2MVAa
Multivariable & Vector Analysis 1

Instructor: **Dr Qianxi Wang**
Research Fields : Applied Mathematics & Fluid Mechanics

Office: **The Watson Building, Room 325**
Tel: **0121 414 6602**
Email: **q.x.wang@bham.ac.uk**
Office hours: **2:30-4:00 pm on Thursday**

Outline of syllabus

2MVAa is a general module, containing both "pure" and "applied" materials. It includes formal definitions, theorems, proofs and applications. The syllabus is outlined as follows:

- Functions of several real variables, partial derivatives, gradient vectors, chain rule, implicit functions, higher order partial derivatives.
- Multiple integrals, cylindrical and spherical polars, change of variables using the Jacobian.
- Taylor series.
- Stationary points, maxima and minima of functions of several variables, leading minor test. Lagrange multipliers.

Notes

- Please email me corrections to the notes, the problem sheets and solutions to (wangqx@maths.bham.ac.uk). Contact me use this email address, as I do not check my Cavas emails everyday.
- Important equations will be framed in the notes.
- You are advised to remember those equations framed with solid lines, as shown in the example,

$$\boxed{\nabla f = \text{grad} f \equiv \frac{\partial f}{\partial x}\boldsymbol{i} + \frac{\partial f}{\partial y}\boldsymbol{j} + \frac{\partial f}{\partial z}\boldsymbol{k}}$$

Lectures & tutorials

- **Please be on time** since it is distracting to have people walking in late.
- I will do my best to be clear but **you must read through and understand the notes before next lecture**. Otherwise, you will get hopelessly lost.
- If anyone is colour blind, please tell me which colour you cannot read.

Assessment

2MVA is a 20 credit module and will be assessed as follows:

- Written unseen examination (3hrs) in June – 80% and continuous assessment – 20%
- 2CAs for 2MVAa and 2CAs for 2MVAb
- CA marks are counted based on the best 3 out of 4 over 2 terms.
- The exam has two sections. Section A is compulsory and is worth 50% (25% on materials from 2MVAa and 25% from 2MVAb); Section B has 4 questions (2 from 2MVAa and 2 from 2MVAb) each with 17%. You should choose 3 out 4 questions in Section B.
- The examination contents for 2MVAa will be the combinations of the **notes**, the **added examples** discussed in class, and the **problem sheets**.

Continuous assessment in Autumn Term: 10%

- **2 problem sheets each 5%**
- **Deadlines for the problem sheets are at 12 noon on Friday in weeks 5, 9.**
- **Please staple each assignment to avoid work going missing, and use only pens (black or blue), not pencils.**
- **Work will be returned via the trays next to front desk of TSO.**
- **Marks will be posed on the Canvas within two weeks after handed in.**
- **You should check to ensure the marks have been entered correctly.**

Example classes

- Examples/Feedback Class at **1PM on Friday** in weeks 2, 4, 6, 8, 10, when you may ask

  Questions on notes

  Questions in problem sheets
- The only way to learn maths is to do maths. Lectures, notes and books tell you what materials you need to learn.
- Doing problems, making mistakes and asking questions is how you learn and understand maths.
- 4 example sheets with solutions will be provided for the purpose of your exercise.

Reference books

The recommended text for 2MVAa is

Adams, R. A. Calculus: a complete course, Addison Wesley.

The following two books also cover most of the materials in 2MVAa and are worth having a look at.

Thomas, R. L. Calculus. Addison Wesley.

Kreyszig, E. Advanced Engineering Mathematics. Wiley.

I. Functions of Several Real Variables

Adams, R. A. Calculus: a complete course, Addison Wesley. 7th Edition

1.1 Functions of several real variables

Definition 1.1 A function f whose domain is $\mathbb{R}^n$ or subset of $\mathbb{R}^n$, for $n \geq 2$ and $n \in \mathbb{N}$, is called a function of several real variables.

We summarize this information using the following notation

$$f : D \subseteq \mathbb{R}^n \to R \subseteq \mathbb{R},$$

where D is the domain of the f and R is the range (codomain) of f.

