 1 + x^2)^2.$$

The graph has been shown as above. In this case the points lying on the parabola are also critical points. Since

$$\frac{\partial f}{\partial x} = 4xy^3(y - 1 - x^2) = 0, \quad \frac{\partial f}{\partial y} = y^2(y - 1 + x^2)(5y - 3 + 3x^2) = 0$$

the critical points are (i) points on x-axis; (ii) points on the parabola given by $y = 1 - x^2$; (iii) $(0, 3/5)$. In this case, all the points on the x-axis are saddle points; points on parabola with $y > 0$ are local min; points on parabola with $y < 0$ are local max. Restrict to the compact region D, the points on the boundary are absolute min, and the critical point $(0, 3/5)$ is an absolute max, by the same argument as above.

A slightly complicated example is

Example 16.4. Consider the function $f : D \to \mathbb{R}$ defined by

$$f(x, y) = 4xy - 2x^2 - y^4$$

on the compact domain

$$D = [-2, 2] \times [-2, 2].$$

Since f is continuous defined on compact set D it attains its max and min on D. First find interior points where the function f has vanishing partials:

$$\frac{\partial f}{\partial x} = 4y - 4x = 0, \quad \frac{\partial f}{\partial y} = 4x - 4y^3 = 0.$$

The critical points are $(0, 0)$, $(1, 1)$ and $(-1, -1)$. Restricting to the boundary, we only need to consider four one-variable functions

$$f(x, 2), \quad f(x, -2), \quad f(2, y), \quad f(-2, y).$$

I will only do $f(x, 2)$ here and the other three follow similarly. Consider

$$f(x, 2) = 8x - 2x^2 - 16, \quad x \in [-2, 2].$$

Since $8x - 2x^2 - 16 = -2(x^2 - 4x + 4) - 8 = -2(x - 2)^2 + 8$, $f(x, 2)$ attains its max at $x = 2$ and min at $x = -2$. Thus we need also take $(2, 2)$ and $(-2, 2)$ into account. For $f(2, y) = 8y - 8 - y^4$ on $y \in [-2, 2]$, we can also find the critical point by setting $8 - 4y^3 = 0$ thus we have the only critical point $y = \sqrt[3]{2}$. Together with the boundary point of $[-2, 2]$ we need to compare $f(2, 2)$, $f(2, -2)$ and $f(2, \sqrt[3]{2})$ to get the absolute extrema for $f(2, y)$ on $y \in [-2, 2]$. One can check that $f(2, y)$ attains its max at $y = \sqrt[3]{2}$ and its min at $y = -2$. Completing the discussion we could get the absolute max and min of $f(x, y)$ on the closed square D.

For more general constraints, one may need to apply Lagrange multiplier which we are not going to cover here (maybe later).

17 March 22 & 23

Now we move on to Inverse and Implicit Function Theorem. The Inverse Function Theorem tells us that under which circumstances a function is locally a diffeomorphism. Namely, recall that

Theorem 17.1. *Let $\boldsymbol{f} : V \to \mathbb{R}^n$ where $V \subset \mathbb{R}^n$ is open and $\boldsymbol{f} \in C^k(V)$ $(1 \leq k \leq \infty)$. Let $\boldsymbol{a} \in V$ with $\boldsymbol{f}(\boldsymbol{a}) = \boldsymbol{b}$ and $D\boldsymbol{f}(\boldsymbol{a})$ is invertible,* i.e., $\det D\boldsymbol{f}(\boldsymbol{a}) \neq 0$, *then*

- *there exists an open set $W \subset V$ containing $\boldsymbol{a}$ such that $\boldsymbol{f}$ is one-to-one on W and $\boldsymbol{f}(W)$ is open in $\mathbb{R}^n$;*
- *$\boldsymbol{f}^{-1} : \boldsymbol{f}(W) \to W$ is also C^k on $\boldsymbol{f}(W)$,*

where the above two mean $\boldsymbol{f} : W \to \boldsymbol{f}(W)$ is a C^k diffeomorphism, and moreover,

- *$D\boldsymbol{f}^{-1}(\boldsymbol{b}) = D\boldsymbol{f}(\boldsymbol{a})^{-1}$.*

Note that in the lecture we proved the theorem by constructing a contraction map and apply the contraction principle. But it can also be proved by applying Implicit Function Theorem to $F : V \times \mathbb{R}^n \to \mathbb{R}^n$ defined by $F(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{y} - \boldsymbol{f}(\boldsymbol{x})$ at $(\boldsymbol{a}, \boldsymbol{b})$ where $\boldsymbol{f}(\boldsymbol{a}) = \boldsymbol{b}$ in order to solve variables $\boldsymbol{x}$ in terms of $\boldsymbol{y}$.

The theorem stated as above is only a local result. Even if we have $D\boldsymbol{f}(\boldsymbol{x})$ invertible everywhere, the result is still local, *i.e.,* it does NOT mean that $\boldsymbol{f}$ has a global inverse on the entire domain V. We have already seen one classical example in class, namely $f(x, y) = (e^x \cos y, e^x \sin y)$. There are other examples such as polar coordinate transformation. The other essential condition for using Inverse Function Theorem is the dimension of the ambient space of the domain and codomain must be the same since otherwise we cannot talk about the invertibility of the Jacobian.

In order to get a global version of the Inverse Function Theorem, we can add conditions on the function $\boldsymbol{f}$ being one-to-one on its domain V and $\det D\boldsymbol{f}(\boldsymbol{x}) \neq 0$ for any $\boldsymbol{x} \in V$.

Note that the Jacobian matrix is a linear operator on tangent spaces. We can define what is called an **immersion**: Given a differentible function $\boldsymbol{f} : U \to V$ where $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$ are open sets, we call $\boldsymbol{f}$ an immersion if the rank of the Jacobian matrix is constant n, *i.e.,* $D\boldsymbol{f}$ has full column rank everywhere, which is equivalent to say that that the linear map $D\boldsymbol{f}$ is everywhere injective on U.

Example 17.2. Give a C^k function $\boldsymbol{f} : U \to V$. Assume $\boldsymbol{f}$ is a bijection and also an immersion, then we immediately get the relation on the dimension: $n \leq m$. Moreover $n < m$ cannot happen, since if it happens, for any $\boldsymbol{a} \in V$ since $\operatorname{rank} D\boldsymbol{f}(\boldsymbol{a}) = n$ the Jacobian $D\boldsymbol{f}(\boldsymbol{a})$ must contain an invertible $n \times n$ minor, say

$$\det \frac{\partial(y_{i_1}, \cdots, y_{i_n})}{\partial(x_1, \cdots, x_n)} \neq 0.$$

By the Inverse Function Theorem, there exists some open neighborhood of $\boldsymbol{a}$ such that $x_1, \cdots, x_n$ can be locally considered as C^k functions in terms of $y_{i_1}, \cdots, y_{i_n}$. But since $n < m$ there are different points $(y_1, \cdots, y_m)$ corresponding to the same point $(x_1, \cdots, x_n)$ which contradict the assumption that $\boldsymbol{f}$ being one-to-one. Thus we must have $n = m$.

Let's do some computation:

Example 17.3. Consider the function $\boldsymbol{f} : \mathbb{R}^2_{x,y} \to \mathbb{R}^2_{u,v}$ given by

$$f(x, y) = (x^2 - y^2, 2xy)$$

where we write $u = f_1(x, y) = x^2 - y^2$ and $v = f_2(x, y) = 2xy$.

(1) Show that there exists a neighborhood U of $(2, 1) \in \mathbb{R}^2$ such that $\boldsymbol{f} : U \to \boldsymbol{f}(U)$ has an inverse $\boldsymbol{f}^{-1} : \boldsymbol{f}(U) \to U$.

First compute the Jacobian of $\boldsymbol{f}(x,y)$:

$$D\boldsymbol{f}(x,y) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y} \\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y} \end{pmatrix} = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix}.$$

And thus

$$D\boldsymbol{f}(2,1) = \begin{pmatrix} 4 & -2 \\ 2 & 4 \end{pmatrix}, \quad \det D\boldsymbol{f}(2,1) \neq 0.$$

By Inverse Function Theorem, there exists some open neighborhood U of $(2,1)$ on which $\boldsymbol{f}$ has an inverse.

(2) Compute $g'(1)$ where $g(t) = h(\boldsymbol{f}^{-1}(3t^2, 4t))$ with $\boldsymbol{f}^{-1}$ is in (1) and $h(x,y) = xy^2$.
By Inverse Function Theorem, locally we have

$$\boldsymbol{f}^{-1}(u,v) = (F(u,v), G(u,v))$$

i.e., locally $x = F(u,v), y = G(u,v)$ and we are solving x, y in terms of u, v, with Jacobian

$$D\boldsymbol{f}^{-1}(u,v) = D\boldsymbol{f}(x,y)^{-1} = \frac{1}{4(x^2+y^2)} \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix}^T = \begin{pmatrix} \dfrac{\partial F}{\partial u} & \dfrac{\partial F}{\partial v} \\ \dfrac{\partial G}{\partial u} & \dfrac{\partial G}{\partial v} \end{pmatrix}.$$

It implies that

$$\frac{\partial F}{\partial u} = \frac{x}{2(x^2+y^2)}, \frac{\partial F}{\partial v} = \frac{y}{2(x^2+y^2)}, \frac{\partial G}{\partial u} = \frac{-y}{2(x^2+y^2)}, \frac{\partial G}{\partial v} = \frac{x}{2(x^2+y^2)}.$$

Now we can compute the derivative $g'(t)$ with $g(t) = h(\boldsymbol{f}^{-1}(3t^2, 4t))$:

$$\begin{aligned} g'(t) = {} & \frac{\partial h}{\partial x}(\boldsymbol{f}^{-1}(3t^2,4t)) \left(\frac{\partial F}{\partial u}(3t^2,4t)6t + \frac{\partial F}{\partial v}(3t^2,4t)4 \right) \\ & + \frac{\partial h}{\partial y}(\boldsymbol{f}^{-1}(3t^2,4t)) \left(\frac{\partial G}{\partial u}(3t^2,4t)6t + \frac{\partial G}{\partial v}(3t^2,4t)4 \right). \end{aligned}$$

Now set $t = 1$ then $\boldsymbol{f}^{-1}(3,4) = (2,1)$ because $\boldsymbol{f}(1,2) = (3,4)$. Notice that we only have formula for these partials in terms of x, y (not u, v). Thus

$$\begin{aligned} g'(1) = {} & (y^2)|_{(2,1)} \left[\frac{x}{2(x^2+y^2)} \cdot 6 + \frac{y}{2(x^2+y^2)} \cdot 4 \right]_{(2,1)} \\ & + (2xy)|_{(2,1)} \left[\frac{-y}{2(x^2+y^2)} \cdot 6 + \frac{x}{2(x^2+y^2)} \cdot 4 \right]_{(2,1)} = \frac{12}{5}. \end{aligned}$$

The Implicit Function Theorem tells us that when the variables $\boldsymbol{y}$ is (locally) implicitly defined by an equation $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y})$ in terms of $\boldsymbol{x}$. Geometrically, suppose we have a hypersurface defined by $z = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y})$, take a level set by setting $z = 0$, we want to know around a given point $(\boldsymbol{a}, \boldsymbol{b})$, whether there exists a function $\boldsymbol{g}(\boldsymbol{x})$ such that $\boldsymbol{y} = \boldsymbol{g}(\boldsymbol{x})$ solves the equation $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y}) = 0$, *i.e.*, the graph of $\boldsymbol{g}$ agrees with the zero set of $\boldsymbol{f}$ locally near $(\boldsymbol{a}, \boldsymbol{b})$.

