MATH 237

2019年5月6日 14:32

MATH 237
Joe Pharaon

jpharaon@uwaterloo.ca
MC 6417

Tutorial Centre
MC 3022

Mondays 3:30 - 4:30
Wednesdays 1:20 - 2:20

Chapter 1 Graphs of Scalar Functions

2019年1月13日 17:46

1.1 Scalar Functions

2019年1月13日 20:47

A function is a RULE that assigns an element of some sets A a unique element in some sets B.

A set is a collection of elements.

Recall that a function $f: A \to B$ associates with each element $a \in A$ a unique element $f(a) \in B$ called the **image** of a under f. The set A is called the **domain** of f and is denoted by $D(f)$. The set B is called the **codomain** of f.

Ex: $f(x) = x$
Domain: $\mathbb{R}$
Codomain: $\mathbb{R}$
Range: $\mathbb{R}$

$f(x) = e^x$
Domain: $\mathbb{R}$
Codomain: $\mathbb{R}$
Range: $(0, +\infty)$

A singular variable scalar valued function is a rule that assigns to an element of $A \subseteq \mathbb{R}$ a unique element in $B \subseteq \mathbb{R}$.

A bi-variate scalar valued function is a rule that assigns to an element of $\mathbb{R}^2$. A unique element in $B \subseteq \mathbb{R}$.

$f(\vec{x}), \vec{x} \in \mathbb{R}^n$
$f(x_1, x_2, x_3, x_4, \dots, x_n)$

Definition: **Scalar Function**
A **scalar function** $f(x_1, \dots, x_n)$ of n-variables in a function whose domain is a subset of $\mathbb{R}^n$ and whose range is a subset of $\mathbb{R}$.

$f(x, y) = x + y$

Domain $\mathbb{R}^2$
Range $\mathbb{R}$

$f(2,0) = 2 + 0 = 2.$

$f(x, y) = x^2 - y$
Quadratic Function
Linear Function
$f(x, y) = 5$

Domain $\mathbb{R}^2$
Range $\{5\}$ singleton

$f(x, y) = x^2 + y^2$
Quadratic Function
Paraboloid
$z = x^2 + y^2$

Domain $\mathbb{R}^2$
Range $[0, +\infty)$

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

Domain: $\{(x,y) \in \mathbb{R}^2 | x \neq 0 \text{ and } y \neq 0\}$
Range : $(-\infty, 0) \cup (0, +\infty)$

$f(x,y) = x^2 - y^2$
Hyperboloid
Domain $\mathbb{R}^2$
Range $\mathbb{R}$

$f(x,y) = \sqrt{8 - 2x^2 - 7y^2}$
Domain: $8 - 2x^2 - 7y^2 \geq 0$
$x^2 + y^2 \leq 4$

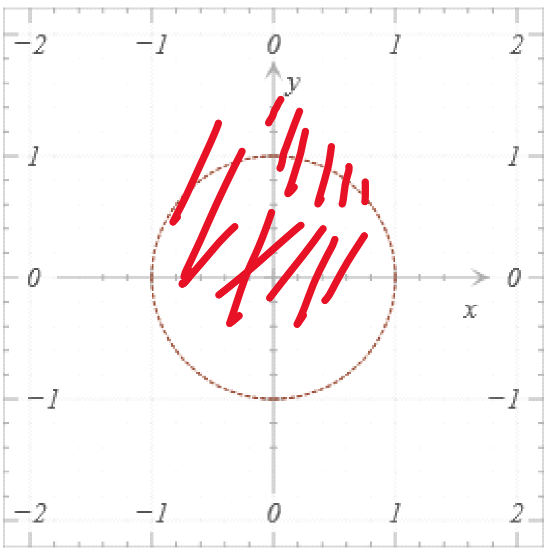
$D = \{(x,y) \in \mathbb{R}^2 \, |x^2 + y^2 \geq 4\}$
Range: $[0, \sqrt{8}]$

Hemisphere
Part of an ellipsoid

Exercise:
$f(x,y) = e^{xy}$
$f(x,y) = \ln(y + x^2 - 1)$

Exersice 1
Sketch the domain and find the range of the following functions:

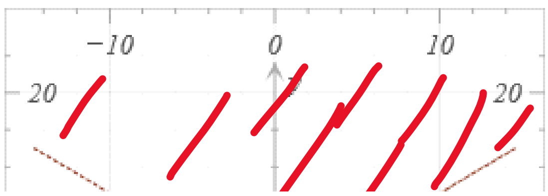
1. $f(x,y) = \ln(1 - x^2 - y^2)$
   Domain:
   $(1 - x^2 - y^2) > 0$
   $1 > x^2 + y^2$



   Range:
   Max Value of $1 - x^2 - y^2$ is 1
   Min is 0

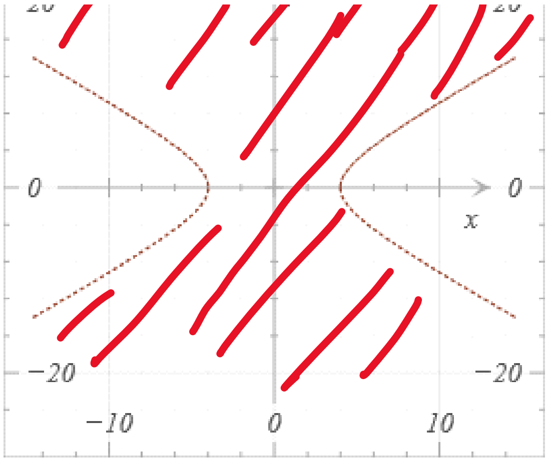
   Thus, range $(-\infty, 0]$

2. $g(x,y) = \sqrt{16 - x^2 + y^2}$
   Domain:
   $16 - x^2 + y^2 > 0$
   $16 > x^2 - y^2$



Domain:

$$16 - x^2 + y^2 > 0$$
$$16 > x^2 - y^2$$



Range:

$\mathbb{R}$

1.2 Geometric Interpretation of $z = f(x, y)$

2019年1月14日 10:00

We define the **graph** of a function $f(x, y)$ as the set of all points $(a, b, f(a, b))$ in $\mathbb{R}^3$ such that $(a, b) \in D(f)$. We think of $f(a, b)$ as representing the height of the graph $z = f(x, y)$ above the xy-plane at the point $(x, y) = (a, b)$

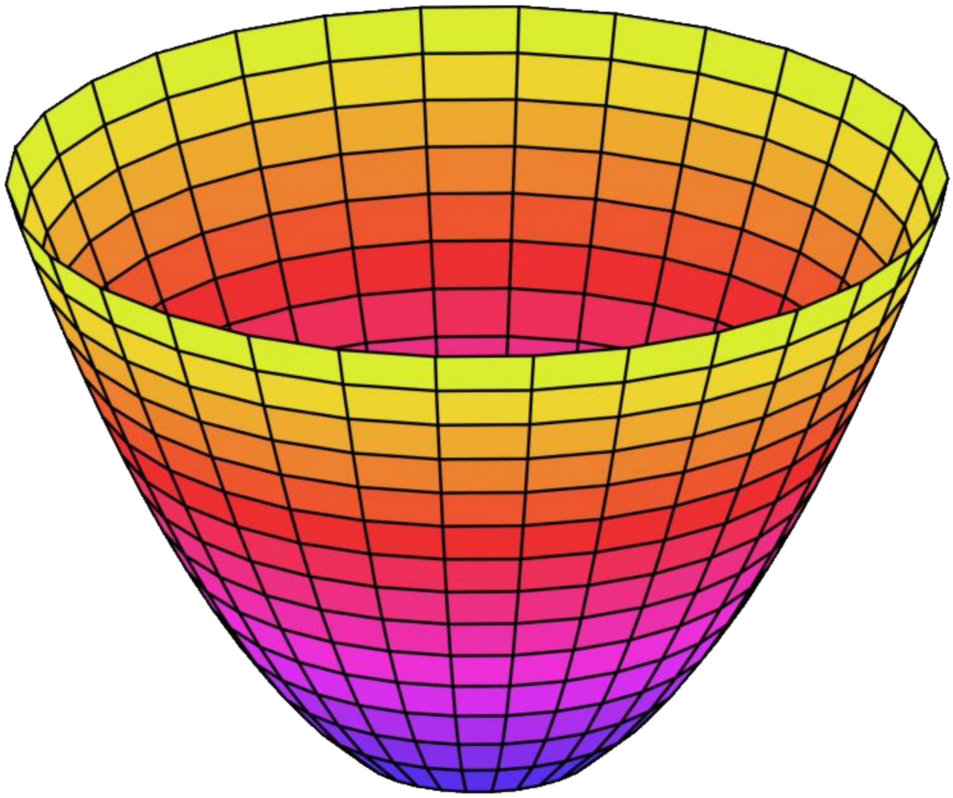
Definition: **Level Curves**
The **level curves** of a function $f(x, y)$ are the curves

$$f(x, y) = k$$

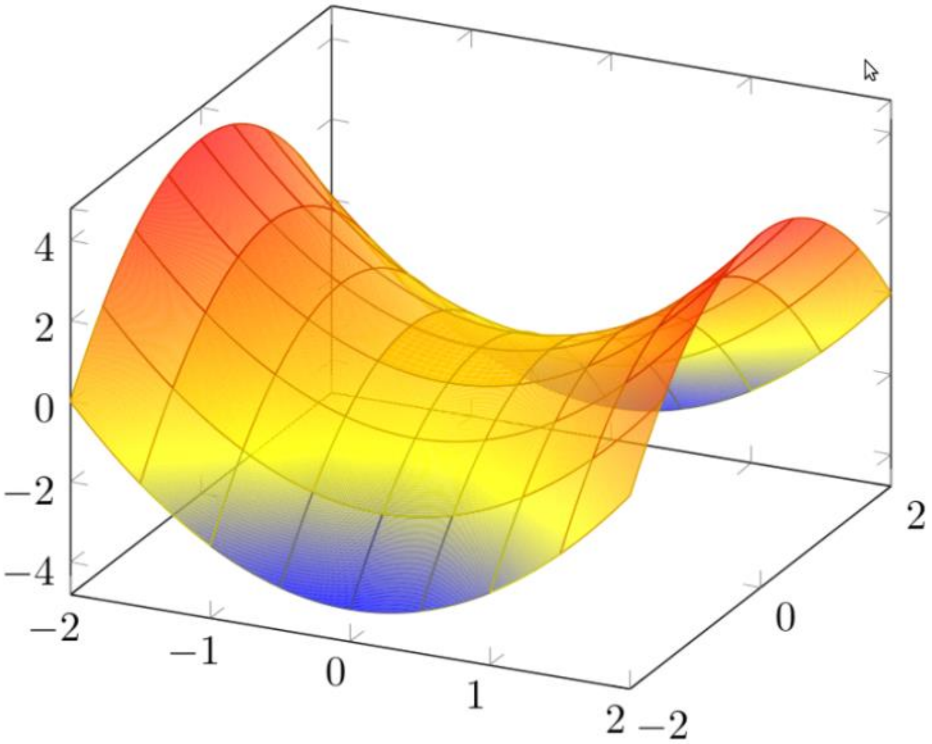
Where k is a constant in the range of f.

The family of level curves is often called a **contour map** or a **topographic** map.

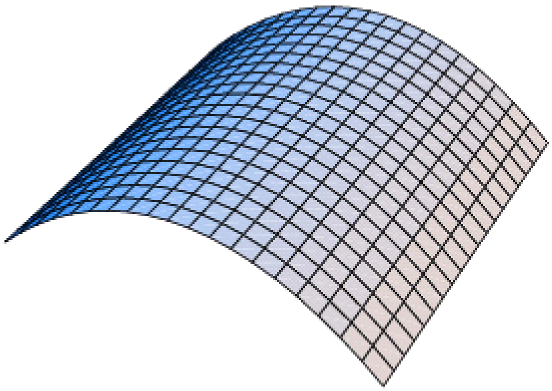
The single point $(0,0)$ is called an **exceptional level curve**.



Paraboloid



Saddle surface



Parabolic cylinder

Some other examples include use in weather maps to show curves of constant temperature called **isotherms**, in marine charts to indicate water depths, and in barometric pressure charts to show curves of constant pressure called **isobars**.

Definition: **Cross Sections**
A **cross section** of a surface $z = f(x, y)$ is the intersection of $z = f(x, y)$ with a plane.

For the purpose of sketching the graph of a surface $z = f(x, y)$, it is useful to consider the cross sections formed by intersecting $z = f(x, y)$ with the vertical planes $x = c$ and $y = d$.

Generalization:
Definition: **Level Surfaces**
A **level surface** of a scalar function $f(x, y, z)$ is defined by

$$f(x, y, z) = k, \qquad k \in R(f)$$

Definition: **Level Sets**
A **level set** a scalar function $f(\mathbf{x}), \mathbf{x} \in \mathbb{R}^n$ is defined by

$$f(\mathbf{x}) = k, \qquad k \in R(f)$$

Chapter 2 Limits

2019年1月14日 22:58

2.1 Definition of a Limit

2019年1月14日 22:58

Definition: **Neighborhood**

A r-**neighbourhood** of a point $(a, b) \in \mathbb{R}^2$ is a set

$$N_r(a, b) = \{(x, y) \in \mathbb{R}^2 \mid \|(x, y) - (a, b)\| < r\}$$

Remark

Recall that $\|(x, y) - (a, b)\|$ is the Euclidean distance in $\mathbb{R}^2$. That is,

$$\|(x, y) - (a, b)\| = \sqrt{(x - a)^2 + (y - b)^2}$$

Definition: **Limit**

Assume $f(x, y)$ is defined in a neighborhood of (a, b), except possibly at (a, b). If for every $\epsilon > 0$ there exists a $\delta > 0$ such that

$$0 < \|(x, y) - (a, b)\| < \delta \quad \text{implies} \quad |f(x, y) - L| < \epsilon$$

Then

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

2.2 Limit Theorems

2019年1月15日 10:17

Theorem 1

If $\lim_{(x,y)\to(a,b)} f(x,y)$ and $\lim_{(x,y)\to(a,b)} g(x,y)$ both exist, then

1. $\lim_{(x,y)\to(a,b)}[f(x,y)+g(x,y)] = \lim_{(x,y)\to(a,b)} f(x,y) + \lim_{(x,y)\to(a,b)} g(x,y)$
2. $\lim_{(x,y)\to(a,b)}[f(x,y)g(x,y)] = \left[\lim_{(x,y)\to(a,b)} f(x,y)\right]\left[\lim_{(x,y)\to(a,b)} g(x,y)\right]$
3. $\lim_{(x,y)\to(a,b)} \dfrac{f(x,y)}{g(x,y)} = \dfrac{\lim_{(x,y)\to(a,b)} f(x,y)}{\lim_{(x,y)\to(a,b)} g(x,y)}$, provide $\lim_{(x,y)\to(a,b)} g(x,y) \neq 0$

Proof:

We will prove (a) and leave (b) and (c) as exercises. Let $\epsilon > 0$. Since $\lim_{(x,y)\to(a,b)} f(x,y) = L_1$ and $\lim_{(x,y)\to(a,b)} g(x,y) = L_2$ both exist, by definition of a limit, there exists a $\delta > 0$ such that

$$0 < \left\|(x,y)-(a,b)\right\| < \delta \quad \text{implies} \quad |f(x,y)-L_1| < \frac{1}{2}\epsilon \text{ and } |g(x,y)-L_2| < \frac{1}{2}\epsilon$$

Thus, if $0 < \left\|(x,y)-(a,b)\right\| < \delta$, then

$$\begin{aligned}|f(x,y)+g(x,y)-(L_1+L_2)| &= |[f(x,y)-L_1]+[g(x,y)-L_2]| \\ &\leq |f(x,y)-L_1| + |g(x,y)-L_2| \\ &< \frac{\epsilon}{2}+\frac{\epsilon}{2} = \epsilon\end{aligned}$$

Theorem 2

If $\lim_{(x,y)\to(a,b)} f(x,y)$ exists, then the limit is unique.

Proof:

Assume that $\lim_{(x,y)\to(a,b)} f(x,y) = L_1$ and $\lim_{(x,y)\to(a,b)} f(x,y) = L_2$. Then,

$$|L_1 - L_2| = \left|\lim_{(x,y)\to(a,b)} f(x,y) - \lim_{(x,y)\to(a,b)} f(x,y)\right| = \left|\lim_{(x,y)\to(a,b)}[f(x,y)-f(x,y)]\right| = 0$$

Hence, $L_1 = L_2$ and so the limit is unique.

2.3 Proving a Limit Does Not Exist

2019年1月15日 11:45

Exercise 1

Let $f(x,y)=\frac{(|x|)}{|x|+y^2}$, for $(x,y)\neq(0,0)$. Show that
$$\lim_{(x,y)\to(0,0)} f(x,mx)=1$$
For all $m\in\mathbb{R}$, but $\lim_{(x,y)\to(0,0)} f(x,y)$ does not exist.

Hint: $y=mx$ does not describe all lines through the origin.

Solution:

Approaching the limit along lines $y=mx$ we get
$$\begin{aligned}\lim_{(x,y)\to(0,0)} f(x,mx) &= \lim_{(x,y)\to(0,0)} \frac{|x|}{|x|+m^2x^2}\\ &= \lim_{(x,y)\to(0,0)} \frac{1}{1+m^2|x|}\\ &= 1\end{aligned}$$

However, if we approach the limit through $x=0$.
We get
$$\begin{aligned}\lim_{(x,y)\to(0,0)} f(0,y) &= \lim_{(x,y)\to(0,0)} \frac{0}{0+y^2}\\ &= 0\end{aligned}$$

Example 3

Let $f(x,y)=\dfrac{x^2y}{x^4+y^4}$, for $(x,y)\neq(0,0)$. Show that $\lim_{(x,y)\to(0,0)} f(x,y)$ does not exist.

Solution:

As before we first test the limit along lines $y=mx$. We get
$$\lim_{(x,y)\to(0,0)} f(x,mx)=\lim_{x\to0}\frac{x^2(mx)}{x^4+(mx)^2}=\lim_{x\to0}\left(\frac{mx}{x^2+m^2}\right)=0$$
And
$$\lim_{(x,y)\to(0,0)} f(0,y)=\lim_{y\to0}\frac{0}{y^2}=\lim_{y\to0}0=0$$
These all give the same value, so we start testing curves. Of course, we don't want to start randomly guessing curves. To get a limit other than 0, we need the power of x everywhere in the denominator to match the power of x in the numerator (so that they cancel out). This prompts us to try the limit along $y=x^2$. We get
$$\lim_{(x,y)\to(0,0)} f(x,x^2)=\lim_{x\to0}\frac{\left(x^2(x^2)\right)}{x^4+(x^2)^2}=\lim_{x\to0}\frac{1}{2}=\frac{1}{2}$$
Since we have two different values along two different paths, the limit does not exist.

Remark

1. We could have done the last example more efficiently by just testing $y=mx^2$ to begin with and showing the limit depends on m.
2. Make sure that all lines or curves you use actually approach the limit. A common error is to approach a limit like in Example 3 along a line $x=1$... Which of course is meaningless as it does not pass through (0,0).
3. Example 3 shows that no matter how many lines and / or curves you test, you cannot use this method to prove a limit exists. Just because you haven't found two paths that give different values does not mean there isn't one!

2.4 Proving a Limit Exists

2019年1月15日 19:35

Theorem 1 (Squeeze Theorem)
If there exists a function $B(x,y)$ such that

$$|f(x,y)-L| \leq B(x,y), \qquad \text{for all } (x,y) \neq (a,b)$$

In some neighborhood of (a,b) and $\lim_{(x,y)\to(a,b)} B(x,y) = 0$, then

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

Proof:
Let $\epsilon > 0$. Since $\lim_{(x,y)\to(a,b)} B(x,y) = 0$ we have that there exists a $\delta > 0$ such that

$$0 < \left\| (x,y)-(a,b) \right\| < \delta \text{ implies } |B(x,y)-0| < \epsilon$$

Hence, if $0 < \left\| (x,y)-(a,b) \right\| < \delta$, then we have

$$|f(x,y)-L| \leq B(x,y) = |B(x,y)| < \epsilon$$

As our hypothesis requires that $B(x,y) \geq 0$ for all $(x,y) \neq (a,b)$ in the neighborhood of (a,b).
Therefore, by definition of a limit, we have

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

Exercise 1
Our statement of the Squeeze Theorem above is not a direct generalization of the Squeeze Theorem we used in single variable calculus. What would the direct generalization of the Squeeze Theorem be? Show how your generalization and the theorem above are related.