For convenience, we often just write it as

$$f : \mathbb{R}^n \to \mathbb{R}.$$

In other words, if a function depends on more than one real variables, it is naturally called a function of several real variables.

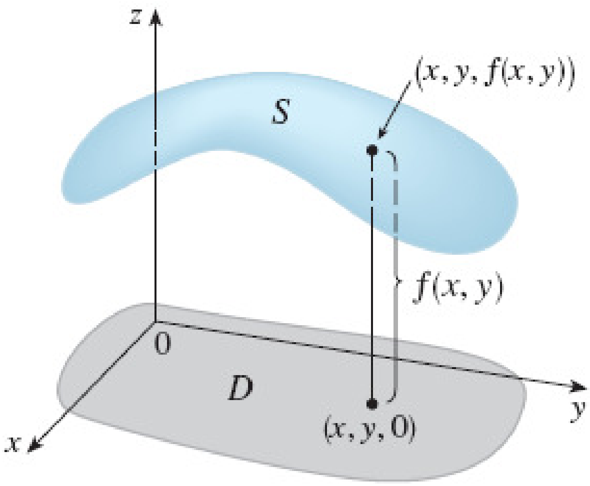
We often use vectors to simplify the notation, so that $\boldsymbol{a}$ might be interpreted as (a, b) or (a, b, c), and $f(\boldsymbol{x})$ might mean $f(x, y)$, $f(x, y, z)$ or $f(x_1, x_2, \ldots, x_n)$ depending on the circumstance.

Geometrical interpretation of $z=f(x, y)$

We will mostly concern with $f: \mathbb{R}^2 \to \mathbb{R}$ and $f: \mathbb{R}^3 \to \mathbb{R}$.

A smooth function $z = f(x, y)$, $(x, y) \in D$ will be a surface in $\mathbb{R}^3$.

For a function with 3 or more variables we are unable to sketch them.



Comments

The complete definition of a function should include its domain and co-domain. The following two functions are not the same:

$f = \sin(x)$ for $x \in [0, 1]$, $\quad g = \sin(x)$ for $x \in [-\pi, \pi]$.

We often write things such as $f: \mathbb{R}^2 \to \mathbb{R}$ where $f(x, y) = \sqrt{x^2 - y^2}$.

Strictly speaking the domain is not $\mathbb{R}^2$ but a proper subset of it.

Usually this will not be a problem and we know what we mean.

Variables does not always occur explicitly in the expression of a function, such as

$f(x, y) = 3$,

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

Graphing of functions

Definition 1.2. For a function $z = f(x, y)$:

A vertical section is the graph of $z = f(x, c)$ or $z = f(c, y)$ for some constant c. A level curve is the curve $c = f(x, y)$ for some constant c.

Vertical sections: $\boldsymbol{z = f(c, y)}$

Graphs in the yz plane

Choose c=-2, -1, 0, 1, 2, 3. The actual values should be chosen according to nature of the function.

Plot these curves in the same figure.