Before getting into the standard local version of the Implicit Function Theorem, we have seen one of the global versions in class,

Theorem 17.4. *Suppose that $f : A \times I \to \mathbb{R}$ where $A \subset \mathbb{R}^n$ is open and $I = (b_1, b_2) \subset \mathbb{R}$ is an interval. For any fixed $\boldsymbol{x} \in A$, suppose*

- *the one-variable function $f(\boldsymbol{x}, -) : y \to f(\boldsymbol{x}, y)$ is differentiable with respect to y and $\frac{\partial f}{\partial y}(\boldsymbol{x}, y) > 0$ for all $y \in I$, and*
- $\lim_{y \to b_1^+} f(\boldsymbol{x}, y) < 0$ *while* $\lim_{y \to b_2^-} f(\boldsymbol{x}, y) > 0$,

then there exists a unique function $g : A \to I$ such that $y = g(\boldsymbol{x})$ is implicitly defined by $f(\boldsymbol{x}, y) = 0$.

The key point for proving this global theorem is to apply the Intermediate Value Theorem to the continuous function $f(\boldsymbol{x}, -)$. The strict monotonicity together with the limit condition guarantee that there exists a uniquely y associated to any fixed $\boldsymbol{x}$ to make $f(\boldsymbol{x}, y) = 0$ and such correspondence precisely constructs the function $g(\boldsymbol{x})$ as we want. Therefore we can use the same idea, to get another global version,

Theorem 17.5. *Suppose that $f : A \times I \to \mathbb{R}$ where $A \subset \mathbb{R}^n$ is open and $I = (b_1, b_2) \subset \mathbb{R}$ is an interval. For any fixed $\boldsymbol{x} \in A$, suppose there is a positive $c > 0$ such that*

$$\frac{\partial f}{\partial y}(\boldsymbol{x}, y) \geq c$$

for any $y \in I$, then there is a unique function $g : A \to I$ such that $y = g(\boldsymbol{x})$ is implicitly defined by $f(\boldsymbol{x}, y) = 0$.

This is a direct consequence of the following lemma in one-variable calculus:

Lemma 17.6. *Suppose that $f : \mathbb{R} \to \mathbb{R}$ is differentiable and there exists a positive real number c such that $f'(x) \geq c$ for any $x \in \mathbb{R}$, then f has a unique root in $\mathbb{R}$.*

Proof. First notice that f is strictly increasing, thus it has at most one root. Consider the value $f(0)$. If $f(0) = 0$ then we are done, and $x = 0$ has to be its unique root. If $f(0) > 0$, apply the MVT to f on the interval $[-f(0)/c, 0]$, then there exists some t in between such that

$$f(0) - f(-f(0)/c) = f'(t)\frac{f(0)}{c} \geq f(0)$$

which implies that $f(-f(0)/c) \leq 0$, meaning that there is at least a root in $[-f(0)/c, 0]$. Likewise we can prove for $f(0) < 0$. $\square$

Let us see one example about the global implicit function theorem. There are several more examples with analysis on the extreme values of the implicit function in the lecture notes and homework.

Example 17.7. Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x, y) = x^2y^3 + 2y - \cos y.$$

For any fixed $x \in \mathbb{R}$, view $f(x, y)$ as a function in terms of y, then

$$\frac{\partial f}{\partial y}(x, y) = 3x^2y^2 + 2 + \sin y \geq 1$$

for any $y \in \mathbb{R}$. Thus there exists a unique $y \in \mathbb{R}$ with $f(x, y) = 0$ for the fixed x. It then defines a function $g : \mathbb{R} \to \mathbb{R}$ with $y = g(x)$ implicitly defined by $f(x, y) = 0$. We can further compute

$$y' = g'(x) = \frac{-2xy^3}{3x^2y^2 + 2 + \sin y}$$

it has only one critical point at $x = 0$. This is because $y = g(x) > 0$ for any $x \in \mathbb{R}$. Let us now prove this claim. Firstly, the implicit function $y = g(x)$ is symmetric about $x = 0$, thus we only need to consider $x \geq 0$. Secondly, $y = g(x) \neq 0$ for any $x \in \mathbb{R}$ since

$$f(x, 0) = -\cos 0 \neq 0.$$

Since $g(x)$ is continuous it is either positive or negative by Intermediate Value Theorem. Moreover, if $x = 0$, then $y_0 = g(0)$ must satisfy

$$f(0, y_0) = 2y_0 - \cos y_0 = 0$$

thus $y_0 = g(0) > 0$. Therefore $y = g(x) > 0$ for all $x \in \mathbb{R}$. Furthermore, $g(x)$ attains its absolute max at $x = 0$ because

$$x^2 = \frac{\cos y - 2y}{y^3} \geq 0$$

meaning that $\cos y - 2y \geq 0$ which implies that $y = g(x) < y_0 = g(0)$ for any $x \in \mathbb{R}$.

Another example goes:

Example 17.8. Consider $f(x, y) = x^4 + y^3 + x^2 + y^2 + 1$. Then there exists a unique $y = g(x)$ implicitly defined by $f(x, y) = 0$ on $\mathbb{R} \times (-\infty, -2/3)$. But there is no such g implicitly defined on $\mathbb{R} \times (0, \infty)$.

For any fixed $x \in \mathbb{R}$, we have

$$\frac{\partial f}{\partial y}(x, y) = 3y^2 + 2y = y(3y + 2) > 0$$

for $y \in (-\infty, -2/3)$. It is also easy to see that $f(x,y)$ is continuous in y for any fixed $x \in \mathbb{R}$. We also have

$$\lim_{y \to -\infty} f(x,y) = -\infty, \qquad \lim_{y \to -\frac{2}{3}^{-1}} f(x,y) = (x^4 + x^2 + 1) + y^2(y+1)|_{-2/3} > 0$$

hold for any fixed $x \in \mathbb{R}$. By the global implicit function theorem there exists a unique $g(x)$ such that $y = g(x)$ is implicitly defined by $f(x,y) = 0$ on $\mathbb{R} \times (-\infty, -2/3)$.

On $\mathbb{R} \times (0, \infty)$, by way of contradiction, if there exists such function $g(x)$, since again for any fixed $x \in \mathbb{R}$

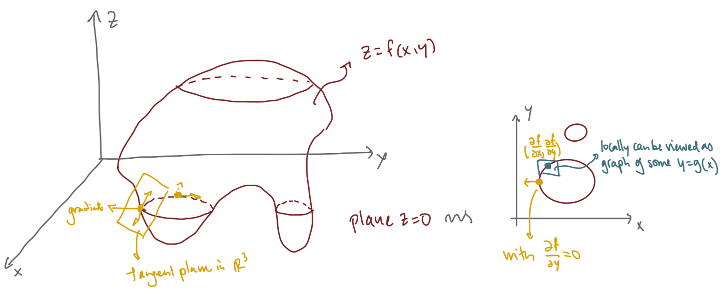
$$\frac{\partial f}{\partial y} = y(3y+2) > 0$$

for any $y > 0$. It means that $f(x,y)$ is strictly increasing in terms of y for any fixed x. Moreover,

$$\lim_{y \to 0+} f(x,y) = x^4 + x^2 + 1 > 0, \qquad \lim_{y \to \infty} f(x,y) = \infty.$$

Together with the fact that $y \mapsto f(x,y)$ is continuous on $(0, \infty)$, we cannot have $f(x,y) = 0$ on $\mathbb{R} \times (0, \infty)$.

For the local Implicit Function Theorem, we can roughly interpret as following:



for instance, consider a surface $z = f(x,y)$ in $\mathbb{R}^3$ with the level set cut out by the hyperplane $z = 0$. The theorem tells us that at which point, in the xy-plane, can the curve locally be viewed as a graph of some function $y = g(x)$. It is to see that locally the curve cannot have vertical tangent line, which is precise when the gradient direction has zero y contribution, *i.e.*, $\frac{\partial f}{\partial y} = 0$.

18 April 5

The local version of implicit function theorem is stated as follows:

Theorem 18.1. *Suppose $\boldsymbol{f} : \Omega \to \mathbb{R}^m$ where $\Omega \subset \mathbb{R}^n \times \mathbb{R}^m$ is an open subset is of class C^1 on Ω. And $(\boldsymbol{x}_0, \boldsymbol{y}_0) \in \Omega$ satisfies $\boldsymbol{f}(\boldsymbol{x}_0, \boldsymbol{y}_0) = 0$ and $D_{\boldsymbol{y}}\boldsymbol{f}(\boldsymbol{x}_0, \boldsymbol{y}_0)$ is invertible. Then there exists an open neighborhood $U \subset \mathbb{R}^n$ of x_0 and a unique function $\boldsymbol{g} : U \to \mathbb{R}^m$ which is implicitly defined by $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y}) = 0$ such that $\boldsymbol{y}_0 = \boldsymbol{g}(\boldsymbol{x}_0)$. The function $\boldsymbol{g} \in C^1(U)$ and*

$$D\boldsymbol{g}(\boldsymbol{x}) = -(D_{\boldsymbol{y}}\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{g}(\boldsymbol{x})))^{-1} D_{\boldsymbol{x}}\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{g}(\boldsymbol{x}))$$

for any $\boldsymbol{x} \in U$.