Remark
Be careful when working with inequalities! For example, the statement

$$x < x^2$$

Is **false** if $|x| < 1$. The appendix at the end of this chapter gives a brief review of inequalities.

Exercise 2.
Prove that

$$\frac{|x^3-y^3|}{x^2+y^2} \leq |x|+|y| \quad \text{for all } \ (x,y) \neq (0,0)$$

Does equality ever hold?

Solution:

$$|x|+|y| \geq |x+y|$$

$$x^3-y^3 = (x-y)(x^2+xy+y^2)$$

$$\begin{aligned} |x^3-y^3| &\leq |x^3|+|-y^3| \\ &= |x^3|+|y^3| \\ &\leq |x^2|(|x|+|y|)+|y^2|(|x|+|y|) \\ &= (|x|+|y|)(x^2+y^2) \end{aligned}$$

Thus,

$$\frac{|x^3-y^3|}{x^2+y^2} \leq \frac{\left((|x|+|y|)(x^2+y^2)\right)}{x^2+y^2} = |x|+|y|$$

Example 3

Determine whether $\lim_{(x,y)\to(0,0)} \frac{x^2-|x|-|y|}{|x|+|y|}$ exists, and if so find its value.

Solution:

Trying lines $y = mx$ we get

$$\lim_{x\to 0} \frac{\left(x^2-|x|-m|x|\right)}{|x|+m|x|} = \lim_{x\to 0} \frac{|x|-(1+m)}{1+m} = -1$$

Since the value along each line is $L = -1$, we try to prove the limit is -1 with the Squeeze Theorem. Thus, we consider

$$\begin{aligned}\left|\frac{x^2-|x|-|y|}{|x|+|y|} - (-1)\right| &= \left|\frac{x^2-|x|-|y|}{|x|+|y|} + \frac{|x|+|y|}{|x|+|y|}\right| \\ &= \frac{x^2}{|x|+|y|} \\ &= \frac{|x|\cdot|x|}{|x|+|y|} \\ &\leq \frac{|x|\left(|x|+|y|\right)}{|x|+|y|} = |x|\end{aligned}$$

Since $|x| \leq \left(|x|+|y|\right)$

Since $\lim_{(x,y)\to(0,0)} |x| = 0$ we get $\lim_{(x,y)\to(0,0)} \frac{x^2-|x|-|y|}{|x|+|y|} = -1$ by the Squeeze Theorem.

Exercise 3

Consider f defined by

$$f(x,y) = \frac{x^2(x-1)-y^2}{x^2+y^2}, \qquad \text{for } (x,y) \neq (0,0)$$

Determine whether $\lim_{(x,y)\to(0,0)} f(x,y)$ exists, and if so find its value.

Solution:

Tying lines $y = mx$ we get

$$\lim_{x\to 0} \frac{\left(x^2(x-1)-(mx)^2\right)}{x^2+(mx)^2} = \lim_{x\to 0} \frac{x^2(x-1)-m^2x^2}{(m^2+1)x^2} = \lim_{x\to 0} \frac{(x-1-m)x^2}{(m^2+1)x^2} = \frac{-1-m}{m^2+1}$$

Clearly, the limit does not exist.

Remark

The concept of a neighbourhood, the definition of a limit, the Squeeze Theorem and the limit theorems are all valid for scalar functions $f(\mathbf{x}), \mathbf{x} \in \mathbb{R}^n$. In fact, to generalize these concepts, one only needs to recall that if $\mathbf{x} = (x_1, \dots, x_n)$ and $\mathbf{a} = (a_1, \dots, a_n)$ are in $\mathbb{R}^n$, then the Euclidean distance from $\mathbf{x}$ to $\mathbf{a}$ is

$$||\mathbf{x} - \mathbf{a}|| = \sqrt{(x_1-a_1)^2 + \cdots + (x_n-a_n)^2}$$

2.5 Appendix: Inequalities

2019年1月17日 13:06

Trichotomy Property: For any real numbers a and b, one and only one of the following holds:

$a = b, \qquad a < b, \qquad b < a$

Transitivity Property: If $a < b$ and $b < c$, then $a < c$.
Addition Property: If $a < b$, then for all c, $a + c < b + c$.
Multiplication Property: If $a < b$ and $c < 0$, then $bc < ac$.

Using these properties one can deduce other results.

The **absolute value** of a real number a is defined by

$$|a| = \begin{cases} a & \text{if } a \geq 0 \\ -a & \text{if } a < 0 \end{cases}$$

Three frequently used results, which follow from the axioms, are listed below.

1. $|a| = \sqrt{a^2}$
2. $|a| < b$ if and only if $-b < a < b$.
3. The Triangle Inequality: $|a + b| \leq |a| + |b|$ for all $a, b \in \mathbb{R}$.

Remark.
When using the Squeeze Theorem, the most commonly used inequalities are:

1. The Triangle Inequality
2. If $c > 0$, then $a < a + c$
3. The cosine inequality $2|x||y| \leq x^2 + y^2$

Chapter 3 Continuous Functions

2019年1月17日 13:21

3.1 Definition of a Continuous Function

2019年1月17日 13:21

In many situations, we shall require that a function $f(x, y)$ is **continuous.** Intuitively, this means that the graph of f (the surface $z = f(x, y)$) has no "breaks" or "holes" in it. As with functions of one variable, continuity is defined by using limits.

Exercise 1
Review the definition of a continuous function of one variable in your first year calculus text. Give an example (formula and graph) of a function $y = f(x)$ which is defined for all $x \in \mathbb{R}$, but is not continuous at $x = 1$.

Definition: Continuous
A function $f(x, y)$ is **continuous** at (a, b) if and only if

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

Additionally, if f is continuous at every point in a set $D \subset \mathbb{R}^2$, then we say that f is continuous on D.

Remark
There are really three requirements in this definition:

1. $\lim_{(x,y)\to(a,b)} f(x, y)$ exists.
2. f is defined at (a, b),
3. The stated equality.

Exercise 2
Let f be defined by $f(x, y) = \begin{cases} \frac{xy}{|x|+|y|} & \text{if } (x, y) \neq (0,0) \\ 0 & \text{if } (x, y) = (0,0) \end{cases}$
Determine whether f is continuous at $(0,0)$.

Solution:

$$\left|\frac{xy}{|x| + |y|}\right| \leq \left|\frac{x(|x| + |y|)}{|x| + |y|}\right| = |x|$$

Since $\lim_{x\to 0} |x| = 0$, we get $\lim_{(x,y)\to(a,b)} \frac{(xy)}{|x|+|y|} = 0$

Thus, f is continuous at $(0,0)$.

3.2 The Continuity Theorems

2019年1月18日 1:45

Definition: Operations on Functions
If $f(x,y)$ and $g(x,y)$ are scalar functions and $(x,y) \in D(f) \cap D(g)$, then:

1. The **sum** $f+g$ is defined by
$$(f+g)(x,y) = f(x,y) + g(x,y)$$
2. The **product** fg is defined by
$$(fg)(x,y) = f(x,y)g(x,y)$$
3. The **quotient** $\frac{f}{g}$ is defined by
$$\left(\frac{f}{g}\right)(x,y) = \frac{f(x,y)}{g(x,y)}, \qquad \text{if } g(x,y) \neq 0$$

Definition: Composite Function
For scalar functions $g(t)$ and $f(x,y)$ the **composite function** $g \circ f$ is defined by
$$(g \circ f)(x,y) = g\left(f(x,y)\right)$$
For all $(x,y) \in D(f)$ for which $f(x,y) \in D(g)$.

We shall refer to the following theorems collectively as the **Continuity Theorems.**

Theorem 1
If f and g are both continuous at (a,b), then $f+g$ and fg are continuous at (a,b).

Proof:
We prove the result for $f+g$ and leave the proof for fg as an exercise. By the hypothesis and the definition of continuous function we have that
$$\lim_{(x,y)\to(a,b)} f(x,y) = f(a,b)$$
$$\lim_{(x,y)\to(a,b)} g(x,y) = g(a,b)$$
Hence, by definition of the sum and limit properties, we get
$$\begin{aligned}\lim_{(x,y)\to(a,b)} (f+g)(x,y) &= \lim_{(x,y)\to(a,b)} f(x,y) + \lim_{(x,y)\to(a,b)} g(x,y) \\ &= f(a,b) + g(a,b) \\ &= (f+g)(a,b)\end{aligned}$$

Exercise 1
Complete the proof of the theorem by proving that fg is continuous at (a,b).

Theorem 2
If f and g are both continuous at (a,b) and $g(a,b) \neq 0$, then the quotient $\frac{f}{g}$ is continous at (a,b).

Exercise 2
Use the Limit Theorems to prove Theorem 2. Where is the hypothesis $g(a,b) \neq 0$ used explicitly?

Theorem 3
If $f(x,y)$ is continuous at (a,b) and $g(t)$ is continuous at $f(a,b)$, then the composition $g \circ f$ is continuous at (a,b).

Proof:
Let $\epsilon > 0$. By definition of continuity we have that

$$\lim_{t\to f(a,b)} g(t) = g\big(f(a,b)\big)$$

So, by definition of a limit there exists a $\delta_1 > 0$ such that

$$|t - f(a,b)| < \delta_1 \quad \text{implies} \quad \big|g(t) - g\big(f(a,b)\big)\big| < \epsilon$$

Similarly, we have

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

Hence, given the above δ_1, there exists a $\delta > 0$ such that

$$\big\|(x,y) - (a,b)\big\| < \delta \quad \text{implies} \quad \big|f(x,y) - f(a,b)\big| < \delta_1$$

Notice that the conclusion of (3.2) is the hypothesis of (3.1) where $t = f(x,y)$. Hence, combining (3.1) and (3.2), we get

$$\big\|(x,y) - (a,b)\big\| < \delta \;\text{ implies }\; \big|f(x,y) - f(a,b)\big| < \delta_1 \;\text{ implies }\; \Big|g\Big(f\big(x,y\big)\Big) - g\big(f(a,b)\big)\Big| < \epsilon$$

Consequently, by definition of a limit,

$$\lim_{(x,y)\to(a,b)} \big(g\circ f\big)(x,y) = \big(g\circ f\big)(a,b)$$

Before we can apply these theorems, we need a list of **basic functions** which are known to be continuous on their domains:

- The constant function $f(x,y) = k$
- The power functions $f(x,y) = x^n, f(x,y) = y^n$
- The logarithm function $\ln(\cdot)$
- The exponential function $e^{(\cdot)}$
- The trigonometric functions, $\sin(\cdot)$, $\cos(\cdot)$, etc.
- The inverse trigonometric functions, $\arcsin(\cdot)$, etc.
- The absolute value function $|\cdot|$

Exercise 3

Prove that the constant function $f(x,y) = k$ and the coordinate functions $f(x,y) = x$, $f(x,y) = y$ are continuous on their domains.

Exercise 4

Prove that $h(x,y) = (xy)^\pi$ is continous for all (x,y) which satisfy $xy > 0$. Which of the theorems and basic functions do you have to use?

Solution:

Basic functions: x^π, y^π

Theorems: Multiply

Exercise 5

Which of the basic functions and theorems do you have to use in order to prove that $h(x,y) = \frac{\sin^2|x+2y|}{x^2+y^2}$ is continuous for all $(x,y) \neq (0,0)$

Example 2

Discuss the continuity of the function f defined by

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

Solution:

For $(x,y) \neq (0,0)$ the Continuity Theorems immediately imply that f is continuous at these points.

Observe the point $(0,0)$ is singled out in the definition of the function. Thus, the Continuity Theorems cannot be appled at $(0,0)$ and so we have to use the definition. That is, we have to determine whether

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

On the line $y = x$ we get

$$\lim_{(x,y)\to(0,0)} f(x,x) = \lim_{x\to 0} \frac{\left(e^{x^2} - 1\right)}{2x^2} = \lim_{x\to 0} \frac{2xe^{x^2}}{4x} = \lim_{x\to 0} \frac{e^{x^2}}{2} = \frac{1}{2}$$

By L'Hospital's Rule. It follows that $\lim_{(x,y)\to(0,0)} f(x,y)$ does not equal $f(0,0)$, and hence by definition, f is not continuous at $(0,0)$.

Exercise 6

Would the function f in Example 2 be continuous at $(0,0)$ if we defined $f(0,0) = \frac{1}{2}$?

Example 3

Discuss the continuity of the function f defined by

$$f(x,y) = \begin{cases} \dfrac{|y-x|}{y-x} & \text{if } x \neq y \\ 0 & \text{if } x = y \end{cases}$$

Solution:

For points (x,y) with $x \neq y$ the Continuity Theorems immediately imply that f is continuous at these points.

We can not apply the continuity theorems at the points (x,y) with $x = y$. Consider any one of these points and denote it by (a,a).

If (x,y) approaches (a,a) with $y - x > 0$, then $|y - x| = y - x$, and $f(x,y)$ approaches (and in fact equals) 1. On the other hand, if (x,y) approaches (a,a) with $y - x < 0$, then $f(x,y)$ approaches -1. Thus,

$$\lim_{(x,y)\to(a,a)} f(x,y)$$

Does not exist. So, by definition of continuity, f is not continuous at (a,a).

The geometric interpretation is simple. The graph of f consists of two parallel half-planes which form a "step" along the line $y = x$.

3.3 Limits Revisited

2019年1月19日 13:21

So far in this chapter, we have shown how to prove that a function is continuous at a point essentially "by inspection" by using the Continuity Theorems. This makes it easy to evaluate $\lim_{(x,y)\to(a,b)} f(x,y)$ if f is continuous at (a,b). In particular, if f is continuous at (a,b), then $\lim_{(x,y)\to(a,b)} f(x,y)$ can be evaluated simply by evaluating $f(a,b)$.

Remark
In applying the Squeeze Theorem one has to prove that $\lim_{(x,y)\to(a,b)} B(x,y) = 0$. One hopes to be able to evaluate this limit by inspection, and so one tries to set up the inequality in the Squeeze Theorem so that $B(x,y)$ is continuous at (a,b).

Chapter 4 The Linear Approximation

2019年1月19日 13:27

4.1 Partial Derivatives

2019年1月19日 13:27

A scalar function $f(x,y)$ can be differentiated in two natural ways:

1. Treat y as a constant and differentiate with respect to x to obtain $\frac{\partial f}{\partial x}$.
2. Treat x as a constant and differentiate with respect to y to obtain $\frac{\partial f}{\partial y}$.

The derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial x}$ are called the (first) **partial derivatives** of f.
Here is the formal definition.

Definition: Partial Derivatives
The **partial derivatives** of $f(x,y)$ are defined by

$$\frac{\partial f}{\partial x} = f_x = \lim_{h\to 0} \frac{f(x+h,y) - f(x,y)}{h}$$
$$\frac{\partial f}{\partial y} = f_y = \lim_{h\to 0} \frac{f(x+h,y) - f(x,y)}{h}$$

Provided that these limits exist.

It is sometimes convenient to use **operator notation** $D_1 f$ and $D_2 f$ for the partial derivatives of $f(x,y)$. The nontation $D_1 f$ means: differentiate f with respect to the variable in the first position, holding the other fixed. If the independent variables are x and y, then

$$D_1 f = \frac{\partial f}{\partial x} = f_x, \qquad D_2 f = \frac{\partial f}{\partial y} = f_y$$

Example 1
Consider the function f defined by $f(x,y) = xe^{kxy}$ where k is a constant. Determine $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$.

Solution:
By using the Product Rule and Chain Rule for differentiation,

$$\frac{\partial f}{\partial x} = (1)e^{kxy} + xe^{kxy} = (1+kxy)e^{kxy}$$
$$\frac{\partial f}{\partial y} = xe^{kxy}(kx) = kx^2 e^{kxy}$$

Exercise 1
A function f is defined by $f(x,y) = \sin(xy^2)$. Determine f_x and f_y.

Solution:

$$\frac{\partial f}{\partial x} = \cos(xy^2) \cdot y^2$$
$$\frac{\partial f}{\partial y} = \cos(xy^2) \cdot 2xy$$

Example 2
A function f is defined by $f(x,y) = (x^3 + y^3)^{\frac{1}{3}}$. Determine whether $\frac{\partial f}{\partial x}(0,0)$ exists.
Solution:
By single-variable differentiation rules,

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

For all (x,y) such that $x^3 + y^3 \neq 0$. One cannot substitute $(x,y) = (0,0)$ in equation (4.1) since the denominator would be zero. Thus, we must use the definition of the partial derivatives at $(0,0)$. We get

$$\begin{aligned}\frac{\partial f}{\partial x}(0,0) &= \lim_{h\to 0}\frac{f(0+h,0)-f(0,0)}{h}\\ &= \lim_{h\to 0}\frac{\left(h^3+0^3\right)^{\frac{1}{3}}-0}{h}\\ &= \lim_{h\to 0} 1 = 1\end{aligned}$$

Example 3

Let $f(x,y)=\begin{cases}\frac{xy}{x^2+y^2} & \text{if } (x,y)\neq(0,0)\\ 0 & \text{if } (x,y)=(0,0)\end{cases}$. Calculate $f_x(0,0)$ and $f_y(0,0)$.

Solution: Since f changes definition at $(0,0)$, we must use the definition of the partial derivaties. We get

$$f_x(0,0)=\lim_{h\to 0}\frac{f(0+h,0)-f(0,0)}{h}=\lim_{h\to 0}\frac{\frac{h(0)}{h^2+0^2}-0}{h}=0$$

$$f_y(0,0)=\lim_{h\to 0}\frac{f(0,0+h)-f(0,0)}{h}=\lim_{h\to 0}\frac{\frac{0(h)}{0^2+h^2}-0}{h}=0$$

Remark

In Example 2, we showed that $f(x,y)=\begin{cases}\frac{xy}{x^2+y^2} & \text{if } (x,y)\neq(0,0)\\ 0 & \text{if } (x,y)=(0,0)\end{cases}$ is not continuous at $(0,0)$, but we have just shown that its partial derivatives exist! This demonstrates that the concept of partial derivatives do not match our concept of differentiability for functions of one variable from Calculus 1. We will look at this more in the next chapter.

Exercise 2

Refer to the function in Example 2. Show that $\frac{\partial f}{\partial x}(a,-a)$ does not exist for $a\neq 0$.

Exercise 3

A function f is defined by $f(x,y)=\left|x(y-1)\right|$. Determine whether $\frac{\partial f}{\partial x}(0,0)$ and $\frac{\partial f}{\partial x}(0,1)$ exist.

Generalization

We can extend what we have done for scalar functions of two variables to scalar functions of n variables $f(\mathbf{x}), \mathbf{x}\in\mathbb{R}^n$. To take the partial derivative of f with respect to its i-th variable, we hold all the other variables constant and differentiate with respect to the i-th variable.

Example 4

Let $f(x,y,z)=xy^2z^3$. Find f_x, f_y, and f_z.

Solution:

We have

$$\begin{aligned}f_x(x,y,z)&=y^2z^3\\ f_y(x,y,z)&=2xyz^3\\ f_z(x,y,z)&=3xy^2z^2\end{aligned}$$

Exercise 4

For $f(x,y,z)$, write the precise definition of $\frac{\partial f}{\partial x},\frac{\partial f}{\partial y},\frac{\partial f}{\partial z}$.

4.2 Higher-Order Partial Derivatives

2019年1月19日 18:32

Second Partial Derivatives

Observe that the partial derivatives of a scalar function of two variables are both scalar functions of two variables. Therefore, we can take the partial derivatives of the partial derivatives of any scalar function.