Vertical sections: $\boldsymbol{z = f(x, c)}$

Graphs in the zx plane

Level curves: $c = f(x, y)$

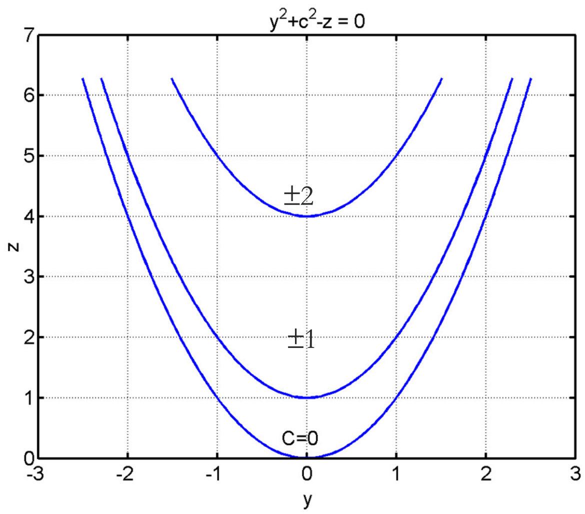
Graphs in the xy plane

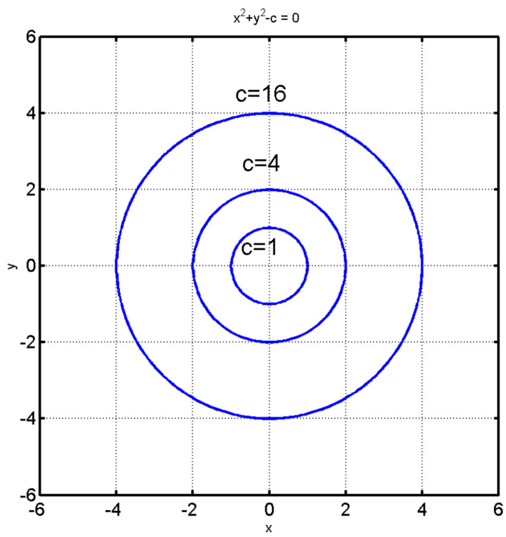
Example 1. ♣

Plot the vertical sections and level curves for $z=x^2+y^2$.

Solution. The vertical sections in the yz plane is $z=c^2+y^2$, which are shown in the left figure. Similarly one can draw the vertical sections in the zx plane.

The level curves $c=x^2+y^2$ is shown in the right figure.





1.2 Partial differentiation and the chain rule

Definition 2.1. The function $f: \mathbb{R}^3 \to \mathbb{R}$ is said to have a partial derivative with respect to x at the point (x_0, y_0, z_0) if the following limit exists

$$\lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x, y_0, z_0) - f(x_0, y_0, z_0)}{\Delta x},$$

which is called the ***partial derivative*** of f with respect to x at the point (x_0, y_0, z_0), denoted as

$$\frac{\partial f(x_0, y_0, z_0)}{\partial x} \quad \text{or} \quad f_x(x_0, y_0, z_0),$$

i.e.

$$\boxed{\frac{\partial f(x_0, y_0, z_0)}{\partial x} \equiv \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x, y_0, z_0) - f(x_0, y_0, z_0)}{\Delta x}}$$

The partial derivative of f with respect to x is the derivative of f with respect to x, when **all other variables are treated as constant**.

Partial derivatives

Derivatives with respect to y and z are defined in a similar way, such as the derivative with respect to y at the point (x, y, z) being given as follows

$$\frac{\partial f(x, y, z)}{\partial y} \equiv \lim_{\Delta y \to 0} \frac{f(x, y + \Delta y, z) - f(x, y, z)}{\Delta y}$$

Example 1. Calculate f_x, f_y and f_z for $f = x\sin y + z$.

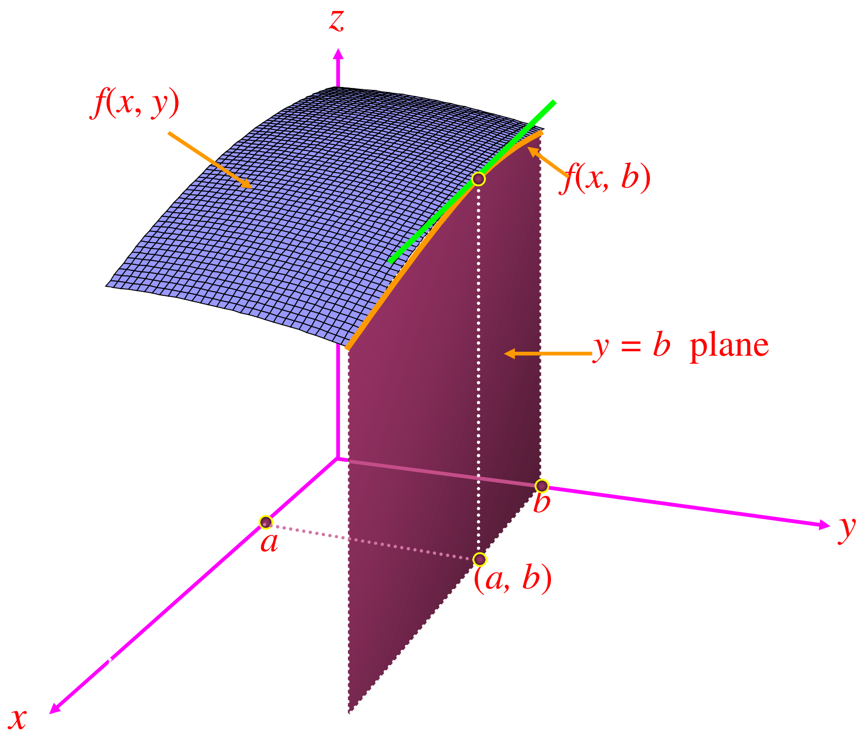
Solution. $\frac{\partial f}{\partial x} = \sin y, \quad \frac{\partial f}{\partial y} = x\cos y, \quad \frac{\partial f}{\partial z} = 1$