Let's see some examples:

Example 18.2. Consider a function $f : \mathbb{R}^3 \to \mathbb{R}$ given by

$$f(x, y, z) = z + x^2 + xz + y^3 + z^3.$$

(1) There exists a function $g : \mathbb{R}^2 \to \mathbb{R}$ such that $z = g(x, y)$ is implicitly defined by $f(x, y, z) = 0$ which is of class C^1 in some neighborhood of $(0, 0) \in \mathbb{R}^2$. First of all

$$\frac{\partial f}{\partial x} = 2x + 2, \quad \frac{\partial f}{\partial y} = 3y^2, \quad \frac{\partial f}{\partial z} = 1 + x + 3z^2$$

all of which are continuous on $\mathbb{R}^3$ thus the function f is of class C^1. Also notice that $f(0, 0, 0) = 0$ now since

$$D_z f(0, 0, 0) = \frac{\partial f}{\partial z}(0, 0, 0) = 1 > 0$$

by implicit function theorem we get the existence of such $g : \mathbb{R}^2 \to \mathbb{R}$ which is of class C^1 in some neighborhood of $(0, 0)$.

(2) We can also compute the partial derivative of the implicit function $g(x, y)$. For example, near $(0, 0)$ we have $f(x, y, g(x, y)) = 0$ thus by Chain rule to differentiate with respect to x:

$$\frac{\partial f}{\partial x}(x, y, g(x, y)) + \frac{\partial f}{\partial z}(x, y, g(x, y))\frac{\partial g}{\partial x}(x, y) = 0$$

together with $g(0, 0) = 0$ we get

$$\frac{\partial g}{\partial x}(0, 0) = 0.$$

Differentiate with respect to x one more time, we get

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial z \partial x}\frac{\partial g}{\partial x} + \left(\frac{\partial^2 f}{\partial x \partial z} + \frac{\partial^2 f}{\partial z^2}\frac{\partial g}{\partial x}\right)\frac{\partial g}{\partial x} + \frac{\partial f}{\partial z}\frac{\partial^2 g}{\partial x^2} = 0$$

for which we can compute $\frac{\partial^2 g}{\partial x^2}(0, 0) = -2$.

(3) But the origin $(0,0)$ is not a local extremum for g.
Consider the sign of $g(0,y)$ around the origin. The implicit function theorem says that $f(0,y,z)=0$ with $z=g(0,y)$ around the origin, then

$$f(0,y,z)=z+y^3+z^3=0$$

and thus $z(z^2+1)=-y^3$. Thus $z=g(0,y)$ changes signs around $y=0$ because $z^2+1>0$ and the sign of y^3 changes around $y=0$. Thus $(0,0)$ cannot be a local extremum for g. (It is a critical point since one can check it has zero gradient, so it has to be a saddle point.)

Consider another example:

Example 18.3. Given an equation

$$F(x,y)=x^2+y+\sin(xy)=0$$

(1) In some sufficiently small neighborhood of $(0,0)$, there is a unique continuous function $y=g(x)$ which is implicitly defined by the equation such that $g(0)=0$. This is simply due to the fact that

$$\frac{\partial F}{\partial y}(0,0)=1+x\cos(xy)|_{(0,0)}=1\neq 0$$

and $F(0,0)=0$. Thus by the implicit function theorem, locally near the origin there is a unique function $y=g(x)$ implicitly determined by the equation and $0=g(0)$. Moreover since F_x' and F_y' are both continuous the implicit function $y=g(x)$ has to be C^1 as well. And we can compute the derivative of the function g:

$$g'(x)=-\frac{F_x'}{F_y'}=\frac{2x+y\cos(xy)}{1+x\cos(xy)}.$$

(2) Now let us discuss the monotonicity of the continuous implicit function $y=g(x)$. Consider the sign of $g'(x)$. Near the origin, the denominator

$$1+x\cos(xy)\geq 1-|x|>0.$$

Thus the monotonicity depends on the sign of $-(2x+y\cos(xy))$. As $g(0)=0$, we have $g'(0)=0$. This means that $y=o(x)$ as $x\to 0$. Therefore $|y\cos(xy)|\leq |y|=o(x)$ near $x=0$ and thus the sign of $g'(x)$ is the same as $-2x$. In conclusion, if $x>0$ then $g'(x)<0$ thus $y=g(x)$ is decreasing; if $x<0$ then $g'(x)>0$ thus $y=g(x)$ is increasing. So the implicit function $y=g(x)$ attains its absolute maximum at $x=0$.

(3) In general there will be no way to see if x can be explicitly written as a function in terms of y which is implicitly defined by $F=0$ near the origin $(0,0)$. This is because

$$\frac{\partial F}{\partial x}(0,0)=(2x+y\cos(xy))_{(0,0)}=0.$$

But from the result in (2), we know that the function $y = g(x)$ attains its absolute max at $x = 0$ thus near $(0, 0)$, when $y < 0$ there are at least two x with $F(x, y) = 0$; but when $y > 0$ there is no such x such that $F(x, y) = 0$. Therefore near $(0, 0)$, there is no function $x = h(y)$ which is uniquely and implicitly defined by $F(x, y) = 0$ with $F(0, 0) = 0$.

19 April 6

One more example:

Example 19.1. Verify the system

$$\begin{cases} e^z + 3x - \cos y + y = 0 \\ e^x - x - z + y - 1 = 0 \end{cases}$$

defines implicitly around $(0, 0, 0)$ a curve in $\mathbb{R}^3$ with parametrized equations $x = t, y = y(t)$ and $z = z(t)$. Find the equation of the tangent line at $(0, 0, 0)$.

Let $f_1(x, y, z) = e^z + 3x - \cos y + y$ and $f_2(x, y, z) = e^x - x - z + y - 1$ and $\boldsymbol{f} = (f_1, f_2)$. Note that $\boldsymbol{f}(0, 0, 0) = (f_1(0, 0, 0), f_2(0, 0, 0)) = (0, 0)$. Moreover

$$\frac{\partial f_1}{\partial x} = 3, \frac{\partial f_1}{\partial y} = \sin y + 1, \frac{\partial f_1}{\partial z} = e^z, \frac{\partial f_2}{\partial x} = e^x - 1, \frac{\partial f_2}{\partial y} = 1, \frac{\partial f_2}{\partial z} = -1$$

which imply $\boldsymbol{f}$ is C^1 and

$$D_{y,z}\boldsymbol{f} = \begin{pmatrix} \dfrac{\partial f_1}{\partial y} & \dfrac{\partial f_1}{\partial z} \\ \dfrac{\partial f_2}{\partial y} & \dfrac{\partial f_2}{\partial z} \end{pmatrix} = \begin{pmatrix} \sin y + 1 & e^z \\ 1 & -1 \end{pmatrix}.$$

It is easy to check that $D_{y,z}\boldsymbol{f}(0, 0, 0)$ is invertible. Thus by implicit function theorem there are functions $y = y(x)$ and $z = z(x)$ which are uniquely and implicitly defined by the system $\boldsymbol{f}(0, 0, 0) = 0$ with $y(0) = 0, z(0) = 0$. Therefore locally near $(0, 0, 0)$, the system has a parametrization $(t, y(t), z(t))$.

To compute the tangent line at $(0, 0, 0)$ we only need to compute $y'(0)$ and $z'(0)$. Since

$$(D_{y,z}\boldsymbol{f}(0, 0, 0))^{-1} = \frac{1}{-2}\begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix}$$

and

$$D_x\boldsymbol{f}(0, 0, 0) = \begin{pmatrix} 3 \\ e^x - 1 \end{pmatrix}_{x=0} = \begin{pmatrix} 3 \\ 0 \end{pmatrix},$$

we have

$$\begin{pmatrix} y'(0) \\ z'(0) \end{pmatrix} = \frac{1}{2}\begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 3 \\ 0 \end{pmatrix} = \begin{pmatrix} -3/2 \\ -3/2 \end{pmatrix}.$$

Thus the tangent line at $(0,0,0)$ is given by $x = t, y = -3t/2, z = -3t/2$.

Alternatively, since the tangent line lies on the tangent planes of the given two surfaces, we only need to compute the intersection of two tangent planes at $(0,0,0)$. Since the normal vectors are

$$(3, \sin y + 1, e^z)_{(0,0,0)} = (3,1,1), \quad (e^x - 1, 1, -1)_{(0,0,0)} = (0,1,-1),$$

the tangent line is given by the solution of the linear system

$$\begin{cases} 3x + y + z = 0 \\ y - z = 0 \end{cases}$$

which is given by $(t, -3t/2, -3t/2)$.

Another application of implicit function theorem is to find constrained extrema, *i.e.,* Lagrange multiplier. The question is formulated in the following: Given $\boldsymbol{f} : A \to \mathbb{R}$ where $A \subset \mathbb{R}^n$ is an open set, and $\boldsymbol{h} : \mathbb{R}^n \to \mathbb{R}^m$. Find the extrema of $\boldsymbol{f}$ subject to the constraints $\boldsymbol{h}(\boldsymbol{x}) = \mathbf{0}$.

In some special cases, we don't need to use the Lagrange Multiplier. For example

Example 19.2. Find the extrema of $f(x,y) = e^{x^2-y^2}$ where (x,y) satisfies the relation $x^2 + y^2 = 1$.

In this case, under the relation, *i.e.,* the points are chosen on the unit circle, we can write $y^2 = 1 - x^2$. Thus $e^{x^2-y^2} = e^{2x^2-1}$ attains its min at $x = 0$ (then $y = \pm 1$), and max at $x = \pm 1$ (then $y = 0$).

Now let us consider a simple case: Find the extreme of $f(x,y)$ over A subject to the relation $h(x,y) = 0$. Suppose that $\nabla h \neq \mathbf{0}$ everywhere in A, *i.e.,* $\operatorname{rank} \nabla h = 1$, and suppose the point (x_0, y_0) is an extremum of $f(x,y)$ with $h(x_0,y_0) = 0$. Since $\nabla h(x_0,y_0) \neq \mathbf{0}$ then by the implicit function theorem, there exists a unique function $y = y(x)$ around x_0 which is implicitly defined by $h(x,y) = 0$. Precisely, we have $h(x, y(x)) = 0$ for $x \in (x_0 - \delta, x_0 + \delta)$ for some $\delta > 0$. Differentiate with respect to x we get

$$\frac{\partial h}{\partial x}(x_0,y_0) + \frac{\partial h}{\partial y}(x_0,y_0)y'(x_0) = 0$$

and since (x_0,y_0) is an extremum of f thus

$$\frac{\partial f}{\partial x}(x_0,y_0) + \frac{\partial f}{\partial y}(x_0,y_0)y'(x_0) = 0.$$

It implies that the following linear system with indeterminants α, β:

$$\begin{cases} \dfrac{\partial h}{\partial x}(x_0,y_0)\alpha + \dfrac{\partial h}{\partial y}(x_0,y_0)\beta = 0 \\ \dfrac{\partial f}{\partial x}(x_0,y_0)\alpha + \dfrac{\partial f}{\partial y}(x_0,y_0)\beta = 0 \end{cases}$$

has a nontrivial solution $(1, y'(x_0))$. In other words, the coefficient matrix must have zero determinant, *i.e.*, $\nabla f(x_0, y_0) = \lambda \nabla h(x_0, y_0)$, which means that (x_0, y_0) is a critical point of $f - \lambda h$.