In how many ways can one calculate a second partial derivative of $f(x,y)$? Since both of the partial derivatives of f have two partial derivatives, there are four possible second partial derivatives of f. They are:

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right), \qquad \text{i.e. differentiate } \frac{\partial f}{\partial x} \text{ with respect to } x, \text{with } y \text{ fixed.}$$

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right), \qquad \text{i.e. Differentiate } \frac{\partial f}{\partial x} \text{ with respect to } y, \text{with } x \text{ fixed.}$$

Similarly

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right), \qquad \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)$$

It is often convenient to use the subscript notation or the operator notation:

$$\frac{\partial^2 f}{\partial x^2} = f_{xx} = D_1^2 f, \qquad \frac{\partial^2 f}{\partial y \partial x} = f_{xy} = D_2 D_1 f$$

$$\frac{\partial^2 f}{\partial x \partial y} = f_{yx} = D_1 D_2 f, \qquad \frac{\partial^2 f}{\partial y^2} = f_{yy} = D_2^2 f$$

The subscript notation suggests that one could write the second partial derivatives in a 2×2 matrix.

Definition: Hessian Matrix

The **Hessian matrix** of $f(x,y)$, denoted by $Hf(x,y)$, is defined as

$$Hf(x,y) = \begin{bmatrix} f_{xx}(x,y) & f_{xy}(x,y) \\ f_{yx}(x,y) & f_{yy}(x,y) \end{bmatrix}$$

Example 1

Let k be a constant. Find all the second partial derivatives of $f(x,y) = xe^{kxy}$.

Solution:

We first calculate the first partial derivatives. We have

$$\frac{\partial f}{\partial x}(x,y) = e^{kxy} + kxye^{kxy}$$

$$\frac{\partial f}{\partial y}(x,y) = kx^2 e^{kxy}$$

Thus, we get

$$\frac{\partial^2 f}{\partial x^2}(x,y) = \frac{\partial}{\partial x}\left[e^{kxy} + kxye^{kxy}\right] = 2kye^{kxy} + k^2xy^2e^{kxy}$$

$$\frac{\partial^2 f}{\partial y \partial x}(x,y) = \frac{\partial}{\partial y}\left[e^{kxy} + kxye^{kxy}\right] = 2kxe^{kxy} + k^2x^2ye^{kxy}$$

$$\frac{\partial^2 f}{\partial x \partial y}(x,y) = \frac{\partial}{\partial x}\left[kx^2e^{kxy}\right] = 2kxe^{kxy} + k^2x^2ye^{kxy}$$

$$\frac{\partial^2 f}{\partial y^2}(x,y) = \frac{\partial}{\partial y}\left[kx^2e^{kxy}\right] = k^2x^3e^{kxy}$$

In the previous example, observe that

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$$

This is a fact a general property of partial derivatives, subject to a continuity requirement, as follows.

Theorem 1
(Clairaut's Theorem)

If f_{xy} and f_{yx} are defined in some neighbourhood of (a,b) and are both continuous (a,b), then

$$f_{xy}(a,b) = f_{yx}(a,b)$$

Exercise 1
Verify that $f(x,y) = \ln(x^2 + y^2)$ satisfies

$$f_{xx} + f_{yy} = 0, \qquad \text{for } (x,y) \neq (0,0)$$

Exercise 2
Verify that $f(x,y) = x^y$ satisfies

$$f_{xy} = f_{yx}, \qquad \text{for } x > 0.$$

Higher-Order Partial Derivatives

If the k-th partial derivatives of $f(x_1, \ldots, x_n)$ are continuous, then we write

$$f \in C^k$$

And say "f is in class C^k."

So, $f(x,y) \in C^2$ means that f has continuous second partial derivatives, and therefore, by Clairaut's Theorem, we have that $f_{xy} = f_{yx}$.

4.3 The Tangent Plane

2019年1月19日 20:32

Definition: Tangent Plane

The **tangent plane** to $z = f(x, y)$ at the point $(a, b, f(a, b))$ is

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

Exercise 1

The graph of the function

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

Is the cone $z = \sqrt{x^2 + y^2}$. Find the equation of the tangent plane at the point $(3, -4, 5)$

Exercise 2

Show that the tangent plane at any point on the cone in Exercise 1 passes through the origin.

Remark

In Exercise 2, you should note that a tangent plane does not exist at the vertex $(0,0,0)$ of the cone, since the cone is not "smooth" there. We shall discuss the question of the existence of a tangent plane in Chapter 5.

4.4 Linear Approximation for $z = f(x,y)$

2019年1月20日 14:39

Review of 1-D case

For a function $f(x)$ the tangent line can be used to approxiamte the graph of the function near the point of tangency. Recall that the equation of the tangent line to $y = f(x)$ at the point $(a, f(a))$ is

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

The function L_a defined by

$$L_a(x) = f(a) + f'(a)(x-a)$$

Is called the **linearization** of f at a since $L_a(x)$ approximates $f(x)$ for x sufficiently close to a.

For x sufficiently close to a, the approximation

$$f(x) \approx L_a(x)$$

Is called the **linear approximation** of f at a.

Exercise 1

Verify each approximation:

1. $\sin x \approx x$, for x sufficiently close to 0.
2. $\sqrt{1+x} \approx 1 + \frac{1}{2}x$, for x sufficiently close to 0
3. $\ln x \approx (x-1)$, for x sufficiently close to 1.

The 2-D case

For a function $f(x,y)$, the tangent plane can be used to approximate the surface $z = f(x,y)$ near the point of tangency.

Definition: Linearization Linear Approximation

For a function $f(x,y)$ we define the **linearization** $L_{(a,b)}(x,y)$ of f at (a,b) by

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

We call the approximation

$$f(x,y) \approx L_{(a,b)}(x,y)$$

The **linear approximation** of $f(x,y)$ at (a,b).

Increment Form of the Linear Approximation

Suppose that we know $f(a,b)$ and want to calculate $f(x,y)$ at a nearby point. Let

$$\Delta x = x - a, \qquad \Delta y = y - b$$

And

$$\Delta f = f(x,y) - f(a,b)$$

The linear approximation is

$$f(x,y) \approx f(a,b) + \frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(y-b)$$

This can be rearranged to yield

$$\Delta f \approx \frac{\partial f}{\partial x}(a,b)\Delta x + \frac{\partial f}{\partial y}(a,b)\Delta y$$

This gives an approximation for the change Δf in $f(x,y)$ due to a change $(\Delta x, \Delta y)$ away from the point (a,b).

We shall refer to equation (4.4) as the **increment form** of the linear approximation.

4.5 Linear Approximation in Higher Dimensions

2019年1月20日 15:10

Linear Approximation in $\mathbb{R}^3$

Consider a function $f(x,y,z)$. By analogy with the case of a function of two variables, we define the linearization of f at $\mathbf{a}=(a,b,c)$ by

$$L_{\mathbf{a}}(x,y,z)=f(\mathbf{a})+f_x(\mathbf{a})(x-a)+f_y(\mathbf{a})(y-b)+f_z(\mathbf{a})(z-c)$$

The notation is becoming cumbersome, but one can improve matters by noting that the final three terms can be represented by the dot product of the vectors

$$(x-a,y-b,z-c)=(x,y,z)-(a,b,c),\qquad \text{and}\quad \nabla f(\mathbf{a})=\left(f_x(\mathbf{a}),f_y(\mathbf{a}),f_z(\mathbf{a})\right)$$

The second vector is called the **gradient** of f at **a.**
Here are the formal definitions.

Definition: Gradient
Suppose that $f(x,y,z)$ has partial derivatives at $\mathbf{a}\in\mathbb{R}^3$. The **gradient** of f at **a** is defined by

$$\nabla f(\mathbf{a})=\left(f_x(\mathbf{a}),f_y(\mathbf{a}),f_z(\mathbf{a})\right)$$

Definition: Linearization Linear Approximation
Suppose that $f(\mathbf{x}),\mathbf{x}\in\mathbb{R}^3$, has partial derivatives at $\mathbf{a}\in\mathbb{R}^3$. The **linearization** of f at **a** is defined by

$$L_{\mathbf{a}}(\mathbf{x})=f(\mathbf{a})+\nabla f(\mathbf{a})\cdot(\mathbf{x}-\mathbf{a})$$

The linear approximation of f at **a** is

$$f(\mathbf{x})\approx f(\mathbf{a})+\nabla f(\mathbf{a})\cdot(\mathbf{x}-\mathbf{a})$$

Linear Approximation in $\mathbb{R}^n$
The advantage of using vector notation is that equations (4.5) and (4.6) hold for a function of n variables $f(\mathbf{x}),\mathbf{x}\in\mathbb{R}^{\mathrm{n}}$. For arbitrary $\mathbf{a}\in\mathbb{R}^{\mathrm{n}}$, we have

$$\mathbf{x}-\mathbf{a}=(x_1-a_1,x_2-a_2,\dots,x_n-a_n)$$

And we define the gradient of f at **a** to be

$$\nabla f(\mathbf{a})=(D_1f(\mathbf{a}),D_2f(\mathbf{a}),\dots,D_nf(\mathbf{a}))$$

Then, the increment form of the linear approximation for $f(\mathbf{x})$ is

$$\Delta f\approx\nabla f(\mathbf{a})\cdot\Delta\mathbf{x}$$

Observe that this formula even works when $n=1$. That is, for a function $g(t)$ of one variable this gives $\nabla g(a)=g'(a)$ and the increment form of the linear approximation is

$$\Delta g\approx\nabla g(a)\cdot\Delta x=g'(a)(x-a)$$

Which is our familiar formula from Calculus 1.
For $f(x,y)$ we have $\nabla f(a,b)=\left(f_x(a,b),f_y(a,b)\right)$ and the increment form of the linear approximation is

$$\Delta f\approx\nabla f(a,b)\cdot\Delta(x,y)=f_x(a,b)(x-a)+f_y(a,b)(y-b)$$

Which matches our work above. Hence, we see that this is a true generalization.

Chapter 5 Differentiable Functions

2019年1月21日 18:25

5.1 Definition of Differentiability

2019年1月21日　18:25

Theorem 1

If $g'(a)$ exists, then $\lim_{x\to a}\frac{|R_{1,a}(x)|}{|x-a|}=0$ where

$$R_{1,a}(x)=g(x)-L_a(x)=g(x)-g(a)-g'(a)(x-a)$$

Proof:

We have

$$\frac{|R_{1,a}(x)|}{|x-a|}=\left|\frac{g(x)-g(a)-g'(a)(x-a)}{x-a}\right|=\left|\frac{g(x)-g(a)}{x-a}-g'(a)\right|$$

The result follows from taking the limit as $x\to a$ (details left as an exercise).

Definition: Differentiable

A function $f(x,y)$ is **differentiable** at (a,b) if

$$\lim_{(x,y)\to(a,b)}\frac{|R_{1,(a,b)}(x,y)|}{\|(x,y)-(a,b)\|}=0$$

Where

$$R_{1,(a,b)}(x,y)=f(x,y)-L_{(a,b)}(x,y)$$

Theorem 2

If a function $f(x,y)$ satisfies

$$\lim_{(x,y)\to(a,b)}\frac{|f(x,y)-f(a,b)-c(x-a)-d(y-b)|}{\|(x,y)-(a,b)\|}=0$$

Then $c=f_x(a,b)$ and $d=f_y(a,b)$.

Proof:

Since

$$\lim_{(x,y)\to(a,b)}\frac{|f(x,y)-f(a,b)-c(x-a)-d(y-b)|}{\|(x,y)-(a,b)\|}=0$$

The limit is 0 along any path. Therefore, along the path along $y=b$, we get

$$\begin{aligned}0&=\lim_{x\to a}\left(\frac{|f(x,b)-f(a,b)-c(x-a)-d(b-b)|}{\|(x,b)-(a,b)\|}\right)\\&=\lim_{x\to a}\frac{(|f(x,b)-f(a,b)-c(x-a)|)}{|x-a|}\\&=\lim_{x\to a}\left|\frac{f(x,b)-f(a,b)}{x-a}-c\right|\\&=f_x(a,b)-c\\c&=f_x(a,b)\end{aligned}$$

Similarly, approaching along $x=a$ we get that $d=f_y(a,b)$.

This implies that the tangent plane gives the best linear approximation to the graph $z=f(x,y)$ near (a,b). Moreover, it tells us that the linear approximation is a "good approximation" if and only if f is differentiable at (a,b).

Remark

Observe that for the linear approximation to exist at (a,b) both partial derivatives of f must exist at (a,b). However, both partial derivatives existing does not guarentee that f will be differentiable. We say that the partial derivatives of f existing at (a,b) is necessary, but not sufficient.

Exercise 1

Prove that

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

Is not differentiable at (0,0).

Solution:

We have $f_x = \frac{3x^2(x^2+y^2)-x^3(2x)}{(x^2+y^2)^2} = \frac{x^4+3x^2y^2}{(x^2+y^2)^2}$ and $f_y = \frac{-x^3(2y)}{(x^2+y^2)^2} = -\frac{2x^3y}{(x^2+y^2)^2}$,

Wrong Procedure

$$f_x(0,0) = \lim_{h\to 0}\frac{f(0+h,0)-f(0,0)}{h} = \lim_{h\to 0}\frac{h^4+3h^2(0)}{(h^2+0)^2} = \lim_{h\to 0} 1 = 1$$

$$f_y(0,0) = \lim_{h\to 0}\frac{f(0,0+h)-f(0,0)}{h} = \lim_{h\to 0} -\frac{0}{h^4} = 0$$

Although the procedure is not correct, it can be imaged as plug in $(h,0)$ for f_x and let the $\lim_{h\to 0}$. Same with f_y.

$$f_x(0,0) = \lim_{h\to 0}\frac{f(0+h,0)-f(0,0)}{h} = \frac{\left(\frac{h^3}{h^2}-0\right)}{h} = 1$$

$$f_y(0,0) = \lim_{h\to 0}\frac{f(0,0+h)-f(0,0)}{h} = \frac{0-0}{h} = 0$$

So the error in the linear approximation is

$$R_{1,(0,0)}(x,y) = f(x,y) - f(0,0) - f_x(0,0)(x-0) - f_y(0,0)(y-0) = \frac{x^3}{x^2+y^2} - x = \frac{x^3-x^2-y^2}{x^2+y^2}$$

For f to be differentiable at (0,0) we need $\lim_{(x,y)\to(0,0)}\frac{|R_{1,(0,0)}(x,y)|}{\sqrt{x^2+y^2}} = 0$.

If we approach the limit along $y = x$, we get

$$\begin{aligned}\lim_{x\to 0}\frac{|R_{1,(0,0)}(x,x)|}{\sqrt{x^2+x^2}} &= \lim_{x\to 0}\frac{\left|\frac{x^3-2x^2}{2x^2}\right|}{\sqrt{2x^2}} \\ &= \lim_{x\to 0}\left|\frac{x^3-2x^2}{2\sqrt{2}x^3}\right| \\ &= \lim_{x\to 0}\left|\frac{1}{2\sqrt{2}}\right| \\ &= \frac{1}{2\sqrt{2}}\end{aligned}$$

Therefore, the limit cannot equal 0 and hence f is not differentiable at (0,0).

Exercise 2

Prove that $f(x,y) = |xy|$ is differentiable at (0,0).

Solution:

$$f_x(0,0) = \lim_{h\to 0}\frac{f(0+h,0)-f(0,0)}{h} = \lim_{h\to 0}\frac{0-0}{h} = 0$$

$$f_y(0,0) = \lim_{h\to 0}\frac{(f(0,0+h)-f(0,0))}{h} = \lim_{h\to 0}\frac{0-0}{h} = 0$$

So the error in the linear approximation is

$$R_{1,(0,0)}(x,y) = f(x,y) - f(0,0) - f_x(0,0)(x-0) - f_y(0,0)(y-0) = |xy|$$

For f to be differentiable at $(0,0)$ we need $\lim_{(x,y)\to(0,0)} \frac{|R_{1,(0,0)}(x,y)|}{\sqrt{x^2+y^2}} = 0$.

If we approach the limit along $y = x$, we get

$$\begin{aligned}\lim_{x\to 0} \frac{|R_{1,(0,0)}(x,x)|}{\sqrt{x^2+x^2}} &= \lim_{x\to 0} \frac{|x^2|}{\sqrt{2x^2}} \\ &= \lim_{x\to 0} \frac{|x|}{\sqrt{2}} \\ &= 0\end{aligned}$$

Therefore, $f(x,y)$ is differentiable at $(0,0)$.

Exercise 3

Prove that $f(x,y) = |xy|$ is not differentiable at $(0,1)$.

Solution:

$$f_x(0,1) = \lim_{h\to 0} \frac{f(0+h,1) - f(0,1)}{h} = \lim_{h\to 0} \frac{|h|}{h}$$

Partial derivatives not exist ⇒ not differentiable? Seems right, but the proof?

The limit does not exist.

Thus, $f(x,y) = |xy|$ is not differentiable at $(0,1)$.

Definition: **Tangent Plane**

Consider a function $f(x,y)$ which is differentiable at (a,b). The **tangent plane** of the surface $z = f(x,y)$ at $(a,b,f(a,b))$ is the graph of the linearization. That is, the tangent plane is given by

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

Since f is assumed to be differentiable at (a,b), by Theorem 2, the tangent plane is the plane that best approximates the surface near the point $(a,b,f(a,b))$. In this case, we say that at the point $(a,b,f(a,b))$ the surface $z = f(x,y)$ is **smooth**.

5.2 Differentiability and Continuity

2019年1月31日 21:21

Theorem 1
If $f(x, y)$ is differentiable at (a, b), then f is continuous at (a, b).

Proof:
The error $R_{1,(a,b)}(x, y)$ is defined by

$$R_{1,(a,b)}(x, y) = f(x, y) - L_{(a,b)}(x, y)$$

Using the definition of $L_{(a,b)}(x, y)$, this equation can be rearranged to read

$$f(x, y) = f(a, b) + \nabla f(a, b) \cdot (x - a, y - b) + R_{1,(a,b)}(x, y)$$

We can write

$$R_{1,(a,b)}(x, y) = \frac{R_{1,(a,b)}(x, y)}{\|(x, y) - (a, b)\|} \|(x, y) - (a, b)\|, \qquad \text{for } \ (x, y) \neq (a, b)$$

Since f is differentiable and by the limit theorems, we get

$$\lim_{(x,y)\to(a,b)} f(x, y) = f(a, b) + 0 + 0 = f(a, b)$$

And so by definition, f is continuous at (a, b).

Exercise 1
Suppose that $f(x, y)$ is not continuous at (a, b). Can you draw a conclusion about whether f is differentiable at (a, b)?

Solution:
Take the contrapositive of Theorem 1, we get

If $f(x, y)$ is not continuous at (a, b), f is not differentiable at (a, b).

5.3 Continuous Partial Derivatives and Differentiability

2019年1月31日 21:37

Theorem 1 (Mean Value Theorem)
If $f(t)$ is continuous on the closed interval $[t_1, t_2]$ and f is differentiable on the open interval (t_1, t_2), then there exists $t_0 \in (t_1, t_2)$ such that

$$f(t_2) - f(t_1) = f'(t_0)(t_2 - t_1)$$

Theorem 2
If $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ are continuous at (a, b), then $f(x, y)$ is differentiable at (a, b).