Example 2. Calculate f_y for $f = y\sin(xy)$.

Solution. $\frac{\partial f}{\partial y} = \frac{\partial y}{\partial y}\sin(xy) + y\frac{\partial \sin(xy)}{\partial y}$

$$= \sin(xy) + y\cos(xy)\frac{\partial (xy)}{\partial y} = \sin(xy) + xy\cos(xy)$$

Geometric interpretation of partial derivative $f_x(a, b)$



The partial derivative $f_x(a, b)$ is the slope of the tangent line to the curve $f(x, b)$ at $x = a$.

High order partial derivatives

As the first order partial derivatives of $f(x, y, z)$ are functions of (x, y, z), one can perform partial derivative to them, which are termed as the second order partial derivatives, such as

$$\frac{\partial^2 f}{\partial x^2} \equiv \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) \equiv f_{xx},$$ Differentiate two times with respect to x

$$\frac{\partial^2 f}{\partial y^2} \equiv \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) \equiv f_{yy},$$ Differentiate two times with respect to y

$$\frac{\partial^2 f}{\partial y \partial x} \equiv \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) \equiv f_{xy},$$ Differentiate with respect to x then y

$$\frac{\partial^2 f}{\partial x \partial y} \equiv \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) \equiv f_{yx},$$ Differentiate with respect to y then x

One can define higher order partial derivatives as follows

$$\frac{\partial^3 f}{\partial x^3} \equiv \frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial x^2}\right) \equiv f_{xxx}, \quad \frac{\partial^3 f}{\partial y \partial x^2} \equiv \frac{\partial}{\partial y}\left(\frac{\partial^2 f}{\partial x^2}\right) \equiv f_{xxy}.$$

Example 3.

Calculate all the second order partial derivatives for $f=x\sin y+yz$.

Solution. $\frac{\partial f}{\partial x}=\sin y,\quad \frac{\partial f}{\partial y}=x\cos y+z,\quad \frac{\partial f}{\partial z}=y$

$$\frac{\partial^2 f}{\partial x^2}=\frac{\partial \sin y}{\partial x}=0,\qquad \frac{\partial^2 f}{\partial y\partial x}=\frac{\partial \sin y}{\partial y}=\cos y,\qquad \frac{\partial^2 f}{\partial z\partial x}=\frac{\partial \sin y}{\partial z}=0,$$

$$\frac{\partial^2 f}{\partial x\partial y}=\frac{\partial(x\cos y+z)}{\partial x}=\cos y,\quad \frac{\partial^2 f}{\partial y^2}=\frac{\partial(x\cos y+z)}{\partial y}=-x\sin y,\quad \frac{\partial^2 f}{\partial z\partial y}=\frac{\partial(x\cos y+z)}{\partial z}=1,$$

$$\frac{\partial^2 f}{\partial x\partial z}=\frac{\partial y}{\partial x}=0,\qquad \frac{\partial^2 f}{\partial y\partial z}=\frac{\partial y}{\partial y}=1,\qquad \frac{\partial^2 f}{\partial z^2}=\frac{\partial y}{\partial z}=0,$$

One can see that $\frac{\partial^2 f}{\partial x\partial y}=\frac{\partial^2 f}{\partial y\partial x},\quad \frac{\partial^2 f}{\partial y\partial z}=\frac{\partial^2 f}{\partial z\partial y},\quad \frac{\partial^2 f}{\partial z\partial x}=\frac{\partial^2 f}{\partial x\partial z}.$

Theorem 2. When a function has the second order continuous partial derivatives, the partial derivations of this function do not depend on the order with respect to the variables.