Above discussion gives a necessary condition for constrained extrema, just as the Fermat theorem for the non-constrained case. It requires further discussion on whether the critical points we obtained are indeed extrema values. Redo the example above, first of all let $h(x, y) = x^2 + y^2 - 1$ and thus $\nabla h = (2x, 2y) \neq 0$ on the unit circle. Now consider

$$F(x, y) = f(x, y) - \lambda h(x, y) = e^{x^2 - y^2} - \lambda(x^2 + y^2 - 1).$$

To find critical points of F, setting $\partial_x F = \partial_y F = 0$ we get

$$\begin{cases} 2xe^{x^2-y^2} - 2\lambda x = 0, \\ 2ye^{x^2-y^2} + 2\lambda y = 0. \end{cases}$$

The points should also be on the unit circle, thus if $x = 0$ then $y = \pm 1$, in this case $\lambda = -e^{-1}$; if $y = 0$ then $x = \pm 1$, in this case $\lambda = e$. Thus possible candidates for the extrema are $(0, \pm 1)$ with $f(0, \pm 1) = e^{-1}$ and $(\pm 1, 0)$ with $f(\pm 1, 0) = e$.

20 April 12

The Lagrange Multiplier Theorem (necessary condition for constrained extrema) allows us to determine the absolute extrema of a multivariate real-valued function $f(\boldsymbol{x})$ subject to the constraint $g(\boldsymbol{x}) = 0$. We introduce a new variable λ and try to find solutions to

$$\nabla f(\boldsymbol{x}) - \lambda \nabla g(\boldsymbol{x}) = 0, \quad g(\boldsymbol{x}) = 0.$$

If we can ensure that f indeed has an absolute extremum on the vanishing set of g (for instance the zero set of g is closed and bounded, and f is continuous), then the absolute extremum of f is also a local extremum and it is attained at the solution $\boldsymbol{x}_0$ of above system for which $\nabla g(\boldsymbol{x}_0) \neq \mathbf{0}$ or at a point where the hypothesis does not hold. Thus we need to check both the solutions to the system, and some exceptional points in the zero set of g at which either ∇f or ∇g does not exist or at which $\nabla g = 0$.

Example 20.1. Given a quadratic form

$$f(x_1, x_2, x_3) = \boldsymbol{x} A \boldsymbol{x}^T = a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + 2a_{12}x_1x_2 + 2a_{23}x_2x_3 + 2a_{13}x_1x_3$$

determined by some 3×3 symmetric matrix $A = (a_{ij})$. Find the extrema on the sphere $h(\boldsymbol{x}) = x_1^2 + x_2^2 + x_3^2 - 1 = 0$.

The function f must attains its max and min because S^2 is compact in $\mathbb{R}^3$. Consider the function

$$F(\boldsymbol{x}) = f(\boldsymbol{x}) - \lambda(x_1^2 + x_2^2 + x_3^2 - 1)$$

which has critical points on S^2 (thus nonzero) as solutions of the following linear system:

$$\begin{cases} (a_{11} - \lambda)x_1 + a_{12}x_2 + a_{13}x_3 = 0, \\ a_{21}x_1 + (a_{22} - \lambda)x_2 + a_{23}x_3 = 0, \\ a_{31}x_1 + a_{32}x_2 + (a_{33} - \lambda)x_3 = 0, \\ x_1^2 + x_2^2 + x_3^2 = 1. \end{cases}$$

In order to make the system have nontrivial solution the coefficient matrix should have zero determinant. It is easy to see that the determinant is just the characteristic polynomial and the solutions to the polynomial is the eigenvalues of the matrix A. Thus for any eigenvalue λ of A, the linear system has a nonzero solution, which can be rescaled to be on S^2. Let (x_1, x_2, x_3) be such a solution, with respect to the eigenvalue λ. Then by multiplying the equations by x_1, x_2, x_3 respectively and summing them up, we get

$$f(\boldsymbol{x}) - \lambda(x_1^2 + x_2^2 + x_3^2) = f(\boldsymbol{x}) - \lambda = 0$$

thus $f(x_1, x_2, x_3) = \lambda$. In other words, the max of f is the max eigenvalue of A while the min of f is the min eigenvalue of A.

The next example show that condition $\nabla g \neq \mathbf{0}$ cannot be dropped in the Lagrange Multiplier Theorem.

Example 20.2. Take a cuspidal cubic curve defined by $y^2 - (x-1)^3 = 0$. We want to find the minimal distance from the origin $(0,0)$ to the point on the cuspidal cubic. In other words, we want to find the minimum value of the function $f(x,y) = x^2 + y^2$ with constraint $g(x,y) = y^2 - (x-1)^3 = 0$. It is easy to see from the picture that the minimum is attained at the cusp $(1,0)$ with shortest distance $\sqrt{f(1,0)} = 1$. But we can compute

$$\nabla f(1,0) = (2,0), \quad \nabla g(1,0) = (0,0).$$

Therefore we don't have $\nabla f(1,0)$ and $\nabla g(1,0)$ are collinear. Thus we can see from this example that the conclusion may not hold if we drop the nonvanishing of the gradient of g from the Lagrange Multiplier Theorem.

The method can also be used to prove some classical inequalities.

Example 20.3. Find the constrained minimum for the function

$$f(x,y,z) = x \ln x + y \ln y + z \ln z$$

subject to the constraint $g(x, y, z) = x + y + z - 1$.

First, the domain for the function f is $x > 0, y > 0, z > 0$. Together with the constraint $x + y + z = 1$, we should look for the minimum over the set

$$E = \{(x, y, z) \in \mathbb{R}^3 \mid 0 < x, y, z < 1\}.$$

Moreover, it is easy to see that $f(x, y, z) < 0$ and $\sup_E f(x, y, z) = 0$ since

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

In order to find the constrained min, using the Lagrange Multiplier. Define

$$L(x, y, z, \lambda) = (x \ln x + y \ln y + z \ln z) - \lambda(x + y + z - 1).$$

Thus we should search for the solutions to the following system of equations:

$$\begin{cases} \ln x + 1 - \lambda = 0 \\ \ln y + 1 - \lambda = 0 \\ \ln z + 1 - \lambda = 0 \\ x + y + z - 1 = 0. \end{cases}$$

The only solution to this system is

$$x = y = z = \frac{1}{3},$$

and this is where the function f attains its absolute min with constraint $g = 0$, which is $f(1/3, 1/3, 1/3) = -\ln 3$. Therefore we have on the hyperplane given by $x + y + z = 1$,

$$x \ln x + y \ln y + z \ln z \geq -\ln 3.$$

This can also be deduced from the Jensen's inequality, by considering the convex function $h(x) = -\ln x$, since $x + y + z = 1$, we have

$$h\left(x\frac{1}{x} + y\frac{1}{y} + z\frac{1}{z}\right) \leq xh\left(\frac{1}{x}\right) + yh\left(\frac{1}{y}\right) + zh\left(\frac{1}{z}\right)$$

which is exactly what we have obtained from Lagrange Multiplier. And it can be easily generalized to n variables.

21 April 14

Let us consider one more example related to multiple constraints.

Example 21.1. Let us find the point on intersection of two hyperplanes given by $x + y + z = 1$ and $3x + 2y + z = 6$ which has the shortest distance to the origin $(0, 0, 0)$. Denote

$$\begin{aligned} f(x, y, z) &= x^2 + y^2 + z^2, \\ h_1(x, y, z) &= x + y + z - 1, \\ h_2(x, y, z,) &= 3x + 2y + z - 6. \end{aligned}$$

We want to find the absolute min of f subject to the constraints $h_1 = h_2 = 0$. Consider the function

$$L(x, y, z, \lambda, \mu) = f - \lambda h_1 - \mu h_2,$$

and we need to solve the system $\nabla L = 0, h_1 = h_2 = 0$, more precisely.

$$\begin{cases} 2x - \lambda - 3\mu = 0 \\ 2y - \lambda - 2\mu = 0 \\ 2z - \lambda - \mu = 0 \\ x + y + z = 1 \\ 3x + 2y + z = 6. \end{cases}$$

It yields

$$x = \frac{\lambda + 3\mu}{2}, y = \frac{\lambda + 2\mu}{2}, z = \frac{\lambda + \mu}{2}$$

by plugging into the constraints, we could get $\lambda = -22/3$ and $\mu = 4$ and thus the point $(7/3, 1/3, -5/3)$ is the solution to the above system.

Now the function f must attain its absolute constrained minimum. Although the intersection of two hyperplanes is not bounded, and function f is continuous and the r-ball is always compact in $\mathbb{R}^3$ for some $r^2 = x_0^2 + y_0^2 + z_0^2$ where (z_0, y_0, z_0) is some point on two hyperplanes. Therefore the constrained absolute min is attained. Moreover we can check that $\nabla h_1 = (1, 1, 1)$ and $\nabla h_2 = (3, 2, 1)$ and they are nonzero everywhere and not collinear. Thus the point $(7/3, 1/3, -5/3)$ has to be point we are looking for.

The integrability of multivariable functions is unlike the single variable case since the region for integration in $\mathbb{R}^n$ is more complicated than intervals. Let us start with integrability on n-rectangles in $\mathbb{R}^n$ and then introduce the notion of Jordan region, *i.e.,* a subset which is Jordan measurable. Then we can define the Riemann integrability over Jordan regions.

Recall that given an n-rectangle $Q \subset \mathbb{R}^n$ and a bounded function $f : Q \to \mathbb{R}$. We say the function f is Riemann integrable on Q, denoted by $f \in R(Q)$ if

$$\underline{\int_Q} f = \sup_G L(f, G) = \inf_G U(f, G) = \overline{\int_Q} f$$

where $G = \{R_1, \cdots, R_p\}$ is a grid of Q and $U(f, G), L(f, G)$ are upper and lower sum defined in the similar way in single variable integration.