Proof:
We derive an expression for the error $R_{1,(a,b)}(x, y)$, given by

$$R_{1,(a,b)}(x, y) = f(x, y) - f(a, b) - f_x(a, b)(x - a) - f_y(a, b)(y - b)$$

Since f_x and f_y are continuous then f_x and f_y exist in some neighborhood $B(a, b)$. For $(x, y) \in B(a, b)$, we write

$$f(x, y) - f(a, b) = [f(x, y) - f(a, y)] + [f(a, y) - f(a, b)]$$

By adding and subtracting $f(a, y)$. The Mean Value Theorem can be applied to each bracket, since one variable is held fixed, and the partial derivatives are assumed to exist. For the first bracket:

$$f(x, y) - f(a, y) = f_x(\bar{x}, y)(x - a)$$

Where $\bar{x}$ lies between a and x. By adding and substracting $f_x(a, b)(x - a)$, we obtain

$$f(x, y) - f(a, y) = f_x(a, b)(x - a) + A(x - a)$$

Where

$$A = f_x(\bar{x}, y) - f_x(a, b)$$

Similarly for the second bracket:

$$\begin{aligned} f(a, y) - f(a, b) &= f_y(a, \bar{y})(y - b) \\ &= f_y(a, b)(y - b) + B(y - b) \end{aligned}$$

Where

$$B = f_y(a, \bar{y}) - f_y(a, b)$$

And $\bar{y}$ lies between b and y.
Substitute equations (5.5) and (5.7) into (5.4) and then substitute equation (5.4) into (5.3) to obtain

$$R_{1,(a,b)}(x, y) = A(x - a) + B(y - b)$$

Where A and B are given by equations (5.6) and (5.8). It follows by the triangle inequality that

$$\begin{aligned} \frac{|R_{1,(a,b)}(x, y)|}{\|(x, y) - (a, b)\|} &\le \frac{|A||x - a|}{\sqrt{(x - a)^2 + (y - b)^2}} + \frac{|B||y - b|}{\sqrt{(x - a)^2 + (y - b)^2}} \\ &\le |A| + |B| \end{aligned}$$

We can now apply the Squeeze Theorem with $L = 0$ and $B(x, y) = |A| + |B|$.
As $(x, y) \to (a, b)$, it follows that

$$(\bar{x}, y) \to (a, b) \quad \text{and} \quad (a, \bar{y}) \to (a, b)$$

Since f_x and f_y are continuous at (a, b), it follows from equations (5.6) and (5.8) that

$$\lim_{(x,y)\to(a,b)} A = 0 \quad \text{and} \quad \lim_{(x,y)\to(a,b)} B = 0$$

Equation (5.9) and the Squeeze Theorem now imply

$$\lim_{(x,y)\to(a,b)} \frac{|R_{1,(a,b)}(x, y)|}{\|(x, y) - (a, b)\|} = 0$$

So that f is differentiable at (a, b), by definition.

Remark
The converse of Theorem 2 is not true. That is, $f(x, y)$ being differentiable at (a, b) does not imply that f_x and f_y are both continuous at (a, b).

Exercise 1

Prove that $f(x, y) = (x^2 + y^2)^{\frac{2}{3}}$ is differentiable at (0,0).

Solution:
By differentiation

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0}\frac{f(0+h,0) - f(0,0)}{h} = \lim_{h\to 0}\frac{(h^2)^{\frac{2}{3}}}{h} = \lim_{h\to 0} h^{\frac{1}{3}} = 0$$
$$\frac{\partial f}{\partial y}(0,0) = \lim_{h\to 0}\frac{f(0,h+0) - f(0,0)}{h} = \lim_{h\to 0}\frac{(h^2)^{\frac{2}{3}}}{h} = \lim_{h\to 0} h^{\frac{1}{3}} = 0$$

$$R_{1,(0,0)}(x, y) = f(x, y) - f(0,0) - \frac{\partial f}{\partial x}(0,0)(x - 0) - \frac{\partial f}{\partial y}(0,0)(y - 0) = (x^2 + y^2)^{\frac{2}{3}}$$
$$\lim_{(x,y)\to(0,0)}\frac{|R_{1,(0,0)}(x, y)|}{\sqrt{x^2 + y^2}} = \lim_{(x,y)\to(0,0)}\frac{(x^2 + y^2)^{\frac{2}{3}}}{(x^2 + y^2)^{\frac{1}{2}}}$$
$$= \lim_{(x,y)\to(0,0)}(x^2 + y^2)^{\frac{1}{6}}$$
$$= 0$$
Thus, $f(x, y)$ is differentiable at (0,0).

Exercise 2
Prove that if $f(x, y) \in C^2$ at (a, b), then f is continuous at (a, b).

Solution:
Since $f_{xx}, f_{xy}, f_{yx}, f_{yy}$ are continuous at (a, b).

f_x and f_y are differentiable at (a, b)
Then f_x and f_y are continuous at (a, b)
Then f is differentiable at (a, b).
Then f is continuous at (a, b).

Summary
Theorem 2 makes it easy to prove that a function f is differentiable at a typical point. One simply differentiates f to obtain the partial derivatives f_x, f_y, and then checks that the partials are continuous functions by inspection, referring to the Continuity Theorems, as in Section 3.2. It is only necessary to use the definition of a differentiable function at an exceptional point.

Generalization
The definition of a differentiable function and theorems 1 and 2 are valid for functions of n variables. The only change is that there are n partial derivatives,

$$\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n}$$

5.4 Linear Approximation Revisited

2019年2月8日 22:16

The error in the linear approximation for $f(x,y)$ is defined by

$$R_{1,(a,b)}(x,y) = f(x,y) - L_{(a,b)}(x,y)$$

Where

$$L_{(a,b)}(x,y) = f(a,b) + \nabla f(a,b) \cdot \left((x,y) - (a,b)\right)$$

It is convenient to rearrange the definition of $R_{1,(a,b)}(x,y)$ to read

$$f(x,y) = f(a,b) + \nabla f(a,b) \cdot (x-a, y-b) + R_{1,(a,b)}(x,y)$$

The linear approximation

$$f(x,y) \approx f(a,b) + \nabla f(a,b) \cdot (x-a, y-b)$$

For (x,y) sufficiently close to (a,b), arises if one neglects the error term. In general, one has no information about $R_{1,(a,b)}(x,y)$, and so it is not clear whether the approximation is reasonable. However, Theorem 2 provides an important piece of information about $R_{1,(a,b)}(x,y)$, namely that if the partial derivatives of f are continuous at (a,b), then f is differentiable and hence

$$\lim_{(x,y)\to(a,b)} \frac{\left|R_{1,(a,b)}(x,y)\right|}{\left\|(x,y) - (a,b)\right\|} = 0$$

In this case, the approximation (5.11) is reasonable for (x,y) sufficiently close to (a,b), and we say that $L_{(a,b)}(x,y)$ is a good approximation of $f(x,y)$ near (a,b).

Chapter 6 The Chain Rule

2019年2月9日 14:08

6.1 Basic Chain Rule in Two Dimensions

2019年2月9日 14:08

Theorem 1 (Chain Rule)
Let $z = f(x(t), y(t))$, and let $a = x(t_0)$ and $b = y(t_0)$. If f is differentiable at (a, b) and $x'(t_0)$ and $y'(t_0)$ exist, then $G'(t_0)$ exists and is given by

$$G'(t_0) = f_x(a,b)x'(t_0) + f_y(a,b)y'(t_0)$$

Proof:
By definition of the derivative,

$$G'(t_0) = \lim_{t\to t_0} \frac{G(t) - G(t_0)}{t - t_0}$$

Provided that this limit exists. By definition of $G(t)$,

$$G(t) - G(t_0) = f(x(t), y(t)) - f\left(x(t_0), y(t_0)\right)$$

Since f is differentiable we can write

$$f(x,y) = f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b) + R_{1,(a,b)}(x,y)$$

Where

$$\lim_{(x,y)\to(a,b)} \frac{\left|R_{1,(a,b)}(x,y)\right|}{\sqrt{(x-a)^2 + (y-b)^2}} = 0$$

Since $a = x(t_0)$, $b = y(t_0)$, it follows from equations (6.7) and (6.8) that

$$\frac{G(t) - G(t_0)}{t - t_0} = f_x(a,b)\left[\frac{x(t) - x(t_0)}{t - t_0}\right] + f_y(a,b)\left[\frac{y(t) - y(t_0)}{t - t_0}\right] + \frac{R_{1,(a,b)}(x(t), y(t))}{t - t_0}$$

You can now see the Chain Rule taking shape. We have to prove that

$$\lim_{t\to t_0} \frac{\left|R_{1,(a,b)}(x(t), y(t))\right|}{|t - t_0|} = 0$$

Define $E(x,y)$ by

$$E(x,y) = \begin{cases} \dfrac{R_{1,(a,b)}(x,y)}{\sqrt{(x-a)^2 + (y-b)^2}} & \text{if } (x,y) \neq (a,b) \\ 0 & \text{if } (x,y) = (a,b) \end{cases}$$

By equation (6.9) and the definition of continuity, E is continuous at (a,b).
From the definition of E,

$$R_{1,(a,b)}(x,y) = E(x,y)\sqrt{(x-a)^2 + (y-b)^2} \quad \text{for all } (x,y)$$

Since $a = x(t_0)$, and $b = y(t_0)$,

$$\frac{\left|R_{1,(a,b)}(x(t), y(t))\right|}{|t - t_0|} = |E(x(t), y(t))| \sqrt{\left[\frac{x(t) - x(t_0)}{t - t_0}\right]^2 + \left[\frac{y(t) - y(t_0)}{t - t_0}\right]^2}$$

Since $x'(t_0)$ and $y'(t_0)$ exist and the fact that E is continuous at (a,b) we get

$$\lim_{t\to t_0} \frac{\left|R_{1,(a,b)}(x(t), y(t))\right|}{|t - t_0|} = E\left(x(t_0), y(t_0)\right)\sqrt{[x'(t_0)]^2 + [y'(t_0)]^2} = 0$$

Since $E(a,b) = 0$.
It now follows from equation (6.6) and (6.10) that $G'(t_0)$ exists, and is given by the desired chain rule formula.

Remark
When first studying the Chain Rule you might think that hypothesis that f is differentiable could be replaced by the weaker hypothesis that $f_x(a,b)$ and $f_y(a,b)$ exist. Exercise 1 shows that this

is not the case.

Exercise 1
With reference to the theorem, let

$$f(x,y) = (xy)^{\frac{1}{3}}, \qquad x(t) = t, \qquad y(t) = t^2$$

Define $G(t) = f(x(t), y(t))$ and show that $G'(0) = 1$. Further show that $f_x(0,0) = 0$ and $f_y(0,0) = 0$, so that the Chain Rule fails. Draw a conclusion about f at $(0,0)$.

Solution:
One simple way to calculate $G'(0)$ is to substitute.

$$f(x(t), y(t)) = t$$
$$G'(0) = f'(0) = 1$$

$$f_x(0,0) = \lim_{h\to 0} \frac{f(0+h,0) - f(0,0)}{h} = \lim_{h\to 0} \frac{0-0}{h} = 0$$

By symmetry, $f_y(0,0) = 0$

Thus, the Chain Rule fails.

Not sure what this means!

Sample Answer: f is not differentiable at $(0,0)$

Remark
In practice it is convenient to use stronger hypotheses in the Chain Rule. In particular, we usually assume that f has continuous partial derivatives at (a, b) and $x'(t)$ and $y'(t)$ are both continuous at t_0. This also allows one to obtain the stronger conclusion that $G'(t)$ is continuous at t_0. These hypotheses can usually be checked quickly, either by using the Continuity Theorems, or in more theoretical situations, by using given information.

Exercise 2 **Be sure to check it again!!!!**
Stupid mistake, $y(t)$ should be $2e^t \cos t$. So that's fine.
Let

~~$T(t) = \ln(1 + x^2 + y^2)$, with $x(t) = e^t \sin t$, $y(t) = 2e^t \sin t$~~

$$T(t) = \ln(1 + x^2 + y^2), \qquad \text{with} \quad x(t) = e^t \sin t, \qquad y(t) = 2e^t \cos t$$

Calculate $\frac{dT}{dt}$ when $t = 0$ in two ways, firstly by substituting x and y in T, and secondly by evaluating $\frac{dx}{dt}(0), \frac{dy}{dt}(0), \frac{\partial T}{\partial x}(0,2)$ and $\frac{\partial T}{\partial y}(0,2)$, and applying the Chain Rule.

Solution:
First, we substitute x and y in T,

$$T(t) = \ln(1 + e^{2t} \sin^2 t + 4e^{2t} \sin^2 t) = \ln(1 + 5e^{2t} \sin^2 t)$$

$$\left.\frac{dT}{dt}\right|_{t=0} = \frac{1}{1 + 5e^{2t} \sin^2 t}\left((5e^{2t})'(\sin^2 t) + (5e^{2t})(\sin^2 t)'\right)$$
$$= \frac{1}{1}(0 + 5 \cdot (2 \sin t) \cdot (\cos t))$$
$$= 0$$

$$\frac{dx}{dt}(0) = e^t \sin t + e^t(\cos t) = 1$$
$$\frac{dy}{dt}(0) = 2$$

$$\frac{\partial T}{\partial x}(0,2) = \lim_{h\to 0}\frac{T(0+h,2)-T(0,2)}{h} = \lim_{h\to 0}\frac{\ln(5+h^2)-\ln 5}{h} = \lim_{h\to 0}\frac{\ln\left(1+\frac{h^2}{5}\right)}{h}$$

Using L'Hospital's Rule

$$\lim_{h\to 0}\frac{\left(\frac{1}{1+\frac{h^2}{5}}\right)\left(\frac{2h}{5}\right)}{1} = 0$$

$$\frac{\partial T}{\partial y}(0,2) = \lim_{h\to 0}\frac{T(0,2+h)-T(0,2)}{h} = \lim_{h\to 0}\frac{\ln(1+(2+h)^2)-\ln 5}{h}$$

Same, using L'Hospital's Rule

$$\lim_{h\to 0}\left(\frac{\frac{1}{1+(2+h)^2}(2(2+h))}{1}\right) = \frac{4}{5}$$

The Chain Rule indicates

$$\frac{dT}{dt} = \frac{\partial T}{\partial x}\frac{dx}{dt} + \frac{\partial T}{\partial y}\frac{dy}{dt}$$

$$\frac{dT}{dt}(0) = 0 + \frac{4}{5}\cdot 2 = \frac{8}{5}$$

But what???

Exercise 3

Define $f(t) = g(1+t^2, 1-t^2)$. If $\nabla g(2,0) = (3,4)$, find $f'(1)$. What condition on g will guarantee the validity of your work?

Solution:

$g = g(u,v)$. Thus, we have $f(t) = g(u(t), v(t))$ where $u(t) = 1+t^2$ and $v(t) = 1-t^2$.
Next, to apply the Chain Rule, we require that f is differentiable.
Assuming this condition, we get

$$\begin{aligned} f'(t) &= g_x(u(t),v(t))u'(t) + g_y(u(t),v(t))v'(t) \\ &= g_x(u(t),v(t))(2t) + g_y(u(t),v(t))(-2t) \end{aligned}$$

Taking $t = 1$ gives

$$f'(1) = g_x(u(0),v(0))(2) + g_y(u(0),v(0))(-2) = 3\cdot 2 + 4\cdot(-2) = -2$$

We need to assume that g is differentiable at $(2,0)$

Exercise 4

A differentiable function $f(x,y)$ is given, and $g(t)$ is defined by

$$g(t) = f(x,y)$$

Where $x(t) = \cos t$ and $y(t) = \sin t$. Write out the Chain Rule for $g'(t)$. Calculate $g'\left(\frac{\pi}{3}\right)$, if $\nabla f\left(\frac{1}{2},\frac{\sqrt{3}}{2}\right) = (\sqrt{3}, 4)$.

Solution:

$$\begin{aligned} g'(t) &= f_x(x(t),y(t))x'(t) + f_y(x(t),y(t))y'(t) \\ &= f_x(x(t),y(t))(-\sin t) + f_y(x(t),y(t))(\cos t) \end{aligned}$$

$$g'\left(\frac{\pi}{3}\right) = \sqrt{3}\left(-\frac{\sqrt{3}}{2}\right) + 4\cdot\frac{1}{2} = -\frac{3}{2} + 2 = \frac{1}{2}$$

The Vector Form of the Basic Chain Rule

We can use the dot product to rewrite the Chain Rule into a vector form. In particular, if we have

$$T(t) = f\big(x(t), y(t)\big)$$

Where $f\big(x, y\big), x(t)$, and $y(t)$ are differentiable, then

$$\begin{aligned}\frac{dT}{dt} &= \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} \\ &= \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) \cdot \left(\frac{dx}{dt}, \frac{dy}{dt}\right) \\ &= \nabla f \cdot \frac{d\mathbf{x}}{dt}\end{aligned}$$

So, we have

$$\frac{d}{dt} f\big(\mathbf{x}(t)\big) = \nabla f\big(\mathbf{x}(t)\big) \cdot \frac{d\mathbf{x}}{dt}(t)$$

With $\mathbf{x}(t) = \big(x(t), y(t)\big)$.

In this vector form, the Chain Rule holds for any differentiable function $f(\mathbf{x}), \mathbf{x} \in \mathbb{R}^{\mathrm{n}}$.

6.2 Extensions of the Basic Chain Rule

2019年2月10日 20:05

$$\frac{\partial u}{\partial s} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial s}$$
$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial t}$$

(6.11)

Remarks

1. It is important to understand the difference between the various partial derivatives in equations (6.11), and to know which variable is held constant. For example

   $\frac{\partial u}{\partial x}$ means:

   Regard u as the given function of x and y, and differentiate with respect to x, holding y fixed.

   $\frac{\partial u}{\partial s}$ means:

   Regard u as the composite function of s and t, and differentiate with respect to s, holding t fixed.

2. Equations of the form $x = x(s,t), y = y(s,t)$ can be thought of as defining a change of coordinates in 2-space.

Remark

In some situation (see the example to follow) it is necessary to write a more precise form of the Chain Rule (6.11), one which displays the functional dependence.

Let g denote the composite function of $f(x,y)$ and $x(s,t), y(s,t)$:

$$g(s,t) = f\big(x(s,t), y(s,t)\big)$$

Then, the first equation in (6.11) can be written as

$$\frac{\partial g}{\partial s}(s,t) = \frac{\partial f}{\partial x}\big(x(s,t), y(s,t)\big)\frac{\partial x}{\partial s}(s,t) + \frac{\partial f}{\partial y}\big(x(s,t), y(s,t)\big)\frac{\partial y}{\partial s}(s,t)$$

With a similar equation for $\frac{\partial g}{\partial t}(s,t)$.

Algorithm

To write the Chain Rule from a dependence diagram we:

1. Take all possible paths from the differentiated variable to the differentiating variable.
2. For each link (-) in a given path, differentiate the upper variable with respect to the lower variable being careful to consider if this is a derivative or a partial derivative. Multiply all such derivatives in that path.
3. Add the products from step 2 together to complete the Chain Rule.

6.3 The Chain Rule for Second Partial Derivatives

2019年2月12日 18:12

Don't Quite Understand...

Chapter 7 Directional Derivatives and the Gradient Vector

2019年2月13日　　8:44

7.1 Directional Derivatives

2019年2月13日　8:45

Definition: **Directional Derivative**

The **directional derivative** of $f(x,y)$ at a point (a,b) in the direction of a unit vector $\hat{u} = (u_1, u_2)$ is defined by

$$D_{\hat{u}} f(a,b) = \frac{d}{ds} f(a + su_1, b + su_2)\Big|_{s=0}$$

Provided the derivative exists.

Theorem 1

If $f(x,y)$ is differentiable at (a,b) and $\hat{u} = (u_1, u_2)$ is a unit vector, then

$$D_{\hat{u}} f(a,b) = \nabla f(a,b) \cdot \hat{u}$$

Proof:

Since f is differentiable at (a,b) we can apply the Chain Rule to get

$$\begin{aligned} D_{\hat{u}} f(a,b) &= \frac{d}{ds} f(a + su_1, b + su_2)\Big|_{s=0} \\ &= \left[D_1 f(a + su_1, b + su_2) \frac{d}{ds}(b + su_1) + D_2 f(a + su_1, b + su_2) \frac{d}{ds}(b + su_2) \right]\Bigg|_{s=0} \\ &= \left[D_1 f(a + su_1, b + su_2) u_1 + D_2 f(a + su_1, b + su_2) u_2 \right]\Big|_{s=0} \\ &= D_1 f(a,b) u_1 + D_2 f(a,b) u_2 \\ &= \nabla f(a,b) \cdot (u_1, u_2) \end{aligned}$$

Remarks

1. Be careful to check the condition of Theorem 1 before applying it. If f is not differentiable at (a,b), then we must apply the definition of the directional derivative.
2. If we choose $\hat{u} = \hat{\imath} = (1,0)$ or $\hat{u} = \hat{\jmath} = (0,1)$, then the directional derivative is equal to the partial derivatives f_x or f_y respectively.