Implicit partial differentiation

Example 4. Find $\frac{\partial z}{\partial x}$ if the equation,

$$yz-\ln z=x+y,$$

defines z as a functions of x and y.

Solution. We differentiate both sides of the eq. to x.

$$\frac{\partial yz}{\partial x}-\frac{\partial \ln z}{\partial x}=\frac{\partial x}{\partial x}+\frac{\partial y}{\partial x}$$

$$yz_x-\frac{z_x}{z}=1+0$$

$$z_x\left(y-\frac{1}{z}\right)=1\qquad z_x=1/\left(y-\frac{1}{z}\right)=\frac{z}{yz\text{-}1}$$

1.3 Differentiable functions and chain rule

L'Hospital's Rule

Let lim stands for the limit of : $\lim\limits_{x\to c}$, $\lim\limits_{x\to+\infty}$, $\lim\limits_{x\to-\infty}$

If $\lim \dfrac{f'(x)}{g'(x)}$ has a finite value or if the limit is $\pm\infty$.

Then $\lim \dfrac{f(x)}{g(x)} = \lim \dfrac{f'(x)}{g'(x)}$

Examples

$$\lim_{x\to 0} \frac{\frac{d\sin(x)}{dx}}{\frac{dx}{dx}} = \lim_{x\to 0} \frac{\cos(x)}{1} = 1 \neq 0 \qquad \Rightarrow \quad \lim_{x\to 0} \frac{\sin(x)}{x} = 1$$

Orders of functions

Let lim stands for the limit of : $\lim\limits_{x\to c}$, $\lim\limits_{x\to+\infty}$, $\lim\limits_{x\to-\infty}$

When $\lim \dfrac{f(x)}{g(x)} = k, \quad k \neq 0,$

We say that $f(x)$ is at the same order of $g(x)$ at the limit, denoted as

$$f(x) = O(g(x))$$

Examples

1. $\lim\limits_{x\to 0} \dfrac{\sqrt{2}x}{x} = \sqrt{2} \neq 0 \qquad \Rightarrow \quad \sqrt{2}x = O(x)$

2. $\lim\limits_{x\to 0} \dfrac{\sin(x)}{x} = \lim\limits_{x\to 0} \dfrac{\frac{d\sin(x)}{dx}}{\frac{dx}{dx}} = \lim\limits_{x\to 0} \dfrac{\cos(x)}{1} = 1 \neq 0 \qquad \Rightarrow \quad \sin(x) = O(x)$

3. $\lim\limits_{x\to 0} \dfrac{\sin(\sqrt{2}x)}{x} = \lim\limits_{x\to 0} \dfrac{\cos(\sqrt{2}x)\sqrt{2}}{1} = \sqrt{2} \qquad \Rightarrow \quad \sin(\sqrt{2}x) = O(x)$

Orders of variables (cont.)

Let lim stands for the limit of : $\lim\limits_{x\to c}$, $\lim\limits_{x\to+\infty}$, $\lim\limits_{x\to-\infty}$

When $\lim \dfrac{f(x)}{g(x)} = 0$,

We say that $f(x)$ is small compared to $g(x)$ at the limit, denoted as

$$f(x) = o(g(x))$$

Examples

1. $\lim\limits_{x\to 0} \dfrac{\sqrt{2}x}{\sqrt{x}} = 0 \quad \Rightarrow \quad \sqrt{2}x = o(\sqrt{x})$

2. $\lim\limits_{x\to 0} \dfrac{\cos(x)-1}{x} = \lim\limits_{x\to 0} \dfrac{\dfrac{d(\cos(x)-1)}{dx}}{1} = \lim\limits_{x\to 0} \dfrac{-\sin(x)}{1} = 0$