Similarly, we also have the Darboux epsilon condition for integrability. Namely if $f : Q \to \mathbb{R}$ is bounded over Q then $f \in R(Q)$ if and only if for any $\epsilon > 0$ there is a grid G of Q such that $U(f, G) - L(f, G) < \epsilon$.

Example 21.2. The function given by

$$f(x, y) = \begin{cases} 1 & \text{if } y > x; \\ 0 & \text{if } y \leq x, \end{cases}$$

is Riemann integrable on unit square $[0, 1]^2$.

First of all f is bounded on $\mathbb{R}^2$. For any $\epsilon > 0$, partition the rectangle $[0, 1]^2$ evenly by dividing each side into n intervals each of length $1/n$. Call this grid G and compute the upper and lower sum of f with respect to the grid G as follows

$$U(f, G) = \sum_{j=1}^{n^2} M_j(f) \frac{1}{n^2} = \frac{n(n+1)}{2} \frac{1}{n^2} = \frac{n(n+1)}{2n^2},$$

$$L(f, G) = \sum_{j=1}^{n^2} m_j(f) \frac{1}{n^2} = \frac{(1 + n - 2)(n - 2)}{2} \frac{1}{n^2} = \frac{(n-2)(n-1)}{2n^2}.$$

Thus

$$U(f, G) - L(f, G) = \frac{1}{2n^2}(4n - 2) < \frac{2n}{n^2} = \frac{2}{n} < \epsilon$$

and we can take n be any natural number greater than $2/\epsilon$.

Using the integrability over rectangles, one can defined a set being Jordan measurable, in other words, a set over which the constant 1 function is integrable. To be precise, let $\Omega \subset \mathbb{R}^n$ be a bounded subset, with $\Omega \subset Q$ for some n-rectangle Q, then Ω is said to be Jordan measurable if the characteristic function χ_Ω is integrable over Q. In this situation, we define the Jordan content $|\Omega|$ of the set Ω as

$$|\Omega| = \int_\Omega \mathrm{d}\boldsymbol{x} = \int_Q \chi_\Omega(x) \mathrm{d}\boldsymbol{x}.$$

The definition for the Jordan region is well-defined, *i.e.,* it does not depend on the choice of rectangle Q. This is because if Q, Q' are two rectangles containing Ω, we can prove using definition that

$$\int_Q \chi_\Omega = \int_{Q \cap Q'} \chi_\Omega = \int_{Q'} \chi_\Omega.$$

The main reason we consider integration over a Jordan region is because we want the boundary of the region to have arbitrarily small contribution to the sum.

This intuitively gives the characterization of a bounded set $\Omega \subset \mathbb{R}^n$ being Jordan measurable: if and only if the boundary $|\partial\Omega|$ has Jordan content zero. Moreover, we can also characterize Jordan content zero sets using finite cover of closed rectangles. Namely, a bounded set $\Omega \subset \mathbb{R}^n$ is Jordan measurable with content zero if and only if for any $\epsilon > 0$ there exists a finite collection of closed rectangles $R_1, \cdots, R_N$ such that

$$\Omega \subseteq \bigcup_{k=1}^{N} R_k, \quad \sum_{k=1}^{N} |R_k| < \epsilon.$$

Let us consider several examples of Jordan content zero sets:

Example 21.3.

- Line segments in $\mathbb{R}^2$ have Jordan content zero. Note that they do NOT have content zero as subsets in $\mathbb{R}$, and thus the concept of Jordan content zero heavily relies on the ambient space.
  For simplicity, let us consider the line segment

  $$L = \{(x, x) \in \mathbb{R}^2 \mid x \in [0, 1]\}.$$

  For any $\epsilon > 0$, take $N = [1/\epsilon] + 1$ and cover L by N rectangles $\{R_1, \cdots, R_N\}$ each of which has side length $1/N$. Thus

  $$\sum_{i=1}^{N} |R_i| = \frac{N}{N^2} = \frac{1}{N} < \epsilon$$

  and hence L has Jordan content zero.
- Unit circle $S^1 \subset \mathbb{R}^2$ has Jordan content zero.
  Divide 2π evenly into 2^n angles, *i.e.*, $\theta = \frac{2\pi}{2^n} = \frac{\pi}{2^{n-1}}$. Cover each circular sector by a rectangle with side lengths $\sin(\frac{\pi}{2^{n-1}})$ and $1 - \cos(\frac{\pi}{2^{n-1}})$. Thus S^1 is covered by 2^n rectangles with total volumes

  $$2^n \sin\frac{\pi}{2^{n-1}} \left(1 - \cos\frac{\pi}{2^{n-1}}\right) \to 0$$

  as $n \to 0$.
- Every finite set in $\mathbb{R}^n$ has Jordan content zero. But there are also infinite sets which also have content zero. For instance the set

  $$E = \left\{ \frac{1}{n} \middle| n \in \mathbb{N} \right\} \subset \mathbb{R}.$$

  (For higher dimensional example, one can take $\{(1/n, 1/m)\}$ similarly.)
  For any $\epsilon > 0$, take the first interval $I_1 = [0, \epsilon/2]$. Then all but finitely many points in E are in I_1. But finite set has content zero so we are done. Alternatively, to be explicit, say $\{1, 1/2, \cdots, 1/m\}$ are not in I_1. Take

  $$I_2 = [1, 1 + \epsilon/2^2], \cdots, I_{m+1} = [m, m + \epsilon/2^{m+1}],$$

then

$$\sum_{j=1}^{m+1} |I_j| = \frac{\epsilon}{2}\left(1 + \frac{1}{2} + \cdots + \frac{1}{2^m}\right) < \epsilon.$$

In general, we have a convergent sequence in $\mathbb{R}^n$ has Jordan content zero. There are also examples(for sure!) of infinite sets which are not even Jordan measurable. For example, $E = [0,1]^n \cap \mathbb{Q}^n \subset \mathbb{R}^n$. This is because

$$\overline{\int}_{[0,1]^n} \chi_E = 1 \neq 0 = \underline{\int}_{[0,1]^n} \chi_E.$$

Thus Jordan measure is not countably additive for nonoverlapping content zero sets, (but Lebesgue measures is).

- Later we will see any curve in $\mathbb{R}^n$ which has a lipschitz continuous parametrization $\gamma : [0,1] \to \mathbb{R}^n$ is of Jordan content zero.
  But note that images of parametric curves in $\mathbb{R}^n$ are not necessarily of content zero. More precisely, there can be continuous functions $x(t), y(t)$ from $[0,1]$ to $\mathbb{R}$ such that $\{(x(t), y(t)) \mid t \in [0,1]\} \subset \mathbb{R}^2$ is not of content zero. For example the space-filling curves, which can be the entire unit square.
- It can be easily deduced from the characterization that if $\Omega \subset \mathbb{R}^n$ has content zero, then every subset of Ω is of content zero; Finite union of content zero sets has content zero.

22 April 21

Let us extend a little bit about the argument above regarding to the parametrized curves. It is true that

- if $\gamma : [0,1] \to \mathbb{R}^n$ is C^1, *i.e.,* continuously differentiable, then $\gamma([0,1])$ has Jordan content zero;
- if $\gamma : [0,1] \to \mathbb{R}^n$ is Lipschitz continuous, then $\gamma([0,1])$ has Jordan content zero.

But, note that it is NOT true that the image of any continuous curve is of content zero, or even Jordan measurable. Moreover, there are examples of continuous functions $(x(t), y(t)) : [0,1] \to \mathbb{R}^2$ such that the image

$$\{(x(t), y(t)) \mid t \in [0,1]\} \subseteq \mathbb{R}^2$$

is not of content zero or not even a Jordan measurable set. Let's consider a special curve in the following examples starting with a one dimensional example of "compact sets are not necessarily Jordan measurable, (but they are Lebesgue measurable)", called *Fat Cantor set.*

Example 22.1 (Fat Cantor set)**.** This is an example of compact set in $\mathbb{R}$ which is not Jordan measurable. The fat Cantor set is a slight modification of the standard Cantor set, *i.e.,* the proportion removed in each step is not a constant 1/3.

Recall that the Cantor set C is compact uncountable with measure zero (thus it is also of Jordan content zero). Moreover, it has empty interior, *i.e.,* it contains no intervals, every point in C is a boundary point, *i.e.,* $\partial C = C$, and it is nowhere dense $\overline{C}^o = \emptyset$.

Peano and Jordan noticed that there is a set even in $\mathbb{R}$, with empty interior, but still cannot be covered with finite intervals with total volume zero. The idea is to start with the construction of Cantor set but try to make it slightly fatter.

The fat Cantor set is constructed as follows, start with the interval $[0,1]$:

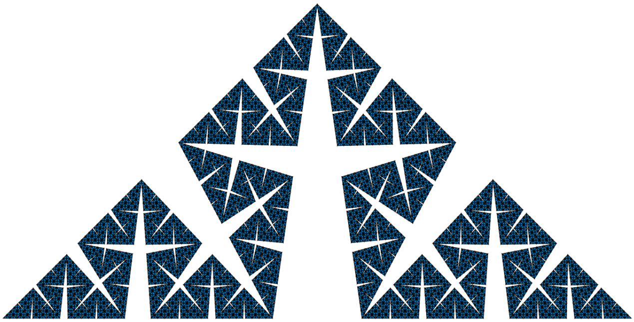
1) Remove the middle (open) $1/4$ from $[0,1]$, *i.e.,* remove $I_1 = (\frac{3}{8}, \frac{5}{8})$ from $[0,1]$, we have the following union left
$$\left[0, \frac{3}{8}\right] \cup \left[\frac{5}{8}, 1\right];$$
2) Remove the middle (open) subintervals of length $1/4^2$ from the remaining 2^1 intervals in the previous step, *i.e.,* remove $I_2 = (\frac{5}{32}, \frac{7}{32}) \cup (\frac{25}{32}, \frac{27}{32})$, the remaining union becomes
$$\left[0, \frac{5}{32}\right] \cup \left[\frac{7}{32}, \frac{3}{8}\right] \cup \left[\frac{5}{8}, \frac{25}{32}\right] \cup \left[\frac{27}{32}, 1\right];$$
n) Remove the middle (open) subintervals of length $1/4^n$ from the remaining 2^{n-1} intervals in the previous step $n-1$. Call the removed interval I_n as used above. Continue this process inductively, the fat Cantor set is defined by
$$C_F = [0,1] \setminus \bigcup_{n=1}^{\infty} I_n.$$

The fat Cantor set C_F has the followings properties:

- We remove a sequence of open sets $\{I_n\}$ each being a finite union of disjoint open intervals. Thus C_F is closed since it is a complement of an open set, and it is bounded, thus compact.
- The total length removed is
$$\sum_{n=1}^{\infty} 2^{n-1} \frac{1}{4^n} = \sum_{n=0}^{\infty} \frac{2^n}{2^{2n+2}} = \frac{1}{4} \sum_{n=0}^{\infty} \frac{1}{2^n} = \frac{1}{2}.$$
Thus C_F has positive measure.
- Similar to the Cantor set, the fat Cantor set C_F has empty interior, *i.e.,* $\partial C_F = C_F$. And it is nowhere dense as well.
- But the fat Cantor set C_F is not Jordan measurable. Let P be any partition of $[0,1]$ and let Q_n be the partition given by the n-th step of iteration. Thus $P_n = P \cup Q_n$ is finer than P thus
$$\frac{1}{2} < U(\chi_{C_F}, P_n) \leq U(\chi_{C_F}, P).$$

It implies that the outer Jordan measure is bounded below by $1/2$ but the inner Jordan measure is zero (it contains no interval) thus it is not Jordan measurable.