Exercise 1

Find the directional derivative of f defined by

$$f(x,y,z) = e^{xyz}$$

At the point $(1,-1,2)$ in the direction of the vector $\vec{u} = (1,2,-2)$.

Solution:

The vector is not a unit vector, so we first normalize it.

$$\vec{u} = \frac{(1,2,-2)}{||1,2,-2||} = \left(\frac{1}{3}, \frac{2}{3}, -\frac{2}{3}\right)$$

$$\frac{\partial f}{\partial x} = yz\, e^{xyz}$$

Similarly, $\frac{\partial y}{\partial x} = xze^{xyz}$

We have

$$\nabla f(x,y,z) = (yze^{xyz}, xze^{xyz}, xye^{xyz})$$

So $\nabla f(1,-1,2) = (-2e^{-2}, 2e^{-2}, -e^{-2})$

$$D_{\hat{u}} f(2,1) = (-2e^{-2}, 2e^{-2}, -e^{-2}) \cdot \left(\frac{1}{3}, \frac{2}{3}, -\frac{2}{3}\right) = -\frac{2}{3}e^{-2} + \frac{4}{3}e^{-2} + \frac{2}{3}e^{-2} = \frac{4}{3}e^{-2}$$

7.2 The Gradient Vector in Two Dimensions

2019年2月21日 8:00

The Greatest Rate of Change

Theorem 1
If $f(x,y)$ is differentiable at (a,b) and $\nabla f(a,b) \neq (0,0)$, then the largest value of $D_{\hat{u}} f(a,b)$ is $\left|\left|\nabla f(a,b)\right|\right|$, and occurs when $\hat{u}$ is in the direction of $\nabla f(a,b)$.

Proof:
Since f is differentiable at (a,b) and $||\hat{u}|| = 1$ we have

$$\begin{aligned} D_{\hat{u}} f(a,b) &= \nabla f(a,b) \cdot \hat{u} \\ &= \left|\left|\nabla f(a,b)\right|\right| \, ||\hat{u}|| \cos\theta \\ &= \left|\left|\nabla f(a,b)\right|\right| \cos\theta \end{aligned}$$

Where θ is the angle between $\hat{u}$ and $\nabla f(a,b)$. Thus, $D_{\hat{u}} f(a,b)$ assumes its largest value when $\cos\theta = 1$ i.e. $\theta = 0$. Consequently, the largest value of $D_{\hat{u}} f(a,b)$ is $\left|\left|\nabla f(a,b)\right|\right|$ and occurs when $\hat{u}$ is in the direction of $\nabla f(a,b)$.

Exercise 1
Find the largest rate of change of $f(x,y) = \ln(x + y^2)$ at the point $(0,1)$, and the direction in which it occurs.

Solution:

$$\frac{\partial f}{\partial x} = \frac{1}{x + y^2}$$

$$\frac{\partial f}{\partial y} = \frac{2y}{x + y^2}$$

$$\nabla f(x,y) = \left(\frac{1}{x + y^2}, \frac{2y}{x + y^2}\right)$$

$\nabla f(0,1) = (1,2)$
The direction is $(1,2)$

Exercise 2
Given a non-constant function $f(x,y)$ and a point (a,b) such that the directional derivative at (a,b) is independent of the direction. What can you say about the tangent plane of the surface $z = f(x,y)$ at the point (a,b)?

Solution:
According to the definition of tangent plane,

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

Since

$$D_{\hat{u}} f(a,b) = \nabla f(a,b) \cdot \hat{u}$$

$\nabla f(a,b) = (0,0)$

~~Plane does not exist?~~
Plane is horizontal.

Theorem 1 also applies in any direction. That is, if $f(\mathbf{x}), \mathbf{x} \in \mathbb{R}^n$, is differentiable at $\mathbf{a}$ and $\hat{u} \in \mathbb{R}^n$

is a unit vector, then the largest value of $D_{\hat{u}}f(\mathbf{a})$ is $\left\|\nabla f(\mathbf{a})\right\|$, and it occurs when $\hat{u}$ is in the direction of $\nabla f(\mathbf{a})$.

The Gradient and the Level Curves of *f*

Theorem 2
If $f(x,y) \in C^1$ in a neighborhood of (a,b) and $\nabla f(a,b) \neq (0,0)$, then $\nabla f(a,b)$ is orthogonal to the level curve $f(x,y) = k$ through (a,b).

Proof:
Since $\nabla f(a,b) \neq (0,0)$, by the Implicit Function Theorem (see Appendix 1), the level curve $f(x,y) = k$ can be described by parametric equations $x = x(t), y = y(t)$ for $t \in I$ where $x(t)$ and $y(t)$ differentiable. Hence, the level curve may be written as $f(x(t),y(t)) = k, t \in I$.
Suppose

$$a = x(t_0), \qquad b = y(t_0) \quad \text{for some } t_0 \in I$$

Since f is differentiable, we can take the derivative of this equation with respect to t using the Chain Rule to get

$$f_x(x(t),y(t))x'(t) + f_y(x(t),y(t))y'(t) = 0$$

On setting $t = t_0$ we get

$$\nabla f(a,b) \cdot \left(x'(t_0), y'(t_0)\right) = 0$$

Thus, $\nabla f(a,b)$ is orthogonal to $\left(x'(t_0), y'(t_0)\right)$ which is tangent to the level curve.

Exercise 3
Prove the level curves of the functions f and g defined by

$$f(x,y) = \frac{y}{x^2}, \qquad x \neq 0, \qquad g(x,y) = x^2 + 2y^2$$

Intersect orthogonally. Illustrate graphically.

The Gradient Vector Field

The gradient of f associates a vector with each point of the domain of f, and is referred to as a **vector field**. It is re[resented graphically by drawing $\nabla f(a,b)$ as a vector emanating from the corresponding point (a,b).

7.3 The Gradient Vector in Three Dimensions

2019年2月21日 15:43

Theorem 1
If $f(x, y, z) \in C^1$ in a neighborhood of (a, b, c) and $\nabla f(a, b, c) \neq (0,0,0)$, then $\nabla f(a, b, c)$ is orthogonal to the level surface $f(x, y, z) = k$ through (a, b, c).

The details are similar to proof of Theorem 7.2.2.

Exercise 1
Find the equation of the tangent plane to the ellipsoid $x^2 + 2y^2 + 3z^2 = 12$ at the point $(1,1,\sqrt{3})$.

Solution:

$\nabla f(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a}) = 0$

Exercise 2
Find the equation of the tangent plane to the surface

$$z = \frac{xy}{3x - 2y} \quad \text{at } (1,2,-2)$$

Hint: Rewrite the equation as $z(3x - 2y) - xy = 0$ and use the above approach.

Chapter 8 Taylor Polynomials and Taylor's Theorem

2019年2月21日 18:57

8.1 The Taylor Polynomial of Degree 2

2019年2月21日 18:57

Definition: 2nd degree Taylor polynomial

The **second degree Taylor polynomial** $P_{2,(a,b)}$ of $f(x,y)$ at (a,b) is given by

$$P_{2,(a,b)}(x,y) = f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b)$$
$$+\frac{1}{2}\left[f_{xx}(a,b)(x-a)^2 + 2f_{xy}(a,b)(x-a)(y-b) + f_{yy}(a,b)(y-b)^2\right]$$

In general, it approximates $f(x,y)$ for (x,y) sufficiently close to (a,b):

$$f(x,y) \approx P_{2,(a,b)}(x,y)$$

With better accuracy than the linear approximation.

Exercise 1

1. Find the Taylor polynomial $P_{2,(a,b)}(x,y)$ for
$$f(x,y) = \frac{1}{2}y^2 + x - \frac{1}{3}x^3$$
At the point $(a,b) = (1,0)$, by calculating the appropriate partial derivatives.
2. Verify your results by letting $u = x-1, v = y$ and writing
$$f(x,y) = \frac{1}{2}v^2 + u + 1 - \frac{1}{3}(u+1)^3$$
Expand and neglect powers higher than 2 and then convert back to x and y. This type of algebraic derivation can only be done for a polynomial function.

Solution:

1.

$\nabla f(x,y) = (1-x^2, y)$

$\nabla f(1,0) = (0,0)$

$Hf(x,y) = \begin{bmatrix} -2x & 0 \\ 0 & 1 \end{bmatrix}$

$Hf(1,0) = \begin{bmatrix} -2 & 0 \\ 0 & 1 \end{bmatrix}$

Thus,

$$P_{2,(1,0)}(x,y) = \frac{2}{3} + \frac{1}{2}\left[-2(x-1)^2 + y^2\right]$$

2. Not sure what the question mean

8.2 Taylor's Formula with Second Degree Remainder

2019年2月24日 12:58

Review of the 1-D case
Theorem 1
If $f''(x)$ exists on $[a,x]$, then there exists a number c between a and x such that

$$f(x)=f(a)+f'(a)(x-a)+R_{1,a}(x)$$

Where

$$R_{1,a}(x)=\frac{1}{2}f''(c)(x-a)^2$$

The 2-D Case
Theorem 2 (Taylor's Theorem)
If $f(x,y)\in C^2$ in some neighborhood $N(a,b)$ of (a,b), then for all $(x,y)\in N(a,b)$ there exist a point (c,d) on the line segment joining (a,b) and (x,y) such that

$$f(x,y)=f(a,b)+f_x(a,b)(x-a)+f_y(a,b)(y-b)+R_{1,(a,b)}(x,y)$$

Where

$$R_{1,(a,b)}(x,y)=\frac{1}{2}\left[f_{xx}(c,d)(x-a)^2+2f_{xy}(c,d)(x-a)(y-b)+f_{yy}(c,d)(y-b)^2\right]$$

Proof:
The idea is to reduce the given function f of two variables to a function g of one variable, by considering only points on the line segment joining (a,b) and (x,y).
We parameterize the line segment L from (a,b) to (x,y) by

$$L(t)=\big(a+t(x-a),b+t(y-b)\big),\qquad 0\le t\le 1$$

For simplicity write $h=x-a$ and $k=y-b$. Then $x-a=h, y-b=k$, and Taylor's formula assumes the form

$$f(x,y)=f(a,b)+f_x(a,b)h+f_y(a,b)k+R_{1,(a,b)}(x,y)$$

Where

$$R_{1,(a,b)}(x,y)=\frac{1}{2}\left[f_{xx}(c,d)h^2+2f_{xy}(c,d)hk+f_{yy}(c,d)k^2\right]$$

Define g by

$$g(t)=f\big(L(t)\big),\qquad 0\le t\le 1$$

Since f has continuous second partials by hypothesis, we can apply the Chain Rule to conclude that g' and g'' are continuous and are given by

$$g'(t)=f_x\big(L(t)\big)h+f_y\big(L(t)\big)k$$

$$g''(t)=f_{xx}\big(L(t)\big)h^2+2f_{xy}\big(L(t)\big)hk+f_{yy}\big(L(t)\big)k^2$$

For $0\le t\le 1$.
Since g'' is continuous on the interval $[0,1]$, Taylor's formula may be applied to g on this interval. That is, we can set $x=1$ and $a=0$ in equations (8.2) and (8.3). It follows that there exists a number $\tilde{c}$, with $0<\tilde{c}<1$, such that

$$g(1)=g(0)+g'(0)+\frac{1}{2}g''(\tilde{c})$$

Each term in this equation can be calculated using equations (8.5)-(8.7), giving

$$g(1)=f\Big((a,b)+\big[(x,y)-(a,b)\big]\Big)=f(x,y)$$

$g(0)=f(a,b)$, and

$$g'(0)=f_x(a,b)h+f_y(a,b)k$$

In addition, if we let $(c,d)=L(\tilde{c})$, then

$$\frac{1}{2}g''(c)=R_{1,(a,b)}(x,y)$$

And equation (8.8) becomes precisely the modified version of Taylor's formula.

Remark

Like the one variable case, Taylor's Theorem for $f(x,y)$ is an existence theorem. That is, it only tells us that the point (c,d) exists, but not how to find it.

Exercise 1
Let $f(x,y)=e^{-2x+y}$. Use Taylor's Theorem to show that the error in the linear approximation $L_{(1,1)}(x,y)$ is at most $6e\left[(x-1)^2+(y-1)^2\right]$ if $0\le x\le 1$ and $0\le y\le 1$.

Solution:
From Taylor's Theorem, we get

$$R_{1,(1,1)}(x,y)=\frac{1}{2}\left[f_{xx}(c,d)(x-1)^2+2f_{xy}(c,d)(x-1)(y-1)+f_{yy}(c,d)(y-1)^2\right]$$

$f_x=(-2)e^{-2x+y}$
$f_y=e^{-2x+y}$

$f_{xx}=4e^{-2x+y}$
$f_{yy}=e^{-2x+y}$
$f_{xy}=-2e^{-2x+y}$

As $0\le x\le 1, 0\le y\le 1$.
$-2\le -2x\le 0, 0\le y\le 1$
$-1\le -2x+y\le 1$

$$\begin{aligned}R_{1,(1,1)}(x,y)&=\frac{1}{2}\left[4e^{-2x+y}(x-1)^2-4e^{-2x+y}(x-1)(y-1)+e^{-2x+y}(y-1)^2\right]\\&<\frac{1}{2}\left[4e(x-1)^2-4e(x-1)(y-1)+e(y-1)^2\right]\\&=\frac{1}{2}e\left[4(x-1)^2-4(x-1)(y-1)+(y-1)^2\right]\\&=\frac{1}{2}e\left[\left(2(x-1)-(y-1)\right)^2\right]\\&=\end{aligned}$$

Remark
The most important thing about the error term $R_{1,(a,b)}(x,y)$ is not its explicit form, but rather its dependence on the magnitude of the displacement $\left\|(x,y)-(a,b)\right\|$. We state the result as a Corollary.

Corollary 3
If $f(x,y)\in C^2$ in some closed neighborhood $N(a,b)$ of (a,b), then there exists a positive constant M such that

$$\left|R_{1,(a,b)}(x,y)\right|\le M\left\|(x,y)-(a,b)\right\|^2,\qquad \text{for all } (x,y)\in N(a,b)$$

8.3 Generalizations

2019年3月5日 15:58

Exercise 1
Write out $P_{3,(a,b)}(x,y)$ explicitly using subscript notation.

Theorem 1
Taylor's Theorem of order k
If $f(x,y) \in C^{k+1}$ at each point on the line segment joining (a,b) and (x,y), then there exists a point (c,d) on the line segment between (a,b) and (x,y) such that

$$f(x,y) = P_{k,(a,b)}(x,y) + R_{k,(a,b)}(x,y)$$

Where

$$R_{k,(a,b)}(x,y) = \frac{1}{(k+1)!}\left[(x-a)D_1 + (y-b)D_2\right]^{k+1} f(c,d)$$

Corollary 2
If $f(x,y) \in C^k$ in some neighborhood of (a,b), then

$$\lim_{(x,y)\to(a,b)} \frac{\left|f(x,y) - P_{k,(a,b)}(x,y)\right|}{\left\|(x,y)-(a,b)\right\|^k} = 0$$

Corollary 3
If $f(x,y) \in C^{k+1}$ in some closed neighborhood $N(a,b)$ of (a,b), then there exists a constant $M > 0$ such that

$$\left|f(x,y) - P_{k,(a,b)}(x,y)\right| \le M\left\|(x,y)-(a,b)\right\|^{k+1}$$

For all $(x,y) \in N(a,b)$.

The final stage in the process of generalization is to consider functions of n variables $f(\mathbf{x}), \mathbf{x} \in \mathbb{R}^n$. One has simply to replace the differential operator

$$\left[(x-a)D_1 + (y-b)D_2\right]$$

By

$$\left[(x_1-a_1)D_1 + \cdots + (x_n-a_n)D_n\right]$$

Letting $\nabla = (D_1, \ldots, D_n)$, we can be write this concisely in vector notation as

$$[(\mathbf{x}-\mathbf{a})\cdot\nabla]$$

Chapter 9 Critical Points

2019年3月5日 20:13

9.1 Local Extrema and Critical Points

2019年3月5日 20:14

Definition
Local Maximum and Minimum

A point (a, b) is a **local maximum point of** f if $f(x, y) \leq f(a, b)$ for all (x, y) in some neighborhood of (a, b).
A point (a, b) is a **local minimum point of** f if $f(x, y) \geq f(a, b)$ for all (x, y) in some neighborhood of (a, b).

Theorem 1
If (a, b) is a local maximum or minimum point of f, then

$$f_x(a, b) = 0 = f_y(a, b)$$

Or at least one of f_x or f_y does not exist at (a, b).

Proof:
Consider the function g defined by $g(x) = f(x, b)$. If (a, b) is a local maximum/minimum point of f, then $x = a$ is a local maximum/minimum point of g, and hence either $g'(a) = 0$ or $g'(a)$ does not exist. Thus it follows that either $f_x(a, b) = 0$ or $f_x(a, b)$ does not exist. A similar argument gives $f_y(a, b) = 0$ or $f_y(a, b)$ does not exist.

Definition
Critical Point

A point (a, b) in the domain of $f(x, y)$ is called a **critical point** of f if

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

Or if at least one of the partial derivatives does not exist at (a, b).

Definition
Saddle Point

A critical point (a, b) of $f(x, y)$ is called a **saddle point** of f if in every neighborhood of (a, b) there exist points (x_1, y_1) and (x_2, y_2) such that

$$f(x_1, y_1) > f(a, b) \text{ and } f(x_2, y_2) < f(a, b)$$

The problem that we are faced with has two parts.
1. Given $f(x, y)$, find all critical points of f.
2. Determine whether the critical points are local maxima, minima or saddle points.

Remark
1. It is essential to solve equations (9.1) and (9.2) systematically, by considering all possible cases, in order to find all critical points.
2. You should be aware that we can only explicitly find the critical points for simple functions f. The equations
   $$f_x(x, y) = 0, \qquad f_y(x, y) = 0$$
   Are a system of equations which are generally non-linear, and there is no general algorithms for solving such systems exactly. There are, however, numerical methods for finding approximate solutions, one of which is a generalization of Newton's method. If you review the one variable case, you might see how to generalize it, using the tangent plane.

It's a challenge!

Exercise 1
Find all critical points of $f(x,y) = xye^{x-y}$.

Solution:
$f_x = y(e^{x-y} + xe^{x-y})$
$f_y = x(e^{x-y} - ye^{x-y})$

$f_x = f_y = 0$

We get two equations
$y(e^{x-y} + xe^{x-y}) = 0$
$x(e^{x-y} - ye^{x-y}) = 0$

Case 1:
$y = 0 \Rightarrow x(e^x) = 0 \Rightarrow x = 0$

Critical point (0,0)

Case 2:
$(e^{x-y} + xe^{x-y}) = 0$
$(1+x)(e^{x-y}) = 0$
$(1+x) = 0$
$x = -1$

$-(e^{-1-y} - ye^{-1-y}) = 0$
$(1-y)e^{-1-y} = 0$
$1 = y$

Critical point$(-1,1)$

Exercise 2
Find all critical points of $f(x,y) = x\cos(x+y)$.
Case 2??? How to deal with this situation?

Solution:
$f_x = (\cos(x+y) - x\sin(x+y))$
$f_y = x(-\sin(x+y))$

We get two equations
$(\cos(x+y) - x\sin(x+y)) = 0$
$(x(-\sin(x+y)) = 0$

Case 1:
$x = 0 \Rightarrow (\cos y) = 0 \Rightarrow y = \frac{\pi}{2} + k\pi\ (k \in \mathbb{Z})$

Case 2:
$\sin(x+y) = 0 \Rightarrow (x+y) = k\pi\ (k \in \mathbb{Z})$

$f_x = (\cos(k\pi)) = 0$

Exercise 3
Give a function $f(x,y)$ with no critical points.