$\Rightarrow \quad \cos(x) - 1 = o(x)$

Differentiable function of single variable

The function $y = f(x,)$ is called differentiable at (x_0) if

$$y - y_0 = f_x(x_0)(x - x_0) + o(x - x_0),$$

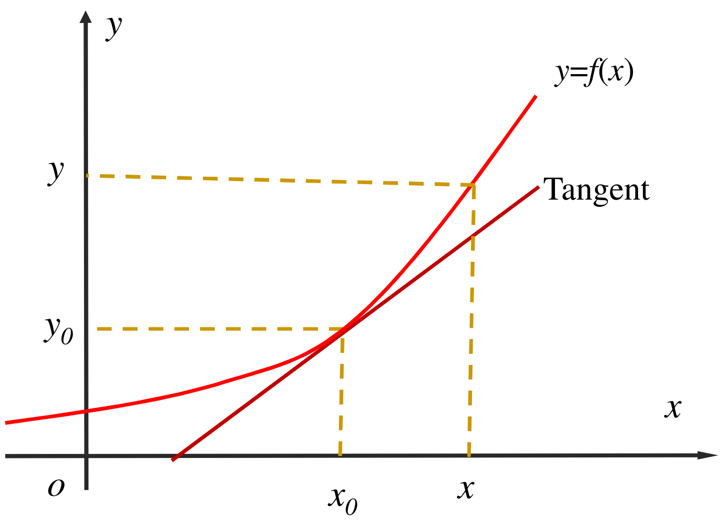
$$\text{or} \quad \Delta y = f_x(x_0)\Delta x + o(\Delta x),$$

where $\Delta x = x - x_0$, $\Delta y = y - y_0$.

we have

$$y - y_0 \approx f_x(x_0)(x - x_0),$$

When a curve is differentiable at x_0, the curve can be approximated by its tangent near x_0.



Differentiable functions with two variables

The function $z = f(x, y)$ is called differentiable at (x_0, y_0) if

$$\Delta z = f_x(x_0, y_0)\Delta x + f_y(x_0, y_0)\Delta y + o(\rho),$$

$$\text{or} \quad z\text{-}z_0 = f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) + o(\rho),$$

where $\Delta x = x\text{-}x_0$, $\Delta y = y\text{-}y_0$, $\Delta z = z\text{-}z_0$, and $\rho = \sqrt{(\Delta x)^2 + (\Delta y)^2}$.

We thus have

$$z\text{-}z_0 \approx f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$$

It is the equation for the tangent plane of the surface $z = f(x, y)$ at (x_0, y_0).

When a surface is differentiable at (x_0, y_0), the surface can be approximated by its tangent plane near (x_0, y_0).

Differentiable function

Definition 3.1. The function $f(x, y, z)$ is called differentiable at (x_0, y_0, z_0) if $\Delta f = f(x, y, z) - f(x_0, y_0, z_0)$ can be expressed as

$$\Delta f = f_x(x_0, y_0, z_0)\Delta x + f_y(x_0, y_0, z_0)\Delta y + f_z(x_0, y_0, z_0)\Delta z + o(\rho),$$

where $\Delta x = x\text{-}x_0$, $\Delta y = y\text{-}y_0$, $\Delta z = z\text{-}z_0$, and $\rho = \sqrt{(\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2}$.

The notation $o(\rho)$ which means going to 0 faster than ρ, i. e.

$$\lim_{\rho \to 0} \frac{o(\rho)}{\rho} = 0.$$

As ρ is infinitively small, we have

$$df = f_x(x_0, y_0, z_0)dx + f_y(x_0, y_0, z_0)dy + f_z(x_0, y_0, z_0)dz.$$

The definition of the differentiable remains the same in the case of n variables as one variable.