Example 22.2 (Osgood curve)**.** A Osgood curve C is a plane curve which is not self-intersecting, *i.e.,* it is the image of a continuous injective function $[0,1] \to \mathbb{R}^2$, compact, and has empty interior, *i.e.,* $\partial C = C$, but has positive area (positive Lebesgue measure). Therefore it cannot be Jordan measurable. The construction can be obtained as a two dimensional version of fat Cantor set on triangle, by removing open wedges with base as the middle $1/4^n$ on each sub-triangle. I am attaching a graph from Wikipedia as below.



Note that open set may not be Jordan measurable. One example is the complement of the fat Cantor set in $\mathbb{R}$, it is bounded open but one can check it is not Jordan measurable. Also open dense sets are not necessarily Jordan measurable.

Example 22.3. Consider the rationals in the interior of $Q = [0,1] \times [0,1]$, *i.e.,* $\mathbb{Q}^2 \cap (0,1)^2$. It is countable so that we can enumerate as $\{r_n\}$. For each r_n let R_n be a small *open* square inside Q such that $\sum_{n=1}^{\infty} |R_n| = 1/2$. For instance one can take the side length for R_n to be $1/2^n$. Then the set $A = \bigcup_{n=1}^{\infty} R_n$ is open and dense, and moreover by density we know that $\overline{A} = Q$. But A is not Jordan measurable. The reason is the following. Consider the characteristic function χ_A on Q. Take any grid $G = \{G_1, \cdots, G_l\}$ of Q, then the upper sum is

$$U(\chi_A, G) = \sum_{G_i \cap A \neq \emptyset} |G_i| = 1.$$

Now set $G' = \{G_i \mid G_i \cap (Q \setminus A) = \emptyset\} = \{G_i \mid G_i \subset A\} \subseteq G$. Then the lower sum is

$$L(\chi_A, G) = \sum_{G_i \in G'} |G_i|.$$

Note that $\cup_{G_i \in G'} G_i$ is compact and admits $\{R_n\}$ as an (infinite) open cover, there is a finite subcover, *i.e.*,

$$\bigcup_{G_i \in G'} G_i \subseteq \bigcup_{i=1}^{N} R_i.$$

Hence

$$L(\chi_A, G) \leq \sum_{i=1}^{N} |R_i| < \frac{\frac{1}{4}}{1-\frac{1}{4}} = \frac{1}{3}.$$

Therefore χ_A is not integrable over Q, which means that the open dense set Q is not Jordan measurable.

There is another example of Thomae function.

Example 22.4 (Thomae function)**.** Consider the function $g : [0,1] \to \mathbb{R}$ given by

$$g(x) = \begin{cases} 1 & \text{if } x = 0; \\ 1/q & \text{if } x \in \mathbb{Q} \text{ and } x = p/q \text{ in the reduced form;} \\ 0 & \text{otherwise.} \end{cases}$$

Consider $\Omega_g = \{(x,y) \mid x \in [0,1], 0 \leq y \leq g(x)\}$. Then Ω_g has Jordan content zero. The reason is as follow. Take any $\epsilon > 0$ we may assume $\epsilon < 1$. Then there are only finitely many of $x \in [0,1]$ such that $f(x) \geq \epsilon/2$. In other words, the set

$$F = \{x \in [0,1] \mid f(x) \geq \epsilon/2\}$$

is finite containing $x = 0$. Now we take the a big rectangle $[0,1] \times [0, \epsilon/2]$ to cover the lines corresponding to $x \in F^c$. What remain are finitely many line segments each of length at most $1 - \epsilon/2$. Therefore we can cover them by finitely many rectangles with total volume $< \epsilon/2$. Put everything together we can cover Ω_g with finite number of rectangles with total volume at most ϵ. Thus Ω_g has Jordan content zero.

23 April 26

The Fubini's Theorem gives us a way to compute multivariable integrals by computing the corresponding iterated integrals, and sometimes it also allows us to interchange the order of integrations to make the computation easier to carry out.

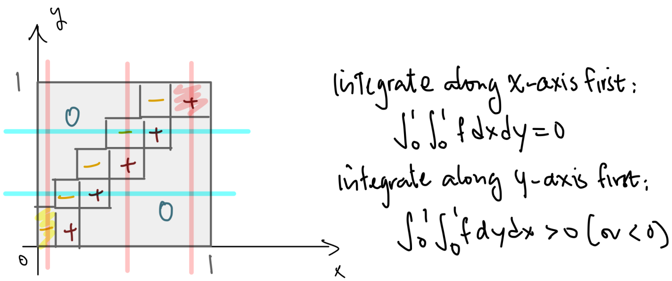
Theorem 23.1 (Fubini's Theorem)**.** *Let $R = [a,b] \times [c,d]$ be a 2-rectangle in $\mathbb{R}^2$. Let $f : R \to \mathbb{R}^2$ be a Riemann integrable function over R. Suppose*

- *For any $x \in [a,b]$ the one variable function $f(x,-) : y \mapsto f(x,y)$ is integrable on $[c,d]$, and*
- *For any $y \in [c,d]$ the one variable function $f(-,y) : x \mapsto f(x,y)$ is integrable on $[a,b]$,*

then
$$\int_R f = \int_a^b \int_c^d f(x,y)\mathrm{d}y\mathrm{d}x = \int_c^d \int_a^b f(x,y)\mathrm{d}x\mathrm{d}y.$$

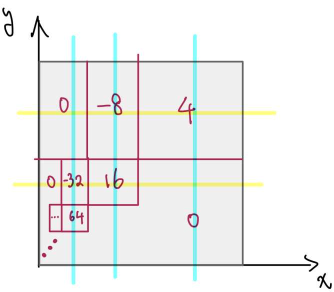
Note that it can happen that both iterated integrals exist but they are not the same, or even two iterated integrals equal but the function is not Riemann integrable. Even we assume two sections f_x, f^y are continuous we may still have that different iterated integrals.

Firstly, we want to construct a function $f(x,y)$ such that f_x and f^y are both integrable, but the iterated integrals are not the same. The idea is the following:



To build up examples following the above idea we need to manipulate the values for each $+/-$ block area.

Example 23.2. Consider the following setup. Each time when we move to the adjacent block we switch the sign of the value but double the absolute value with area shrinked into halves.



The function $f(x,y) : [0,1]^2 \to \mathbb{R}$ can be explicitly written down:

$$f(x,y) = \begin{cases} 2^{2n} & \text{if } (x,y) \in [\frac{1}{2^n}, \frac{1}{2^{n-1}}) \times [\frac{1}{2^n}, \frac{1}{2^{n-1}}); \\ -2^{2n+1} & \text{if } (x,y) \in [\frac{1}{2^{n+1}}, \frac{1}{2^n}) \times [\frac{1}{2^n}, \frac{1}{2^{n-1}}); \\ 0 & \text{otherwise.} \end{cases}$$

First consider, for any fixed $y \in [0,1]$,

$$\int_0^1 f(x,y)\mathrm{d}x = 0$$

since for any horizontal slice, f has only two nonzero values, and the lengths is differed by a multiple of $1/2$. Thus

$$\int_0^1 \int_0^1 f(x,y)\mathrm{d}x\mathrm{d}y = 0.$$

Now consider the other iterated integral. It is easy to see that

$$\int_0^1 f(x,y)\mathrm{d}y = \begin{cases} 0 & \text{if } x \in [0,1/2); \\ \frac{4}{2} & \text{if } x \in [1/2,1). \end{cases}$$

Thus

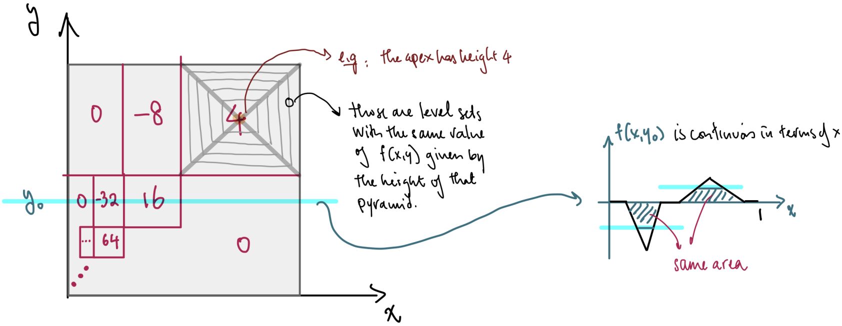
$$\int_0^1 \int_0^1 f(x,y)\mathrm{d}y\mathrm{d}x = \int_{1/2}^1 \int_0^1 f(x,y)\mathrm{d}y\mathrm{d}x = 1.$$

Therefore these two iterated integrals are not equal. Moreover, the function f is not even bounded on $[0,1]^2$ so it cannot be Riemann integrable.

This example can be further modified to a function $f : [0,1]^2 \to \mathbb{R}$ satisfying

- f_x, f^y are continuous on $[0,1]$ (not just integrable);
- one iterated integral is zero the other is positive;
- f is not integrable over $[0,1]^2$.

To make f_x, f^y continuous, we cannot take piecewise constant functions anymore. To modify this, we take each rectangle as a base of a rectangular equilateral pyramid with the value of the function given by the height.



Even if both iterated integrals exist and equal, the function itself may not be Riemann integrable. Here is an example.