$f(x, y) = x$

9.2 The Second Derivative Test

2019年3月7日 20:57

Review of the 1-D case
For a function $f(x)$ of one variable, the second degree Taylor polynomial approxiamtion is

$$f(x) \approx f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2$$

For x sufficiently close to a. If $x = a$ is a critical point of f, then $f'(a) = 0$, and the approximation can be rearranged to give

$$f(x) - f(a) \approx \frac{1}{2}f''(a)(x-a)^2$$

Thus, for x sufficiently close to a, $f(x) - f(a)$ has the same sign as $f''(a)$. If $f''(a) > 0$, then $f(x) - f(a) > 0$ for x sufficiently close to a and $x = a$ is a minimum point. If $f''(a) < 0$, then $f(x) - f(a) < 0$ for x sufficiently close to a and $x = a$ is a local maximum point. There is no conclusion if $f''(a) = 0$.

The 2-D case
For $f(x, y) \in C^2$, the second degree Taylor polynomial approximation is

$$f(x, y) \approx P_{2,(a,b)}(x, y)$$

For (x, y) sufficiently close to (a, b). If (a, b) is a critical point of f such that

$$f_x(a, b) = 0 = f_y(a, b)$$

Then the approximation can be rearranged to yield

$$f(x, y) - f(a, b) \approx \frac{1}{2}\left[f_{xx}(a, b)(x-a)^2 + 2f_{xy}(a, b)(x-a)(y-b) + f_{yy}(a, b)(y-b)^2\right]$$

For (x, y) sufficiently close to (a, b). The sign of the expression on the right will determine the sign of $f(x, y) - f(a, b)$, and hence whether (a, b) is a local maximum, local minimum or saddle point.

The expression on the right is called a **quadratic form**, and at this stage it is necessary to discuss some properties of these objects.

Quadratic Forms
Definition
Quadratic Form

A function Q of the form

$$Q(u, v) = a_{11}u^2 + 2a_{12}uv + a_{22}v^2$$

Where a_{11}, a_{12} and a_{22} are constants, is called a **quadratic form** on $\mathbb{R}^2$.

Remark
Semidefinite quadratic forms may be split into two classes, positive semidefinite and negative semidefinite. The matrix C above would be classified as positive semidefinite.

If A is not a diagonal matrix, the nature of A (or of $Q(u, v)$) is not immediately obvious. For example, even if all entries of A are positive, it does not follow that A is a positive definite matrix.

Theorem 1
Second Partial Derivatives Test
Suppose that $f(x, y) \in C^2$ in some neighborhood of (a, b) and that

$$f_x(a, b) = 0 = f_y(a, b)$$

1. If $Hf(a, b)$ is positive definite, then (a, b) is a local minimum point of f.
2. If $Hf(a, b)$ is negative definite, then (a, b) is a local maximum point of f.
3. If $Hf(a, b)$ is indefinite, then (a, b) is a saddle point of f.

Remarks

1. The argument preceding the theorem is not a proof, since it involves an approximation. One can use Taylor's formula and a continuity argument to give a proof. See Section 9.3.
2. Note the analogy with the second derivative test for functions of one variable. The requirement $g''(a) > 0$, which implies a local minimum, is replaced by the requirement that the matrix of second partial derivatives $Hf(a,b)$ be positive definite.

Theorem 2

If $Q(u,v) = a_{11}u^2 + 2a_{12}uv + a_{22}v^2$ and $D = a_{11}a_{22} - a_{12}^2$, then

1. Q is positive definite if and only if $D > 0$ and $a_{11} > 0$
2. Q is negative definite if and only if $D > 0$ and $a_{11} < 0$
3. Q is indefinite if and only if $D < 0$
4. Q is semidefinite if and only if $D = 0$

Remark

Observe that D is the determinant of the associated symmetric matrix.

Example 3

Omitted

Exercise 1

Fill in the details of Example 3 above.

Exercise 2

Find and classify all critical points of the function $f(x,y) = x^2 + 6xy + 2y^2$.

Exercise 3

Find and classify all critical points of the function $f(x,y) = (x^2 + y^2 - 1)y$.

Remark

Another way of classifying the Hessian matrix is by finding its eigenvalues. In particular, a symmetric matrix is positive definite if all of its eigenvalues are positive, negative definite if all of its eigenvalues are negative, and indefinite if has both positive and negative eigenvalues.

Degenerate Critical Points

If $Hf(a,b)$ is semidefinite, so that the second partial derivative test gives no conclusion, we say that the crucial point (a,b) is **degenerate**. In order to classify the critical point, one has to investigate the sign of $f(x,y) - f(a,b)$ in a small neighborhood of (a,b).

Generalizations

The definitions of local maximum point, local minimum point and critical point can be generalized in the obvious way to functions f of n variables. The Hessian matrix of f at $\mathbf{a}$ is the $n \times n$ symmetric matrix given by

$$Hf(\mathbf{a}) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a})\right]$$

Where $i,j = 1,2,\ldots,n$. The Hessian matrix can be classified as positive definitem negative definite, indefinite or semidefinite by considering the associated quadratic form in $\mathbb{R}^n$:

$$Q(\mathbf{u}) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a,b)u_i u_j$$

As in $\mathbb{R}^2$. The second derivative test as stated in $\mathbb{R}^2$ now holds in $\mathbb{R}^n$. It can be justified heuristically by using the second degree Taylor polynomial approximation,

$$f(\mathbf{x}) \approx P_{2,\mathbf{a}}(\mathbf{x})$$

Which leads to

$$f(\mathbf{x}) - f(\mathbf{a}) \approx \frac{1}{2!}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a,b)(x_i - a_i)(x_j - a_j)$$

Generalizing equation (9.3).

Level Curves Near a Critical Point

A point at which $\nabla f(a,b) \neq (0,0)$ is called a **regular point** of f.

Convex Functions
1-D Case

We say that a twice differentiable function $f(x)$ is **strictly convex** if $f''(x) > 0$ for all x and f is **convex** is $f''(x) \geq 0$ for all x. Thus the term convex means "concave up". Convex functions have two interesting properties.

1. $f(x) > L_a(x) = f(a) + f'(x)(x-a)$ for all $x \neq a$, for any $a \in \mathbb{R}$.
2. For $a < b$, $f(x) < f(a) + \frac{f(b)-f(a)}{x-a}(x-a)$ for $x \in (a,b)$.

Proof:

1. Follows from Taylor's Theorem: $f(x) = L_a(x) + \frac{f''(c)}{2}(x-a)^2$ where c is between a and x. Thus $R_{1,a}(x) > 0$ for $x \neq a$, giving $f(x) > L_a(x)$ for all $x \neq a$.
2. Let $g(x) = f(x) - \left[f(a) + \frac{f(b)-f(a)}{b-a}(x-a)\right]$. Then $g(a) = g(b) = 0$ and $g''(x) = f''(x) > 0$. We must show that $g(x) < 0$ for $x \in (a,b)$. By the Mean Value Theorem $\frac{f(b)-f(a)}{b-a} = f'(c)$ for some $c \in (a,b)$. Note that $g'(x) = f'(x) - \frac{f(b)-f(a)}{b-a} = f'(x) - f'(c)$. Thus $g'(c) = 0$. Since $g''(x) > 0$ then $g'(x)$ is strictly increasing. Since $g'(c) = 0$ then $g'(x) < 0$ on $[a,c)$ and $g'(x) > 0$ on $(c,b]$. This implies that $g(x)$ is strictly decreasing on $[a,c]$ and strictly increasing on $[c,b]$. Since $g(a) = 0$ and $g(b) = 0$ we get that $g(x) < 0$ on $(a,c]$ and on $[c,b)$. Therefore, $g(x) < 0$ on (a,b), as required.

Remark
(1) says that the graph of f lies above any tangent line, and (2) says that any secant line lies above the graph of f.

2-D Case
Suppose $f(x,y)$ has continuous second partial derivatives. We say that f is **strictly convex** if $Hf(x,y)$ is positive definite for all (x,y). By Theorem 2, f is strictly convex means $f_{xx} > 0$ and $f_{xx}f_{yy} - f_{xy}^2 > 0$ for all (x,y). We get a result which is analogous to Theorem 3.

Theorem 4
If $f(x,y)$ has continuous second partial derivatives and is strictly convex, then

1. $f(x,y) > L_{(a,b)}(x,y)$ for all $(x,y) \neq (a,b)$, and
2. $f\left(a_1 + t(b_1 - a_1), a_2 + t(b_2 - a_2)\right) < f(a_1,a_2) + t[f(b_1,b_2) - f(a_1,a_2)]$ for $0 < t < 1$, $(a_1,a_2) \neq (b_1,b_2)$.

Proof:

1. Follows from Taylor's Theorem:
$$f = L_{(a,b)}(x,y) + \frac{1}{2}\left[f_{xx}(c,d)(x-a)^2 + 2f_{xy}(c,d)(x-a)(y-b) + f_{yy}(c,d)(y-b)^2\right]$$
Where (c,d) is on the line segment from (a,b) to (x,y). Since $f_{xx}(c,d) > 0$, $f_{xx}(c,d)f_{yy}(c,d) - f_{xy}(c,d)^2 > 0$, $R_{1,(a,b)}(x,y) > 0$ for $(x,y) \neq (a,b)$ by Theorem 2. Therefore, $f(x,y) > L_{(a,b)}(x,y)$ for $(x,y) \neq (a,b)$.
2. We parameterize the line segment L from (a_1,a_2) to (b_1,b_2) by

$$L(t) = \left(a_1 + t(b_1 - a_1), a_2 + t(b_2 - a_2)\right), \qquad 0 \le t \le 1$$

For simplicity write $h = b_1 - a_1$ and $k = b_2 - a_2$. Define $g(t)$ by

$$g(t) = f(L(t)), \qquad 0 \le t \le 1$$

Since f has continuous second partials by hypothesis, we can apply the Chain Rule to conclude that g' and g'' are continuous and are given by

$$g'(t) = f_x(L(t))h + f_y(L(t))k$$

$$g''(t) = f_{xx}(L(t))h^2 + 2f_{xy}(L(t))hk + f_{yy}(L(t))k^2$$

For $0 \le t \le 1$. Since $f_{xx}(L(t)) > 0$ and $f_{xx}(L(t))f_{yy}(L(t)) - f_{xy}(L(t))^2 > 0$ for all t, $g''(t) > 0$ by Theorem 2. Thus, by Theorem 3, part (2):

$$g(t) < g(0) + \frac{g(1) - g(0)}{1 - 0}(t - 0), \qquad \text{for } 0 < t < 1$$

Therefore, $f\left(a_1 + t(b_1 - a_1), a_2 + t(b_2 - a_2)\right) < f(a_1, a_2) + t[f(b_1, b_2) - f(a_1, a_2)]$ for $0 < t < 1$ as required.

Remark

(1) says that the graph of f lies above the tangent plane and (2) says that the cross-section of the graph of f above the line segment from (a_1, a_2) to (b_1, b_2) lies below the secant line.

Theorem 5

If $f(x, y) \in C^2$ is strictly convex and has a critical point (c, d), then $f(x, y) > f(c, d)$ for all $(x, y) \neq (c, d)$ and f has no other critical point.

Proof:

Note that $L_{(c,d)}(x, y) = f(c, d)$. Thus, $f(x, y) > f(c, d)$ for all $(x, y) \neq (c, d)$ by Theorem 4, part (1). If f has a second critical point (c_1, d_1), then by the reasoning just given $f(c_1, d_1) > f(c, d)$ and $f(c, d) > f(c_1, d_1)$ which is a contradiction.

9.3 Proof of the Second Partial Derivative Test

2019年3月11日 8:35

Lemma 1

Let $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$ be a positive definite matrix. If $|\tilde{a}-a|, |\tilde{b}-b|$ and $|\tilde{c}-c|$ are sufficiently small, then $\begin{bmatrix} \tilde{a} & \tilde{b} \\ \tilde{b} & \tilde{c} \end{bmatrix}$ is positive definite.

Proof:

Let Q and $\tilde{Q}$ be the quadratic forms determined by the given matrices.

$$Q(u,v) = au^2 + 2buv + cv^2$$

And similarly for $\tilde{Q}(u,v)$. We perform the change of variables

$$u = r\cos\theta, \qquad v = r\sin\theta$$

To obtain

$$Q(u,v) = r^2 p(\theta)$$

Where

$$p(\theta) = a\cos^2\theta + 2b\cos\theta\sin\theta + c\sin^2\theta$$

Since for $r = 1, Q(u,v) = p(\theta)$, and Q is positive definite, we must have $p(\theta) > 0$ for all θ, $0 \le \theta \le 2\pi$.

Let

$$k = \min_{0\le\theta\le 2\pi} p(\theta)$$

Then $k > 0$ and bu equation (9.11)

$$Q(u,v) \ge kr^2 \text{ for all } (u,v) \neq (0,0)$$

We are given that $|\tilde{a}-a|, |\tilde{b}-b|$ and $|\tilde{c}-c|$ are sufficiently small. Let

$$\delta = \max\{|\tilde{a}-a|, |\tilde{b}-b|, |\tilde{c}-c|\}$$

By equation (0.10) and the triangle inequality,

$$\begin{aligned} |Q(u,v) - \tilde{Q}(u,v)| &\le |\tilde{a}-a|u^2 + 2|\tilde{b}-b||u||v| + |\tilde{c}-c|v^2 \\ &\le \delta(u^2 + 2|u||v| + v^2) \\ &= \delta(|u|+|v|)^2 \\ &= \delta r^2(|\cos\theta| + |\sin\theta|)^2 \\ &< 4\delta r^2 \end{aligned}$$

We now choose $\delta = \frac{1}{8}k$. Then

$$|Q(u,v) - \tilde{Q}(u,v)| < \frac{1}{2}kr^2$$

Which implies

$$\begin{aligned} \tilde{Q}(u,v) &\ge Q(u,v) - \frac{1}{2}kr^2 \\ &\ge kr^2 - \frac{1}{2}kr^2, \qquad \text{by (9.12)} \\ &= \frac{1}{2}kr^2 \end{aligned}$$

This shows that $\tilde{Q}(u,v) > 0$ for all $(u,v) \neq (0,0)$. Therefore, $\tilde{Q}(u,v)$ is positive definite.

Remark

The lemma is also true if "positive definite" is replaced by "negative definite" or "indefinite".

Theorem 2

(The Second Partial Derivative Test)

Suppose that $f(x,y) \in C^2$ in some neighborhood of (a,b) and that

$$f_x(a,b) = 0 = f_y(a,b)$$

1. If $Hf(a,b)$ is positive definite, then (a,b) is a local minimum point of f.

2. If $Hf(a,b)$ is negative definite, then (a,b) is a local maximum of f.
3. If $Hf(a,b)$ is indefinite, then (a,b) is a saddle point of f.

Proof:
We will prove (1).
We apply Taylor's formula with second order remainder. Since

$$f_x(a,b) = 0 = f_y(a,b)$$

Taylor's formula can be written as

$$f(x,y) - f(a,b) = \frac{1}{2}\left[f_{xx}(c,d)(x-a)^2 + 2f_{xy}(c,d)(x-a)(y-b) + f_{yy}(c,d)(y-b)^2\right]$$

Where (c,d) lies on the line segment joining (a,b) and (x,y). The coefficient matrix in the quadratic expression on the right side of (9.13) is hte Hessian matrix $Hf(c,d)$.

We are given that $Hf(a,b)$ is positive definite. By the lemma, there exists $\epsilon > 0$ such that if

$$|f_{xx}(x,y) - f_{xx}(a,b)| < \epsilon, \qquad |f_{xy}(x,y) - f_{xy}(a,b)| < \epsilon, \qquad |f_{yy}(x,y) - f_{yy}(a,b)| < \epsilon$$

Then $Hf(x,y)$ is positive definite. Since the second partials of f are continuous at (a,b), the definition of continuity implies that there exists a $\delta > 0$ such that

$$\|(x,y) - (a,b)\| < \delta$$

Implies (9.14) and hence that $Hf(x,y)$ is positive definite. Since

$$\|(c,d) - (a,b)\| < \|(x,y) - (a,b)\|$$

It follows that $Hf(c,d)$ is also positive definite. It now follows from equation (9.13) and the definition of positive definite matrix, that if $0 < \|(x,y) - (a,b)\| < \delta$, then $f(x,y) - f(a,b) > 0$. Thus, by definition (a,b) is a local minimum point of f.

Parts (2) and (3) of the second derivative test can be proved in a similar way using the modified lemma.

Chapter 10 Optimization Problems

2019年3月13日 8:18

10.1 The Extreme Value Theorem

2019年3月13日 8:20

Definition
Absolute Maximum and Minimum

Given a function $f(x,y)$ and a set $S \subseteq \mathbb{R}^2$,

1. A point $(a,b) \in S$ is an absolute maximum point of f on S if
   $$f(x,y) \le f(a,b) \quad \text{for all } (x,y) \in S$$
   The value $f(a,b)$ is called the **absolute maximum value** of f on S.
2. A point $(a,b) \in S$ is an **absolute minimum point** of f on S if
   $$f(x,y) \ge f(a,b) \quad \text{for all } (x,y) \in S$$
   The value $f(a,b)$ is called the **absolute minimum value** of f on S.

The Extreme Value Theorem

Theorem 1
(The Extreme Value Theorem)

If $f(x)$ is continuous on a finite closed interval I, then there exists $c_1, c_2 \in I$ such that
$$f(c_1) \le f(x) \le f(c_2) \quad \text{for all } x \in I$$

Exercise 1
Given a function $f(x)$ and an interval I such that

1. I is closed, but f does not have an absolute maximum of I.
2. I is finite and f is continuous on I, but f does not have an absolute maximum on I.
3. I is finite and f is continuous on I, but f does not have an absolute minimum.

Definition
Bounded Set

A set $S \subset \mathbb{R}^2$ is said to be **bounded** if and only if it is contained in some neighbourhood of the origin.

Definition
Boundary Point

Given a set $S \subset \mathbb{R}^2$, a point $(a,b) \in \mathbb{R}^2$ is said to be a **boundary point** of S if and only if every neighbourhood of (a,b) contains at least one point in S and one point not in S.

Definition
Boundary of S

The set $B(S)$ of all boundary points of S is called the **boundary** of S.

Definition
Closed Set

A set $S \subseteq \mathbb{R}^2$ is said to be **closed** if S contains all of its boundary points.

Theorem 2
If $f(x,y)$ is continuous on a closed and bounded set $S \subset \mathbb{R}^2$, then there exists points $(a,b), (c,d) \in S$ such that
$$f(a,b) \le f(x,y) \le f(c,d) \quad \text{for all } (x,y) \in S$$

Remark
A function $f(x,y)$ may have an absolute maximum and/or an absolute minimum on a set $S \subseteq \mathbb{R}^2$ even if the conditions are not satisfied. We just cannot guarantee the existence with the theorem.

10.2 Algorithm for Extreme Values

2019年3月13日 8:45

Algorithm

Let $S \subset \mathbb{R}^2$ be closed and bounded. To find the maximum and/or minimum value of a function $f(x,y)$ that is continuous on S,

1. Find all critical points of f that are contained in S. Evaluate f at each such point.
2. Find the maximum and minimum points of f on the boundary $B(S)$.
3. The maximum value of f on S is the largest value of the function found in steps (1) and (2). The minimum value of f on S is the smallest value of the function found in steps (1) and (2).

Remarks

1. The absolute maximum and/or minimum value may occur at more than one point in S.
2. It is not necessary to determine whether the critical points are local maximum or minimum points.

Exercise 1

Find the maximum of $f(x,y) = x^2y - y$ on the set $S = \{(x,y) | 9x^2 + 4y^2 \leq 36\}$.