Differentiable function

In the case of a function of one variable, differentiable is equivalent to having derivative, i.e.

$$f \in D \quad \Leftrightarrow \quad \exists f_x \,.$$

In the case of a function of several variables, this is not true. In fact, a differential function has all the partial derivatives

$$f \in D(x_1, x_2, \cdots, x_n) \;\Rightarrow\; \exists f_{x_1}, f_{x_2}, \cdots, f_{x_n} \,.$$

But the existence of partial derivatives does not mean the function is differentiable

$$f \in D(x_1, x_2, \cdots, x_n) \quad \not\Leftarrow \quad \exists f_{x_1}, f_{x_2}, \cdots, f_{x_n}$$

Counterexample: the function $f(x, y) = \sqrt[3]{xy}$ has both derivatives at (0, 0), but it is not differentiable at (0, 0).

Sufficiently smoothed function

Informal Definition 3.2. A function is said to be *sufficiently smooth* if the function and **as many of its partial derivatives as required** are continuous where they need to be.

Below, we will assume all our functions to be sufficiently smooth and differentiable.

The chain rule for functions of single variable

Suppose $y=y(x)$ and $x=x(u)$. Denoting $y=y(x(u))=Y(u)$, we have

$$\frac{dY}{du} = \frac{dy}{dx}\frac{dx}{du}.$$

Conventionally we just write

$$\frac{dy}{du} = \frac{dy}{dx}\,\frac{dx}{du}.$$

We will use this convention for multi-variable functions.

The chain rule for functions with two variables

Consider a differentiable function $f(x, y)$. We have $df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy$

Assume $x=x(u, v)$, $y=y(u, v)$ being differentiable,

$$dx = \frac{\partial x}{\partial u}du + \frac{\partial x}{\partial v}dv \qquad dy = \frac{\partial y}{\partial u}du + \frac{\partial y}{\partial v}dv$$

Combining the above 3 equations

$$\begin{aligned} df &= \frac{\partial f}{\partial x}\left(\frac{\partial x}{\partial u}du + \frac{\partial x}{\partial v}dv\right) + \frac{\partial f}{\partial y}\left(\frac{\partial y}{\partial u}du + \frac{\partial y}{\partial v}dv\right) \\ &= \left(\frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u}\right)du + \left(\frac{\partial f}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial v}\right)dv \end{aligned}$$

since $f = f\left(x(u,v), y(u,v)\right) = f(u,v) \quad \Rightarrow \quad df = \frac{\partial f}{\partial u}du + \frac{\partial f}{\partial v}dv$

This gives the *Chain Rule* $\frac{\partial f}{\partial u} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u}, \quad \frac{\partial f}{\partial v} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial v}.$

Examples 1

Assume $z=2xy-y^2$, $x=t^2+1$, $y=t^2-1$. Calculate dz/dt.

Solution 1.

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial t} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}$$

Calculate the derivatives needed and substitute them back into the above equation yields

$$\begin{aligned} \frac{dz}{dt} &= 2y2t + (2x-2y)2t = 4t(y+x-y) = 4tx \\ &= 4t(t^2+1) \quad \text{(express the results in terms of } t\text{)} \end{aligned}$$

Solution 2. One can express z in terms of t

$$z = 2(t^2+1)(t^2-1) - (t^2-1)^2 = 2(t^4-1) - (t^2-1)^2$$

$$\frac{dz}{dt} = 8t^3 - 2(t^2-1)2t = 4t(t^2+1)$$

The Chain rule

In a general case, we have the following chain rule.

Fact 2.5. Let f be a function of the variables $x_1, x_2, \cdots, x_n$,

$$f = f(x_1, \cdots, x_n),$$

where each x_j is a function of (some of) the variables $t_1, t_2, \cdots, t_m$

$$x_j = x_j(t_1, t_2, \cdots, t_m),\ j=1, 2, \cdots, n.$$

If f and x_j are sufficiently smooth, then

$$\boxed{\frac{\partial f}{\partial t_i} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial t_i} + \frac{\partial f}{\partial x_2}\frac{\partial x_2}{\partial t_i} + \cdots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial t_i},\ i = 1, 2, \cdots, \mathrm{m}.}$$

Example 2

Assume $w = e^x yz,\ x = ts,\ y = s^2,\ z = t\text{-}s$. Find w_t.