Example 23.3. Define a function $f : [0,1]^2 \to \mathbb{R}$ by

$$f(x,y) = \begin{cases} 1 & \text{if } (x,y) \in (\frac{p}{2^n}, \frac{q}{2^n}) \in \mathbb{Q}^2; \\ 0 & \text{otherwise.} \end{cases}$$

For each fixed $y \in [0,1]$, if y is of the form $\frac{q}{2^n}$ then there are only finitely many x's such that $f(x,y) = 1$, namely $x = \frac{1}{2^n}, \frac{2}{2^n}, \cdots, \frac{2^n-1}{2^n}$. If y is not of the form $\frac{q}{2^n}$, then $f(x,y) = 0$ for any $x \in [0,1]$. Thus for any $y \in [0,1]$, the section f^y differs from the the constant zero function only at a finite set of points (content zero). Thus f^y is Riemann integrable on $[0,1]$ and

$$\int_0^1 f(x,y)\mathrm{d}x = 0.$$

Similarly, we get

$$\int_0^1 f(x,y)\mathrm{d}y = 0.$$

Hence

$$\int_0^1 \int_0^1 f(x,y)\mathrm{d}x\mathrm{d}y = \int_0^1 \int_0^1 f(x,y)\mathrm{d}y\mathrm{d}x = 0.$$

But f is NOT Riemann integrable on $[0,1]^2$. Since the set of rationals with denominator a power of 2 is dense in $\mathbb{R}$, we have

$$E = \left\{ \left(\frac{p}{2^n}, \frac{q}{2^n}\right) \middle| p, q \in \mathbb{Z}, n \in \mathbb{Z}_{\geq 0}) \right\} \subset \mathbb{R}^2$$

is dense in $\mathbb{R}^2$ as well. And the complement E^c is also dense in $\mathbb{R}^2$ by the density of irrationals. Thus given any grid G on $[0,1]^2$ any subrectangle $R_i \in G$ contains points in E and also points in E^c. Thus $U(f;G) = 1$ and $L(f;G) = 0$. Therefore f is not Riemann integrable.

One more computational example of application of Fubini's.

Example 23.4. If the region of integration D is Jordan measurable but is not closed, and the function can be continuously extended to $\overline{D}$, then

$$\int_D f = \int_{\overline{D}} f$$

because $|\partial D| = 0$.

Consider the double integral

$$\int_D \frac{\sin y^2}{y} \mathrm{d}x\mathrm{d}y$$

where $D = \{(x, y) \in \mathbb{R}^2 \mid 0 \le x \le y^2, 0 < y \le \sqrt{\pi}\}$. Since

$$\lim_{y\to 0} \frac{\sin y^2}{y} = 0$$

so we can extend f to the boundary of D, by defining

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

We can express $\overline{D}$ as type I and type II region:

$$\begin{aligned} \overline{D} &= \{(x, y) \mid 0 \le x \le y^2, 0 \le y \le \sqrt{\pi}\} \\ &= \{(x, y) \mid 0 \le x \le \pi, \sqrt{x} \le y \le \sqrt{\pi}\}. \end{aligned}$$

As a type I region

$$\int_D f = \int_0^\pi \mathrm{d}x \int_{\sqrt{x}}^{\sqrt{\pi}} \frac{\sin y^2}{y} \mathrm{d}y$$

it is not easy to compute; As a type II region,

$$\int_D f = \int_0^{\sqrt{\pi}} \mathrm{d}y \int_0^{y^2} \frac{\sin y^2}{y} \mathrm{d}x = \int_0^{\sqrt{\pi}} y^2 \frac{\sin y^2}{y} \mathrm{d}y = -\frac{1}{2} \cos y^2 |_0^{\sqrt{\pi}}$$

is doable.

24 April 28

Computing double integrals with or without using Fubini's theorem depends on different situations. The followings are two easy examples.

Example 24.1. • Consider the double integral where $D = [-1, 1] \times [0, 1]$:

$$\iint_D xe^{x^2+y^2} \mathrm{d}x\mathrm{d}y.$$

First notice that $\int e^{y^2} \mathrm{d}y$ has no closed form, thus

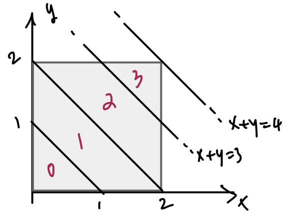
$$\iint_D xe^{x^2+y^2} \mathrm{d}x\mathrm{d}y = \int_0^2 \mathrm{d}y \int_{-1}^1 xe^{x^2+y^2} \mathrm{d}x = 0$$

where the last equality is due to the fact that $xe^{x^2+y^2}$ is an odd function in terms of x and the interval for integration is symmetric about $x = 0$.

- Consider the double integral where $D = [0,2] \times [0,2]$:

$$\iint_D \lfloor x+y \rfloor \mathrm{d}x\mathrm{d}y.$$

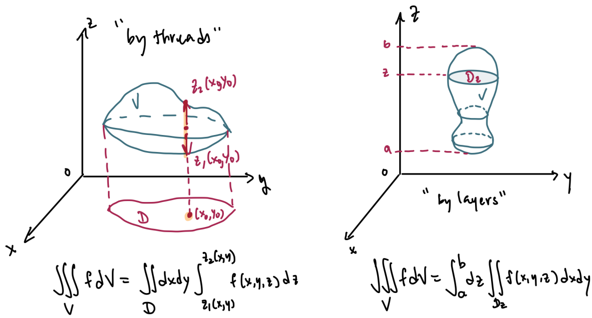
We can of course rewrite the function by dividing the domain D and use the Fubini's theorem by computing iterated integrals. But it is easy to compute the double integral directly since the function is constant on certain regions:



Thus the double integral can easily be computed as

$$\iint_D \lfloor x+y \rfloor \mathrm{d}x\mathrm{d}y = \left(2 - \frac{1}{2}\right) + 2\left(2 - \frac{1}{2}\right) + 3\frac{1}{2} = 6.$$

There is also an n-dimensional Fubini's theorem similar to the 2D case. In particular, for visualization purpose, for triple integrals, we have integration by threads and by layers as follows:



It is formulated as the following Cavalieri's principle:

Proposition 24.2 (Cavalieri's Principle). *Let V be a bounded subset of $\mathbb{R}^3$ and $f : V \to \mathbb{R}$ a Riemann integrable function. And let I denote the triple integral of f over V.*

- *Suppose $V = \{(x,y,z) \mid a \le z \le b, (x,y) \in D_z\}$ where for each $z \in [a,b]$, D_z is a subset of $\mathbb{R}^2$ whose boundary is of content zero in $\mathbb{R}^2$* i.e., *Jordan set, and* *for each fixed $z \in [a,b]$ the double integral $\iint_{D_z} f(x,y,z)\mathrm{d}(x,y)$ exists. Then*

$$\int_a^b \mathrm{d}z \iint_{D_z} f(x,y,z)\mathrm{d}(x,y) = I;$$

- *Suppose $V = \{(x,y,z) \mid (x,y) \in D_0, z_1(x,y) \le z \le z_2(x,y)\}$ where D_0 is the subset of $\mathbb{R}^2$ in the xy-plane which is Jordan measurable, $z_1, z_2 : D_0 \to \mathbb{R}$ are integrable functions such that $z_1 \le z_2$ and for each fixed $(x,y) \in D_0$, the integral $\int_{z_1(x,y)}^{z_2(x,y)} f(x,y,z)\mathrm{d}z$ exists. Then*

$$\iint_{D_0} \mathrm{d}(x,y) \int_{z_1(x,y)}^{z_2(x,y)} f(x,y,z)\mathrm{d}z = I.$$

Example 24.3. One application for Cavalieri's Principle is to compute the volumes of rotations in $\mathbb{R}^3$, *i.e.*, $f(x,y,z) \equiv 1$.

- Let $f : [a,b] \to \mathbb{R}$ be a positive continuous function. Denote by Ω the set in $\mathbb{R}^3$ obtained by revolving the graph of f about the x-axis. Thus the volume of Ω is

$$|\Omega| = \int_a^b |\Omega_x|\mathrm{d}x = \int_a^b \pi f(x)^2 \mathrm{d}x.$$

- Let T be a solid torus in $\mathbb{R}^3$ obtained by revolving the circle $(y-a)^2 + z^2 \le b^2$ in the yz-plane about the z-axis. The the volume of T can be computed as

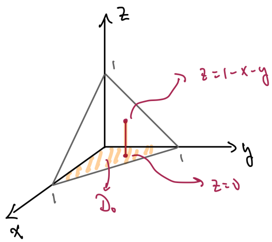
$$\begin{aligned} |T| &= \int_{-b}^b \pi \left[(a + \sqrt{b^2 - z^2})^2 - (a - \sqrt{b^2 - z^2})^2 \right] \mathrm{d}z \\ &= 4\pi a \int_{-b}^b \sqrt{b^2 - z^2}\mathrm{d}z = 2\pi^2 a b^2. \end{aligned}$$

For triple integral by threads, consider the following example.

- Consider the triple integral

$$I = \iiint_\Omega \frac{1}{(1+x+y+z)^3}\mathrm{d}(x,y,z)$$

where $\Omega = \{(x,y,z) \mid x+y+z \le 1, x,y,z \ge 0\}$.



Then by threads, the integral can be written as

$$I = \iint_{D_0} \mathrm{d}(x,y) \int_0^{1-x-y} \frac{\mathrm{d}z}{(1+x+y+z)^3},$$

where $D_0 = \{(x,y) \in \mathbb{R}^2 \mid x, y \geq 0, x+y \leq 1\}$.
By layers, we can write

$$I = \int_0^1 \mathrm{d}z \iint_{D_z} \frac{1}{(1+x+y+z)} \mathrm{d}(x,y)$$

where for each fixed $z \in [0,1]$, $D_z = \{(x,y) \mid x+y \leq 1-z, x, y \geq 0\}$.

For change of variable formula, let us first recall the definition of coordinate transformation. Let $V \subset \mathbb{R}^n$ be an open set and $\phi : V_{\boldsymbol{u}} \to \mathbb{R}^n_{\boldsymbol{x}}$ a vector-valued function such that
- ϕ is $1-1$ on V;
- ϕ is of class C^1 on V;
- for any $\boldsymbol{u} \in V$ the Jacobian $D\phi(\boldsymbol{u})$ is invertible.

Such ϕ is called a coordinate transformation.

Now if $E \subset V$ with $\overline{E} \subset V$ (compact), then ϕ preserves the Jordan measurability and

$$\int_{\phi(E)} f(\boldsymbol{x}) \mathrm{d}\boldsymbol{x} = \int_E f(\phi(\boldsymbol{u})) |\det D\phi(\boldsymbol{u})| \mathrm{d}\boldsymbol{u}.$$

Remark 24.4.