Solution:

Skip for now....

Exercise 2

Find the maximum value of the function $f(x,y) = x^2y + xy^2$ on the triangular region with vertices (0,0), (0,1) and (1,0).

10.3 Optimization with Constraints

2019年3月13日 9:05

Method of Lagrange Multipliers
We want to find the maximum (minimum) value of a differentiable function $f(x,y)$ subject to the constraint $g(x,y)=k$ where $g\in C^1$, or, more geometrically, find the maximum (minimum) value of $f(x,y)$ on the level set $g(x,y)=k$.

If $f(x,y)$ has a local maximum (or minimum) at (a,b) relative to nearby points on the curve $g(x,y)=k$ and $\nabla g(a,b)\neq(0,0)$, then, by the Implicit Function Theorem (see Appendix 1), $g(x,y)=k$ can be described by parametric equations

$$x=p(t),\qquad y=q(t)$$

With p and q differentiable, and $(a,b)=\left(p(t_0),q(t_0)\right)$ for some t_0. Define

$$u(t)=f\left(p(t),q(t)\right)$$

The function u gives the values of f on the constraint curve, and hence has a local maximum (or minimum) at t_0. It follows that

$$u'(t_0)=0$$

Assuming f is differentiable we can apply the Chain Rule to get

$$u'(t)=f_x\left(p(t),q(t)\right)p'(t)+f_y\left(p(t),q(t)\right)q'(t)$$

Evaluating this at t_0 and using (10.2) gives

$$0=f_x(a,b)p'(t_0)+f_y(a,b)q'(t_0)$$

This can be written as

$$\nabla f(a,b)\cdot\left(p'(t_0),q'(t_0)\right)=0$$

Recall the geometric interpretation of the gradient vector $\nabla g(a,b)$ proven in Theorem 7.2.2 that $\nabla g(a,b)$, if non-zero, is orthogonal to the level curve $g(x,y)=k$ at (a,b). Thus, since $\left(p'(t_0),q'(t_0)\right)$ is the tangent vector to the constraint curve (10.1) we also have

$$\nabla g(a,b)\cdot\left(p'(t_0),q'(t_0)\right)=0$$

Since we are working in two dimensions, equations (10.3) and (10.4) imply that $\nabla f(a,b)$ and $\nabla g(a,b)$ are scalar multiples of each other. That is, there exists a constant λ such that

$$\nabla f(a,b)=\lambda\nabla g(a,b)$$

This leads to the following procedure, called the **Method of Lagrange Multipliers.**

Algorithm (Lagrange Multiplier Algorithm)
Assume that $f(x,y)$ is a differentiable function and $g\in C^1$. To find the maximum value and minimum value of f subject to the constraint $g(x,y)=k$, evaluate $f(x,y)$ at all points (a,b) which satisfy one of the following conditions.

1. $\nabla f(a,b)=\lambda\nabla g(a,b)$ and $g(a,b)=k$
2. $\nabla g(a,b)=(0,0)$ and $g(a,b)=k$
3. (a,b) is an end point of the curve $g(x,y)=k$

The maximum/minimum value of $f(x,y)$ is the largest/smallest value of f obtained at the points found in $(1)-(3)$.

Remark

1. The variable λ, called the **Lagrange multiplier**, is not required for our purposes and so should be eliminated. However, in some real world applications, the value of λ can be extremely useful.
2. Case (2) and (3) are both exceptional. Observe that case (2) must be included since we assume that $\nabla g(a,b)\neq(0,0)$ in the derivation.
3. In the curve $g(x,y)=k$ is unbounded, then one must consider $\lim_{\|(x,y)\|\to\infty}f(x,y)$ for

(x, y) satisfying $g(x, y) = k$.

Exercise 1
Exercise 2
Omitted

Functions of Three Variables
Algorithm
To find the maximum/minimum value of a differentiable function $f(x, y, z)$ subject to $g(x, y, z) = k$ such that $g \in C^1$, we evaluate $f(x, y, z)$ at all points (a, b, c) which satisfy one of the following:

1. $\nabla f(a, b, c) = \lambda \nabla g(a, b, c)$ and $g(a, b, c) = k$
2. $\nabla g(a, b, c) = (0,0,0)$ and $g(a, b, c) = k$
3. (a, b, c) is a boundary point of the surface $g(x, y, z) = k$

The maximum/minimum value of $f(x, y, z)$ is the largest/smallest value of f obtained from all points found in (1) - (3).

Remark
If condition (1) in the algorithm holds, it follows that the level surface $f(x, y, z) = f(a, b, c)$ and the constraint surface $g(x, y, z) = k$ are tangent at the point (a, b, c), since their normals coincide (See Theroem 7.3.1)

Remark
Keep in mind the geometric interpretation. The level sets $f(x, y, z) = k$ are spheres centered on the point (1,2,2). The minimum distance occurs when one of the spheres touches (i.e. Is tangent to) the constraint surface which is the sphere $g(x, y, z) = 1$. At the point of tangency the normals are parallel, i.e. $\nabla f = \lambda \nabla g$.

Exercise 3
Omitted

Generalization
The method of Lagrange multipliers can be generalized to functions of n variables $f(\mathbf{x}), \mathbf{x} \in \mathbb{R}^n$ and with r constraints of the form

$$g_1(\mathbf{x}) = 0, \qquad g_2(\mathbf{x}) = 0, \qquad \dots, \qquad g_r(\mathbf{x}) = 0$$

In order to find the possible maximum and minimum points of f subject to the constraints (10.18), one has to find all points **a** such that

$$\nabla f(\mathbf{a}) = \lambda_1 \nabla g_1(\mathbf{a}) + \cdots + \lambda_r \nabla g_r(\mathbf{a}), \;\; and\ g_i(\mathbf{a}) = 0, \;\; 1 \le i \le r$$

The scalars $\lambda_1, \dots, \lambda_r$ are the Lagrange multipliers. When $r = 1$, and $n = 2$ or 3, this reduces to previous algorithms.

Chapter 11 Coordinate Systems

2019年3月23日 18:46

11.1 Polar Coordinates

2019年3月23日 18:46

In a plane we choose a point O called the pole (or origin). From O we draw a ray called the **polar axis**.
Let P be any point in the plane. We will represent the position of P by the ordered pair (r, θ) where $r \geq 0$ is the length of the line OP and θ is the angle between the polar axis and OP. We call r and θ the polar coordinates of P.

Remarks

1. We assume, as usual, that an angle θ is considered positive if measured in the counterclockwise direction from the polar axis and negative if measured in the clockwise direction.
2. We represent the point O by the polar coordinates $(0, \theta)$ for any value of θ.
3. We are restricting r to be non-negative to coincide with te interpretation of r as distance. Many textbooks do not put this restriction on r.
4. Since we use the distance r from the pole in our representation, polar coordinates are suited for solving problems in which there is symmetry about the pole.

Relationship to Cartesian Coordinates

$x = r\cos\theta, \qquad r = \sqrt{x^2 + y^2}$

$y = r\sin\theta, \qquad \tan\theta = \dfrac{y}{x}$

Remark

The equation $\tan\theta = \frac{y}{x}$ does not uniquely determine θ, since over $0 \leq \theta \leq 2\pi$ each value of $\tan\theta$ occurs twice. One must be careful to choose the θ which lies in the correct quadrant.

Graphs in Polar Coordinates

The graph of an explicitly defined polar equation $r = f(\theta)$ or $\theta = f(r)$, or an implicitly defined polar equation $f(r, \theta) = 0$, is a curve that consists of all points that have at least one polar representation (r, θ) that satisfies the equation of the curve.

Exercise 1

Sketch the polar equation $\theta = \frac{\pi}{4}$.

Omitted

Remark

The polar equation $r = e^{\theta}$ gives a **logarithmic spiral** which often appears in nature.

Exercise 2

Sketch the polar equations $r = \sin\theta$ and $r = 1 - 2\cos\theta$.

Exercise 3

Convert the equation of the curve $x^2 - y^2 = 1$ to polar coordinates.

Area in Polar Coordinates

$$A = \lim_{\left|\left|\Delta\theta_i\right|\right| \to 0} \sum_{i=0}^{n-1} \frac{1}{2}\left[f(\theta_i^*)\right]^2 \Delta\theta = \int_a^b \frac{1}{2}\left[f(\theta)\right]^2 d\theta$$

Exercise 4

Find the area inside the lemniscate $r = 2\sqrt{\sin 2\theta}$.

Algorithm
To find the area between two curves in Polar coordinates, we use the same method we used for doing this in Cartesian coordinates.

1. Find the points of intersections.
2. Graph the curves and split the desired region into easily integrable regions.
3. Integrate

Remark
Finding points of intersection can be tricky, especially at the pole/origin which does not have a unique representation: $(0, \theta)$ for any θ represents the origin, so simply setting expressions equal to each other may 'miss' that point. It is essential to sketch the region when finding points of intersection.

Exercise 5
Find the area between the curves $r = \cos\theta$ and $r = \sin\theta$.

11.2 Cylindrical Coordinates

2019年3月23日 22:58

Observe that we can extend polar coordinates to 3-dimensional space by introducing another axis, called the **axis of symmetry**, through the pole perpendicular to the polar plane.

Remark
Notation for cylindrical coordinates may vary from author to author. In particular, in the sciences they generally use the Standard ISO 31-11 notation which gives the cylindrical coordinates as (ρ, ϕ, z).

$$x = r\cos\theta, \qquad r = \sqrt{x^2 + y^2}$$
$$y = r\sin\theta, \qquad \tan\theta = \frac{y}{x}$$
$$z = z, \qquad z = z$$

Remark
Cylindrical coordinates are useful when there is symmetry about an axis. Thus, it is sometimes desirable to lie the polar axis and axis of symmetry along different axes.

Graphs in Cylindrical Coordinates
Exercise 1
Sketch the graph of $z = r^2$ in cylindrical coordinates.

Exercise 2
Find the equation of $z = \frac{y}{\sqrt{x^2+y^2}}$ in cylindrical coordinates.

11.3 Spherical Coordinates

2019年3月25日 6:04

We now extend this idea to another 3-dimensional coordinate system called **spherical coordinates**.

Remark
The symbols used for spherical coordinates also vary from author to author as does the order in which they are written. In mathematics, it is not uncommon to find ρ replaced by r. The standard ISO 31-11 convention uses ϕ as the polar angle and θ as the angle with the positive z-axis. Therefore, it is very important to understand which notation is being used when reading an article.

$$x = \rho \sin\phi \cos\theta\,, \qquad \rho = \sqrt{x^2 + y^2 + z^2}$$
$$y = \rho \sin\phi \cos\theta\,, \qquad \tan\theta = \frac{y}{x}$$
$$z = \rho \cos\phi\,, \qquad \cos\phi = \frac{z}{\sqrt{x^2 + y^2 + z^2}}$$

Graphs in Spherical Coordinates

Exercise 1
Convert $x^2 + y^2 + z^2 = 2x$ to spherical coordinates.

Chapter 12 Mappings of $\mathbb{R}^2$ into $\mathbb{R}^2$

2019年3月25日 6:13

Definition
Vector-Valued Function

A function whose domain is a subset of $\mathbb{R}^n$ and whose codomain is $\mathbb{R}^m$ is called a **vector-valued** function.

Remark
While we represent $(f(t), g(t))$ as a point in $\mathbb{R}^2$, remember that it can also be thought of as a position vector.

Definition
Mapping

A vector-valued function whose domain is a subset of $\mathbb{R}^n$ and whose codomain is a subset of $\mathbb{R}^n$ is called a **mapping** (or transformation).

12.1 The Geometry of Mappings

2019年3月25日 6:22

A pair of equations

$$u = f(x, y)$$
$$v = g(x, y)$$

Associates with each point $(x, y) \in \mathbb{R}^2$ a single point $(u, v) \in \mathbb{R}^2$, and thus defines a vector-valued function

$$(u, v) = F(x, y) = \big(f(x, y), g(x, y)\big)$$

The scalar functions f and g are called the **component functions** of the mapping.

In general, if a mapping F from $\mathbb{R}^2$ to $\mathbb{R}^2$ acts on a curve C in its domain, it will determine a curve in its range, denoted by $F(C)$ and called the **image of C under F**.

More generally, if a mapping F from $\mathbb{R}^2$ to $\mathbb{R}^2$ acts on any subset S in its domain it will determine a set $F(S)$ in its range, called the **image of S under F**.

Exercise 1
Find the image of the circle $(x-1)^2 + y^2 = 1$ under the mapping F defined in Example 1.

Remarks
1. Observe that each of the images are exactly what we could get if we sketched the equations as in Chapter 11.
2. The mapping from polar coordinates to Cartesian coordinates in non-linear. The image of a straight line is not necessarily a straight line.

Exercise 2
Find the image of the square

$$S = \{(x, y) \mid 1 \le x \le 2, 2 \le y \le 3\}$$

Under the mapping defined by

$$(u, v) = F(x, y) = (xy, y)$$

12.2 The Linear Approximation of a Mapping

2019年3月25日 13:06

Definition
Derivative Matrix

The **derivative matrix** of a mapping defined by

$$F(x,y) = \big(f(x,y), g(x,y)\big)$$

Is denoted DF and defined by

$$DF = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{bmatrix}$$

If we introduce the column vectors

$$\Delta\mathbf{u} = \begin{bmatrix} \Delta u \\ \Delta v \end{bmatrix}, \qquad \Delta\mathbf{x} = \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}$$

Then the **increment form of the linear approximation for mappings** becomes

$$\Delta\mathbf{u} \approx DF(a,b)\Delta\mathbf{x}$$

For $\Delta\mathbf{x}$ sufficiently small. Thus, the **linear approximation for mappings** is

$$F(x,y) \approx F(a,b) + DF(a,b)\Delta\mathbf{x}$$

Exercise 1
Consider the mapping defined by

$$(u,v) = F(x,y) = \big(\ln(x+y), \ln(x-y)\big)$$

Approximate the image of the point $(0.95, 0.1)$ under F.

Generalization
A mapping F from $\mathbb{R}^n$ to $\mathbb{R}^m$ is defined by a set of m component functions:

$$u_1 = f_1(x_1, \dots, x_n)$$
$$\vdots$$
$$u_m = f_m(x_1, \dots, x_n)$$

Or, in vector notation

$$\mathbf{u} = F(\mathbf{x}) = \big(f_1(\mathbf{x}), \dots, f_m(\mathbf{x})\big), \qquad \mathbf{x} \in \mathbb{R}^n$$

If we assume that F has continuous partial derivatives, then the derivative matrix of F is the $m \times n$ matrix defined by

$$DF(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

As expected, the linear approximation for F at $\mathbf{a}$ is

$$F(\mathbf{x}) \approx F(\mathbf{a}) + DF(\mathbf{a})\Delta\mathbf{x}$$

Where

$$\Delta\mathbf{u} = \begin{bmatrix} \Delta u_1 \\ \vdots \\ \Delta u_m \end{bmatrix} \in \mathbb{R}^m, \qquad \Delta\mathbf{x} = \begin{bmatrix} \Delta x_1 \\ \vdots \\ \Delta x_n \end{bmatrix} \in \mathbb{R}^n$$

12.3 Composite Mappings and the Chain Rule

2019年3月25日 18:05

Theorem 1
(Chain Rule in Matrix Form)

Let F and G be mappings from $\mathbb{R}^2$ to $\mathbb{R}^2$. If G has continuous partial derivatives at (x,y) and F has continuous partial derivatives at $(u,v)=G(x,y)$, then the composite mapping $F\circ G$ has continuous partial derivatives at (x,y) and

$$D(F\circ G)(x,y)=DF(u,v)DG(x,y)$$

Proof:
Define the component functions for F, G, and $F\circ G$ as in equations (12.1) and (12.2). Then, the chain rule for scalar functions gives

$$\begin{aligned}
DF(u,v)DG(x,y) &= \begin{bmatrix} \frac{\partial p}{\partial u} & \frac{\partial p}{\partial v} \\ \frac{\partial q}{\partial u} & \frac{\partial q}{\partial v} \end{bmatrix} \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix} \\
&= \begin{bmatrix} \frac{\partial p}{\partial u}\frac{\partial u}{\partial x}+\frac{\partial p}{\partial v}\frac{\partial v}{\partial x} & \frac{\partial p}{\partial u}\frac{\partial u}{\partial y}+\frac{\partial p}{\partial v}\frac{\partial v}{\partial y} \\ \frac{\partial q}{\partial u}\frac{\partial u}{\partial x}+\frac{\partial q}{\partial v}\frac{\partial v}{\partial x} & \frac{\partial q}{\partial u}\frac{\partial u}{\partial y}+\frac{\partial q}{\partial v}\frac{\partial v}{\partial y} \end{bmatrix} \\
&= \begin{bmatrix} \frac{\partial p}{\partial x} & \frac{\partial p}{\partial y} \\ \frac{\partial q}{\partial x} & \frac{\partial q}{\partial y} \end{bmatrix} \\
&= D(F\circ G)(x,y)
\end{aligned}$$

As required.

Exercise 1
Consider the mappings defined by

$$F(u,v)=\left(u^2v, e^{uv-1}\right), \qquad G(x,y)=\left(\sqrt{2x^2+2y^2}, 2x+y^2\right)$$

1. Use the chain rule in matrix form to find the derivative matrix $D(F\circ G)$.
2. Calculate $D(G\circ F)(1,1)$.
3. Use the linear approximation of mappings to approximate the image of $(u,v)=(1.01, 0.98)$ under $G\circ F$.

Chapter 13 Jacobians and Inverse Mappings

2019年3月25日 18:29

13.1 The Inverse Mapping Theorem

2019年3月25日 18:30

Definition
Invertible Mapping
Inverse Mapping

Let F be a mapping from a set D_{xy} onto a set D_{uv}. If there exists a mapping F^{-1}, called the **inverse of** F which maps D_{uv} onto D_{xy} such that

$$(x,y) = F^{-1}(u,v) \text{ if and only if } (u,v) = F(x,y)$$

Then F is said to be **invertible** on D_{xy}.

Definition
One-to-One

A mapping F from $\mathbb{R}^2$ to $\mathbb{R}^2$ is said to be **one-to-one** on a set D_{xy} if and only if $F(a,b) = F(c,d)$ implies $(a,b) = (c,d)$, for all $(a,b),(c,d) \in D_{xy}$.

Theorem 1
Let F be a mapping from a set D_{xy} onto a set D_{uv}. If F is one-to-one on D_{xy}, then F is invertible on D_{xy}.

Theorem 2
Consider a mapping F which maps D_{xy} onto D_{uv}. If F has continuous partial derivatives at $\mathbf{x} \in D_{xy}$ and there exists an inverse mapping F^{-1} of F which has continuous partial derivatives at $\mathbf{u} = F(\mathbf{x}) \in D_{uv}$, then

$$DF^{-1}(\mathbf{u})DF(\mathbf{x}) = I$$

Proof:
By the Chain Rule in Matrix Form we get

$$DF^{-1}(\mathbf{u})DF(\mathbf{x}) = D(F^{-1} \circ F)(\mathbf{x})$$

Then, by equation (13.1) we have

$$D(F^{-1} \circ F)(\mathbf{x}) = D\mathbf{x} = \begin{bmatrix} \frac{\partial x}{\partial x} & \frac{\partial x}{\partial y} \\ \frac{\partial y}{\partial x} & \frac{\partial y}{\partial y} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

As required.

Remark
The fact that we could solve and obtain a unique solution for x and y in the preceding example proves that F has an inverse mapping on $\mathbb{R}^2$. It is only in simple examples that one can carry out this step. Hence it is useful to develop a test to determine if a mapping F has an inverse mapping.