- In practice some of the assumptions can be relaxed. For example ϕ can be $1-1$ with invertible Jacobian aways from a Jordan content zero set.
- The $1-1$ is essential. It guarantees that $\phi(E)$ is only traced out exactly once. For instance, consider the polar transformation $\phi : \mathbb{R}^2 \to \mathbb{R}^2$ given by $(r,\theta) \mapsto (r\cos\theta, r\sin\theta)$ which is C^1 with Jacobian of determinant r. So the Jacobian is invertible when $r \neq 0$ (but since if $r = 0$ the origin has Jordan content zero so it is okay). Now we take $E = [0,1] \times [0, 2\pi]$ for (r,θ) which is a Jordan region in $\mathbb{R}^2$ and $\phi(E) = \overline{B_1(\mathbf{0})}$ is also Jordan measurable.
  Note that ϕ is not $1-1$ on $E = [0,1] \times [0,2\pi]$ since

  $$\phi(r,0) = \phi(r, 2\pi), \quad \text{for all } r \in [0,1].$$

  But we can still apply the change of variable formula since the set on which ϕ fails to be $1-1$ is two line segments in $\mathbb{R}^2$ which has content zero.
  But if we take $E = [0,1] \times [0, 4\pi]$ the image $\phi(E)$ does not change. Now the change of variable formula cannot apply and indeed,

  $$\iint_{\phi(E)} \mathrm{d}(x,y) = \pi$$

but
$$\iint_E r\mathrm{d}(r,\theta) = \int_0^{4\pi} \mathrm{d}\theta \int_0^1 r\mathrm{d}r = 2\pi.$$
This is because ϕ fails to be $1-1$ when $r \in [0,1]$ and $\theta \in [0, 2\pi]$ since
$$\phi(r,\theta) = \phi(r, \theta + 2\pi).$$
and $[0,1] \times [0, 2\pi]$ does not has Jordan content zero in $\mathbb{R}^2$. Geometrically allowing θ to vary from 0 to 4π the unit disc is traced out twice (thus the area is getting doubled).

25 May 3

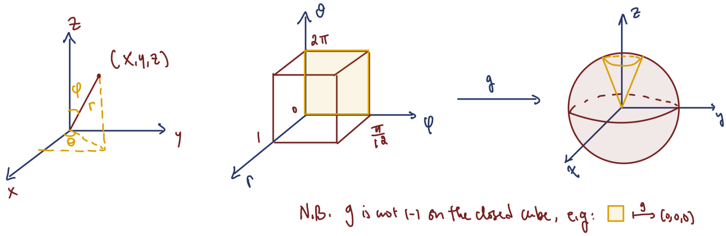
Recall the Change of Variables formula:

Theorem 25.1. *Let $\Omega \subset \mathbb{R}^n$ be an open set and $\boldsymbol{g} : \Omega \to \mathbb{R}^n$ be a C^1-diffeomorphism on Ω. Then for any bounded $E \subset \mathbb{R}^n$ such that $\overline{E} \subseteq \Omega$,*

- *E is Jordan measurable if and only if $\boldsymbol{g}(E)$ is Jordan measurable;*
- *If E is Jordan measurable then f is Riemann integrable on $\boldsymbol{g}(E)$ if and only if $(f \circ \boldsymbol{g})|\det D\boldsymbol{g}|$ is Riemann integrable on E. In this case we have the change of variables formula*
$$\int_{\boldsymbol{g}(E)} f(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = \int_E f(\boldsymbol{g}(\boldsymbol{u}))|\det D\boldsymbol{g}(\boldsymbol{u})|\mathrm{d}\boldsymbol{u}.$$

Remark 25.2. The conclusion is still true if the C^1 mapping $g : \Omega \to \mathbb{R}^n$ is only assumed to be a C^1 diffeomorphism on the interior of E, instead of on the entire neighborhood Ω of E. This extended version has been used in most practical situations. For instance the polar coordinate transformation, spherical transformation etc.

Example 25.3. Let Ω be a "cone" bounded above by unit sphere and below by a cone with vertex angle $\frac{\pi}{6}$. Use the spherical coordinate transformation $\boldsymbol{g} : \mathbb{R}^3 \to \mathbb{R}^3$ given by
$$\boldsymbol{g}(r, \varphi, \theta) = (r \sin\varphi \cos\theta, r\sin\varphi\sin\theta, r\cos\varphi).$$
Therefore the cone Ω is the image under the transformation $\boldsymbol{g}$ of the cube
$$Q = \{(r, \varphi, \theta) \in \mathbb{R}^3 \mid r \in [0,1], \varphi \in [0, \pi/12], \theta \in [0, 2\pi]\} :$$



The transformation map is not $1-1$ on the entire Q but it is a C^1 diffeomorphism in the interior of Q, together with the Fubini's theorem we have

$$\int_{\Omega} f = \int_0^{2\pi} \int_0^{\pi/12} \int_0^1 f(r\sin\varphi\cos\theta, r\sin\varphi\sin\theta, r\cos\varphi) r^2 \sin\varphi \mathrm{d}r\mathrm{d}\varphi\mathrm{d}\theta$$

due to the fact that $|\det D\boldsymbol{g}(r,\varphi,\theta)| = r^2\sin\varphi$.

We use the change of variables formula to simplify not only the expression of the integrand, but also the region for integration so that other techniques such as Fubini's theorem may apply.

Here is another example of cylindrical coordinate transformation.

Example 25.4. Consider the following triple integral:

$$\iiint_{\Omega} e^{x^2+y^2} \mathrm{d}(x,y,z)$$

where $\Omega = \{(x,y,z) \in \mathbb{R}^3 \mid x^2+y^2 \le R^2, 0 \le z \le h\}$ for some fixed position $R, h \in \mathbb{R}$ is a cylinder region in $\mathbb{R}^3$ with base radius R and height h.

We use the cylindrical coordinate transformation $\boldsymbol{g} : \mathbb{R}^3 \to \mathbb{R}^3$ given by

$$\boldsymbol{g}(r,\varphi,z) = (r\cos\varphi, r\sin\varphi, z).$$

Therefore the region Ω is the image under the transformation $\boldsymbol{g}$ of the cube

$$Q = \{(r,\varphi,z) \in \mathbb{R}^3 \mid r \in [0,R], \varphi \in [0,2\pi], z \in [0,h]\}.$$

Hence by Fubini's theorem, we can compute

$$\begin{aligned}\iiint_{\Omega} e^{x^2+y^2}\mathrm{d}(x,y,z) &= \iiint_Q e^{r^2} r \mathrm{d}(r,\varphi,z) \\ &= \int_0^h \mathrm{d}z \int_0^{2\pi} \mathrm{d}\varphi \int_0^R e^{r^2} r\mathrm{d}r \\ &= 2\pi h \frac{1}{2} e^{r^2}|_0^R = \pi h(e^{R^2}-1).\end{aligned}$$

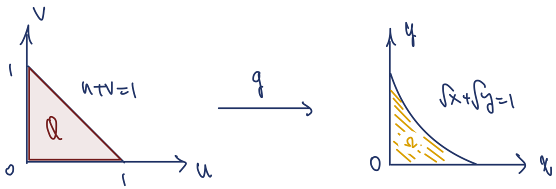
Other transformation as

Example 25.5. Compute the double integral

$$\iint_{\Omega} (x+y)^{1/2} \mathrm{d}(x,y)$$

where Ω is the region in the first quadrant which is bounded by the axes and the parabola $\sqrt{x}+\sqrt{y}=1$.

Consider the substitution $u=\sqrt{x}, v=\sqrt{y}$ thus the transformation mapping is $\boldsymbol{g}:(u,v)\mapsto(u^2,v^2)$



with Jacobian matrix

$$D\boldsymbol{g}(u,v)=\begin{pmatrix} 2u & 0 \\ 0 & 2v \end{pmatrix} \implies \det D\boldsymbol{g}(u,v)=4uv.$$

Then together with Fubini's theorem

$$\iint_{\Omega} (x+y)^{1/2}\mathrm{d}(x,y)=4\iint_{Q} uv(u^2+v^2)^{1/2}\mathrm{d}(u,v)=4\int_0^1 \mathrm{d}v \int_0^{1-v} uv(u^2+v^2)\mathrm{d}u.$$

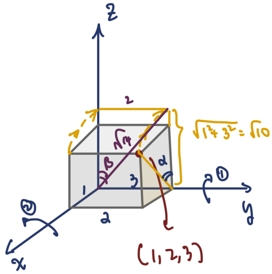
The other commonly used transformation is the orthogonal transformation or rigid rotation.

Example 25.6. Consider the triple integral

$$\iiint_{\Omega} \cos(x+2y+3z)\mathrm{d}(x,y,z)$$

where Ω is the unit ball in $\mathbb{R}^3$.

The idea is using some rotation to make the hyperplane given by $x+2y+3z$ to be either the plane spanned by any of the two axes directions. Or in other words, rotate the normal vector $(1,2,3)$ to either $(0,0,1)$ or $(0,1,0)$ or $(1,0,0)$. For example to rotation $(1,2,3)$ to $(0,0,1)$ we can compose two rotations



1st rotation (about y) by $-\alpha$

$$\begin{pmatrix} \frac{3}{\sqrt{10}} & 0 & -\frac{1}{\sqrt{10}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{10}} & 0 & \frac{3}{\sqrt{10}} \end{pmatrix}$$

2nd rotation (about x) by β

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{\sqrt{10}}{\sqrt{14}} & -\frac{2}{\sqrt{14}} \\ 0 & \frac{2}{\sqrt{14}} & \frac{\sqrt{10}}{\sqrt{14}} \end{pmatrix}$$

Hence the transformation matrix is

$$\begin{pmatrix} \frac{3}{\sqrt{10}} & 0 & -\frac{1}{\sqrt{10}} \\ -\frac{2}{\sqrt{140}} & \frac{\sqrt{100}}{\sqrt{140}} & -\frac{6}{\sqrt{140}} \\ \frac{1}{\sqrt{14}} & \frac{2}{\sqrt{14}} & \frac{3}{\sqrt{14}} \end{pmatrix}$$

From the third row we can rewrite the original integral as

$$\iiint_{\Omega} \cos(\sqrt{14}w)\mathrm{d}(u, v, w).$$

To proceed one can apply another spherical transformation as follows

$$2\pi \int_0^1 \int_0^{\pi} \cos(\sqrt{14}r \cos\varphi) r^2 \sin\varphi \mathrm{d}r \mathrm{d}\varphi.$$