Definition
Jacobian

The **Jacobian** of a mapping

$$(u,v) = F(x,y) = \big(u(x,y), v(x,y)\big)$$

Is denoted $\frac{\partial(u,v)}{\partial(x,y)}$, and is defined by

$$\frac{\partial(u,v)}{\partial(x,y)} = \det[DF(x,y)] = \det\begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix}$$

Exercise 1
Calculate the Jacobian $\frac{\partial(x,y)}{\partial(r,\theta)}$ of the mapping F given by

$$(x,y) = F(r,\theta) = (r\cos\theta, r\sin\theta)$$

Corollary 3
Consider a mapping defined by

$$(u,v) = F(x,y) = \big(f(x,y), g(x,y)\big)$$

Which maps a subset D_{xy} onto a subset D_{uv}. Suppose that f and g have continuous partials on D_{xy}. If F has an inverse mapping F^{-1}, with continuous partials on D_{uv}, then the Jacobian of F is non-zero:

$$\frac{\partial(u,v)}{\partial(x,y)} \neq 0, \qquad \text{on } D_{xy}$$

Remark
The notation $\frac{\partial(u,v)}{\partial(x,y)}$ for the Jacobian reminds one which partial derivatives have to be calculated. Thus, if F maps $(x,y) \to (u,v)$ and is one-to-one, then the inverse mapping F^{-1} maps $(u,v) \to (x,y)$, and the Jacobian of the inverse mapping is denoted by

$$\frac{\partial(x,y)}{\partial(u,v)} = \det[F^{-1}(u,v)] = \det\begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix}$$

Recall from linear algebra that $\det(AB) = \det A \det B$ for all $n \times n$ matrices A, B. Thus, we can deduce from Theorem 2 a simple relationship between the Jacobian of a mapping and the Jacobian of the inverse mapping. We state this as a corollary to Theorem 2.

Corollary 4
(Inverse Property of the Jacobian)

If the hypotheses of Theorem 2 hold, then

$$\frac{\partial(x,y)}{\partial(u,v)} = \frac{1}{\frac{\partial(u,v)}{\partial(x,y)}}$$

Proof:
By Theorem 2,

$$I = DF^{-1}(u,v)DF(x,y)I$$

Taking the determinant of this equation gives

$$\det I = \det\left(DF^{-1}(u,v)DF(x,y)\right)$$
$$1 = \det(DF^{-1}(u,v))\det\left(DF(x,y)\right)$$

Thus, by definition of the Jacobian,

$$1 = \frac{\partial(x,y)}{\partial(u,v)}\frac{\partial(u,v)}{\partial(x,y)}$$

Since $DF(x,y)$ is invertible, we have $\frac{\partial(u,v)}{\partial(x,y)} \neq 0$. Therefore, we get

$$\frac{\partial(x,y)}{\partial(u,v)} = \frac{1}{\frac{\partial(u,v)}{\partial(x,y)}}$$

Theorem 5
(Inverse Mapping Theorem)

If a mapping $(u, v) = F(x, y)$ has continuous partial derivatives in some neighborhood of (a, b) and $\frac{\partial(u,v)}{\partial(x,y)} \neq 0$ at (a, b), then there is a neighborhood of (a, b) in which F has an inverse mapping $(x, y) = F^{-1}(u, v)$ which has continuous partial derivatives.

Exercise 2
Referring to Example 3, show that the inverse mapping is given by

$$(x, y) = F^{-1}(u, v) = \left(\frac{1}{4}\left(v + \sqrt{v^2 - 8u}\right), \frac{1}{4}\left(3v - \sqrt{v^2 - 8u}\right)\right)$$

13.2 Geometrical Interpretation of the Jacobian

2019年3月26日 12:38

Exercise 1
Let $F(x,y) = (x^2y, -xy)$ and let S be the square pictured in the diagram. Will the image of S under F have more or less area? Explain your answer.

Remark
For a linear mapping $(u,v) = F(x,y) = (ax+by, cx+dy)$ where a,b,c,d are constants, the derivative matrix is

$$DF(x,y) = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

And this the linear approximation is exact by Taylor's Theorem since all second partials are zero. Therefore, for a linear mapping, the approximation (13.3) becomes an exact relation.

Exercise 2
Show that the linear mapping $(u,v) = F(x,y) = (x+2y, x+y)$ preserves area. Illustrate the action of the mapping by finding the image of the square with vertices (0,0), (0,1), (1,0) and (1,1).

Exercise 3
Use the Jacobian to verify the well-known result that any linear mapping F which is a rotation,

$$(u,v) = F(x,y) = (x\cos\theta + y\sin\theta, -x\sin\theta + y\cos\theta)$$

Where θ is a constant, preserves areas.

Generalization
At the end of Section 12.2, we generalized the concept of a mapping F from $\mathbb{R}^2$ to $\mathbb{R}^2$ to a mapping F from $\mathbb{R}^n$ to $\mathbb{R}^m$, and defined the Jacobian of the mapping, as follows.

Definition
Jacobian

For a mapping defined by

$$\mathbf{u} = F(\mathbf{x}) = (f_1(\mathbf{x}), \ldots, f_n(\mathbf{x}))$$

Where $\mathbf{u} = (u_1, \ldots, u_n)$ and $\mathbf{x} = (x_1, \ldots, x_n)$, the **Jacobian** of F is

$$\frac{\partial(u_1,\ldots,u_n)}{\partial(x_1,\ldots,x_n)} = \det[DF(x,y)] = \det\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}$$

Geometrical Interpretation of the Jacobian in 3-D

13.3 Constructing Mappings

2019年3月26日 20:16

Exercise 1
Find a linear mapping F which will transform the ellipse $3x^2 + 2xy + y^2 = 4$ into the circle $u^2 + v^2 = 4$.

Exercise 2
Find an invertible mapping which will transform the region D_{xyz} in the first octant bound by $xy = 1, xy = 3, xz = 1, xz = 3, yz = 2, yz = 4$ into a cube in the uvw-space.

Chapter 14 Double Integrals

2019年3月28日 11:49

14.1 Definition of Double Integrals

2019年3月28日 11:49

Definition
Integrable

Let $D \subset \mathbb{R}^2$ be closed and bounded. Let P be a partition of D as described above, and let $|\Delta P|$ denote the length of the longest side of all rectangles in the partition P. A function $f(x, y)$ which is bounded on D is **integrable** on D if all Riemann sums approach the same value as $|\Delta P| \to 0$.

Definition
Double Integral

If $f(x, y)$ is integrable on a closed bounded set D, then we define the **double integral** of f on D as

$$\iint_D f(x, y)\ dA = \lim_{\Delta P \to 0} \sum_{i=1}^{n} f(x_i, y_i) \Delta A_i$$

Interpretation of the Double Integral
When you encounter the double integral symbol

$$\iint_D f(x, y)\, dA$$

Think of "limit of a sum". In itself, the double integral is a mathematically defined object. It has many interpretations depending on the meaning that you assign to the integrand $f(x, y)$. The "dA" in the double integral symbol should remind you of the area of a rectangle in a partition of D.

Double Integral as Area:
The simplest interpretation is when you specialize f to be the constant function with value unity:

$$f(x, y) = 1, \qquad \text{for all } (x, y) \in D$$

Then the Riemann sum (14.1) simply sums the areas of all rectangles in D, and the double integral serves to define the area $A(D)$ of the set D:

$$A(D) = \iint_D 1\, dA$$

Double Integral as Volume:
If $f(x, y) \geq 0$ for all $(x, y) \in D$, then the double integral

$$\iint_D f(x, y)\, dA$$

Can be interpreted as the volume $V(S)$ of the origin defined by

$$S = \{(x, y, z) \,|\, 0 \leq z \leq f(x, y), (x, y) \in D\}$$

Which represents the solid below the surface $z = f(x, y)$ and above the set D in the xy-plane.
The justification is as follows.
The partition P of D decomposes the solid S into vertical "columns". The height of the column above the i-th rectangle is approximately $f(x_i, y_i)$, and so its volume is approximately

$$f(x_i, y_i) \Delta A_i$$

The Riemann sum (14.1) thus approximates the volume $V(S)$:

$$V(S) \approx \sum_{i=1}^{n} f(x_i, y_i) \Delta A_i$$

As $|\Delta P| \to 0$ the partition becomes increasingly fine, so the error in the approximation will tend to zero. Thus, the volume $V(S)$ is

$$V(S) = \iint_D f(x,y)\, dA$$

Double Integral as Mass:

Think of a thin flat plate of metal whose density varies with position. Since the plate is thin, it is reasonable to describe the varying density by an "area density", that is a function $f(x,y)$ that gives the mass per unit area at position (x,y). In other words, the mass of a small rectangle of area ΔA_i located at position (x_i, y_i) will be approximately

$$M_i \approx f(x_i, y_i)\Delta A_i$$

The Riemann sum (14.1) corresponding to a partition P of D will approximate the total mass M of the plate D, and the double integral of f over D, being the limit of the sum, will represent the total mass:

$$M = \iint_D f(x,y)\, dA$$

Double Integral as Probability:

Let $f(x,y)$ be the probability density of a continuous 2-D random variable (X,Y). The probability that $(X,Y) \in D$, a given subset $\mathbb{R}^2$, is

$$P\big((X,Y) \in D\big) = \iint_D f(x,y)\, dA$$

Average Value of a Function:

The double integral is also used to define the average value of a function $f(x,y)$ over a set $D \subset \mathbb{R}^2$.

Recall for a function of one variable, $f(x)$, the average value of f over an interval $[a,b]$, denoted f_{av}, is defined by

$$f_{av} = \frac{1}{b-a}\int_a^b f(x)\, dx$$

Similarly, for a function of two variables $f(x,y)$, we can define the average value of f over a closed and bounded subset D of $\mathbb{R}^2$ by

$$f_{av} = \frac{1}{A(D)}\iint_D f(x,y)\, dA$$

Exercise 1

A city occupies a region D of the xy-plane. The population density in the city (measured as people/unit area) depends on position (x,y), and is given by a function $p(x,y)$. Interpret the double integral $\iint_D p(x,y)\, dA$

Properties of the Double Integral

Theorem 1

(Linearity)

If $D \subset \mathbb{R}^2$ is a closed and bounded set and f and g are two integrable functions on D, then for any constant c:

$$\iint_D (f+g)\, dA = \iint_D f\, dA + \iint_D g\, dA$$

$$\iint_D cf\, dA = c\iint_D f\, dA$$

Theorem 2

(Basic Inequality)

If $D \subset \mathbb{R}^2$ is a closed and bounded set and f and g are two integrable functions on D such that $f(x,y) \le g(x,y)$ for all $(x,y) \in D$, then

$$\iint_D f\, dA \le \iint_D g\, dA$$

Theorem 3

(Absolute Value Inequality)

If $D \subset \mathbb{R}^2$ is a closed and bounded set and f is an integrable function on D, then

$$\left|\iint_D f\, dA\right| \le \iint_D |f|\, dA$$

Theorem 4
(Decomposition)

Assume $D \subset \mathbb{R}^2$ is a closed and bounded set and f is an integrable function on D. If D is decomposed into two closed and bounded subsets D_1 and D_2 by a piecewise smooth curve C, then

$$\iint_D f\, dA = \iint_{D_1} f\, dA + \iint_{D_2} f\, dA \qquad D = D_1 \cup D_2$$

Remarks

1. The Basic Inequality can be used to obtain an estimate for a double integral that cannot be evaluated exactly.
2. The decomposition property is essential for dealing with complicated regions of integration and with discontinuous integrands.

14.2 Iterated Integrals

2019年3月28日 21:51

Theorem 1
Let $D \subset \mathbb{R}^2$ be defined by

$y_\ell(x) \le y \le y_u(x), \quad$ and $\quad x_\ell \le x \le x_u$

Where $y_\ell(x)$ and $y_u(x)$ are continuous for $x_\ell \le x \le x_u$. If $f(x,y)$ is continuous on D, then

$$\iint_D f(x,y)\,dA = \int_{x_\ell}^{x_u}\int_{y_\ell(x)}^{y_u(x)} f(x,y)\,dy\,dx$$

Remark
Although the parentheses around the inner integral are usually omitted, we must evaluate it first. Moreover, as in our interpretation of volume above, when evaluating the inner integral, we are integrating with respect to y while holding x constant. That is, we are using **partial integration**.

Suppose now that the set D can be described by inequalities of the form

$x_\ell(y) \le x \le x_u(y), \quad$ and $y_\ell \le y \le y_u$

Where y_ℓ, y_u are constants and $x_\ell(y), x_u(y)$ are continuous functions of y on the interval

$y_\ell \le y \le y_u$

Then, by reversing the roles of x and y in Theorem 1, the double integral $\iint_D f(x,y)\,dA$ can be written as in iterated integral in the order "x first, then y":

$$\iint_D f(x,y)\,dA = \int_{y_\ell}^{y_u}\int_{x_\ell(y)}^{x_u(y)} f(x,y)\,dx\,dy$$

Exercise 1
Describe the set D by inequalities in two ways. Evaluate the double integral

$$\iint_D (x+y)\,dA$$

In two ways.

Exercise 2
Let D be the triangular region with vertices (0,0), (1,1), and (0,2). Evaluate

$$\iint_D y\,dA$$

Exercise 3
Let D be the triangular region with vertices (0,0), (0,1), and (1,1). Evaluate

$$\iint_D e^{-y^2}\,dA$$

Exercise 4
Find the volume of the solid bounded above by the paraboloid $z = 4 - x^2 - y^2$, and below by the rectangle $D = \{(x,y) | 0 \le x \le 1, 0 \le y \le 1\}$.

14.3 The Change of Variable Theorem

2019年3月29日 20:15

Theorem 1
(Change of Variable Theorem)

Let each of D_{uv} and D_{xy} be a closed bounded set whose boundary is a piecewise-smooth closed curve. Let

$$(x, y) = F(u, v) = (f(u, v), g(u, v))$$

Be a one-to-one mapping of D_{uv} onto D_{xy}, with $f, g \in C^1$, and $\frac{\partial(x,y)}{\partial(u,v)} \neq 0$ expect for possibly on a finite collection of piecewise-smooth curves in D_{uv}. If $G(x, y)$ is continuous on D_{xy}, then

$$\iint_{D_{xy}} G(x, y)\, dx\, dy = \iint_{D_{uv}} G(f(u, v), g(u, v)) \left|\frac{\partial(x, y)}{\partial(u, v)}\right| du\, dv$$

Exercise 1
Omitted

Double Integrals in Polar Coordinates

Remark
Because polar coordinates have a simple geometric interpretation, one can obtain the r and θ limits of integration directly from the diagram in the xy-plane, without drawing the region $D_{r\theta}$ in the same way as we did for finding areas in polar coordinates in Chapter 11. The method is illustrated in the diagram.

Exercise 2
Omitted
Exercise 3
Omitted
Exercise 4
Omitted

Chapter 15 Triple Integrals

2019年3月29日 20:42

15.1 Definition of Triple Integrals

2019年3月29日 20:43

Definition
Integrable

A function $f(x, y, z)$ which is bounded on a closed bounded set $D \subset \mathbb{R}^3$ is said to be **integrable** on D if and only if all Riemann sums approach the same value as $\Delta P \to 0$.

Definition
Triple Integral

If $f(x, y, z)$ is integrable on a closed bounded set D, then we define the **triple integral** of f over D, as

$$\iiint_D f(x, y, z)\, dV = \lim_{\Delta P \to 0} \sum_{i=1}^{n} f(x_i, y_i, z_i) \Delta V_i$$

Interpretation of the Triple Integral
When you encounter the triple integral symbol

$$\iiint_D f(x, y, z)\, dV$$

You should think of "limit of a sum". In itself, the triple integral is a mathematically defined object. It has many interpretations, depending on the interpretation that you assign to the integrand $f(x, y, z)$. The "dV" in the triple integral symbol should remind you of the volume of a rectangular block in a partition of D.

Triple Integral as Volume:
The simplest interpretation is when you specialize f to be the constant function with value unity:

$$f(x, y, z) = 1, \qquad \text{for all } (x, y, z) \in D$$

Then, the Riemann sum (15.1) simply sums the volumes of all rectangular blocks in D, and the triple integral over D serves to define the volume $V(D)$ of the set D:

$$V(D) = \iiint_D 1\, dV$$

Triple Integral as Mass:
Think of a planet or star whose density varies with position. Let D denote the subset of $\mathbb{R}^3$ occupied by the star. Let $f(x, y, z)$ denote the density (mass per unit volume) at position (x, y, z). The mass of a small rectangular block located within the star at position (x_i, y_i, z_i) will be approximately

$$\Delta M_i \approx f(x_i, y_i, z_i) \Delta V_i$$

Thus, the Riemann sum corresponding to a partition P of D

$$\sum_{i=1}^{n} f(x_i, y_i, z_i) \Delta V_i$$

Will approximate the total mass M of the star, and the triple integral of f over D, being the limit of the Riemann sum, will represent the total mass:

$$M = \iiint_D f(x, y, z)\, dV$$

Average Value of a Function:

Definition
Average Value

Let $D \subset \mathbb{R}^3$ be closed and bounded with volume $V(D) \neq 0$, and let $f(x, y, z)$ be a bounded and integrable function on D. The **average value** of f over D is defined by

$$f_{avg} = \frac{1}{V(D)} \iiint_D f(x, y, z)\, dV$$

Remark
If you have the impression that you have read this section someplace else, you're right. Compare it with Section 14.1. The only essential change is to replace "area" with "volume".

Properties of the Triplpe Integral
Theorem 1
(Linearity)

If $D \subset \mathbb{R}^3$ is a closed and bounded set, c is a constant, and f and g are two integrable functions on D, then

$$\iiint_D (f + g)\, dV = \iiint_D f\, dV + \iiint_D g\, dV$$
$$\iiint_D cf\, dV = c \iiint_D f\, dV$$

Theorem 2
(Basic Inequality)

If $D \subset \mathbb{R}^3$ is a closed and bounded set and f is an integrable function on D. If D is decomposed into two closed and bounded subsets D_1 and D_2 by a piecewise smooth curve C, then

$$\iiint_D f\, dV = \iiint_{D_1} f\, dV + \iiint_{D_2} f\, dV$$

15.2 Iterated Integrals

2019年3月30日 1:03

Theorem 1
Let D be the subset of $\mathbb{R}^3$ defined by

$$z_\ell(x,y) \le z \le z_u(x,y) \text{ and } (x,y) \in D_{xy}$$

Where z_ℓ and z_u are continuous functions on D_{xy}, and D_{xy} is a closed bounded subset in $\mathbb{R}^2$, whose boundary is a piecewise smooth closed curve. If $f(x,y,z)$ is continuous on D, then

$$\iiint_D f(x,y,z)\, dV = \iint_{D_{xy}} \int_{z_\ell(x,y)}^{z_u(x,y)} f(x,y,z)\, dz\, dA$$

Remark
As with double iterated integrals, we are doing partial integration. That is, to evaluate the inner integral of

$$\iint_{D_{xy}} \int_{z_\ell(x,y)}^{z_u(x,y)} f(x,y,z)\, dz\, dA$$

We hold x and y constant and integrate with respect to z.

Exercise 1
Exercise 2
Exercise 3
Exercise 4
Exercise 5
Omitted

15.3 The Change of Variable Theorem

2019年3月30日 1:12

Theorem 1
(Change of Variable Theorem)

Let

$$x = f(u, v, w), \qquad y = g(u, v, w), \qquad z = h(u, v, w)$$

Be a one-to-one mapping of D_{uvw} onto D_{xyz}, with f, g, h having continuous partials, and

$$\frac{\partial(x, y, z)}{\partial(u, v, w)} \neq 0 \quad \text{on } D_{uvw}$$

If $G(x, y, z)$ is continuous on D_{xyz}, then

$$\iiint_{D_{xyz}} G(x, y, z)\, dV = \iiint_{D_{uvw}} G\big(f(u, v, w), g(u, v, w), h(u, v, w)\big) \left| \frac{\partial(x, y, z)}{\partial(u, v, w)} \right| dV$$

Exercise 1
Exercise 2
Omitted

Triple Integrals in Cylindrical Coordinates
Exercise 3
Exercise 4
Omitted

Triple Integrals in Spherical Coordinates
Exercise 5
Exercise 6
Exercise 7
Omitted