Solution. Using the chain rule,

$$\frac{\partial w}{\partial t} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial t} \tag{1}$$

Calculate the derivatives needed

$$\frac{\partial w}{\partial x} = e^x yz,\quad \frac{\partial w}{\partial y} = e^x z,\quad \frac{\partial w}{\partial z} = e^x y,$$

$$\frac{\partial x}{\partial t} = s,\quad \frac{\partial y}{\partial t} = 0,\quad \frac{\partial z}{\partial t} = 1.$$

Substituting them into (1) yields

$$\frac{\partial w}{\partial t} = e^x yz \cdot s + e^x z \cdot 0 + e^x y \cdot 1 = e^x y(zs+1)$$

$$= e^{ts} s^2((t-s)s+1) \qquad \text{(express } x, y, z \text{ in terms of } s, t\text{)}$$

Key points of the chain rule

Suppose that f is a function of variables $x, y, \cdots, z$, and that each of $x, y, \cdots, z$ is a function of variables $r, s, \cdots, t$.

To calculate f_s:

(1) First use the Chain Rule to write

$$f_s = f_x x_s + f_y y_s + \cdots + f_z z_s .$$

(2) Using the given functions to work out $f_x, f_y, \cdots, f_z$ and $x_s, y_s, \cdots z_s$ as far as is possible.

(3) Substitute these back into the expression for f_s and simplify, leaving your answer in terms of the variables $r, s, \cdots, t$ whenever possible.

If you do not have an explicit expression for f in terms of $x, y, \cdots, z$, or x in terms of $r, s, \cdots, t$, then your answer will involve terms of the form f_x or x_s, etc.

Example 3

Consider $f(g(x,y),h(x,y))$, where $x = x(t)$ and $y = y(t, s)$.

Find f_t and f_s.

Solution.

$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial g}\frac{\partial g}{\partial t} + \frac{\partial f}{\partial h}\frac{\partial h}{\partial t} = \frac{\partial f}{\partial g}\left(\frac{\partial g}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial g}{\partial y}\frac{\partial y}{\partial t}\right) + \frac{\partial f}{\partial h}\left(\frac{\partial h}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial h}{\partial y}\frac{\partial y}{\partial t}\right)$$

$$= \frac{\partial f}{\partial g}\left(\frac{\partial g}{\partial x}\frac{dx}{dt} + \frac{\partial g}{\partial y}\frac{\partial y}{\partial t}\right) + \frac{\partial f}{\partial h}\left(\frac{\partial h}{\partial x}\frac{dx}{dt} + \frac{\partial h}{\partial y}\frac{\partial y}{\partial t}\right)$$

$$\frac{\partial f}{\partial s} = \frac{\partial f}{\partial g}\frac{\partial g}{\partial s} + \frac{\partial f}{\partial h}\frac{\partial h}{\partial s} = \frac{\partial f}{\partial g}\left(\frac{\partial g}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial g}{\partial y}\frac{\partial y}{\partial s}\right) + \frac{\partial f}{\partial h}\left(\frac{\partial h}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial h}{\partial y}\frac{\partial y}{\partial s}\right)$$

$$= \frac{\partial f}{\partial g}\frac{\partial g}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial f}{\partial h}\frac{\partial h}{\partial y}\frac{\partial y}{\partial s}$$

Example 4

In polar coordinates: $x=r\cos\theta$, $y=r\sin\theta$. Calculate $r_x, r_y, \theta_x, \theta_y$

Solution.

Using $x = r\cos\theta,\quad y = r\sin\theta,$ we have

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

Perform derivatives w.r.t. x to the eq.

$$\frac{\partial r^2}{\partial x} = \frac{\partial}{\partial x}\left(x^2 + y^2\right) \quad\Rightarrow\quad 2rr_x = 2x \;\Rightarrow\; r_x = x/r = \cos\theta$$

Perform derivative w.r.t. y

$$\frac{\partial r^2}{\partial y} = \frac{\partial}{\partial y}\left(x^2 + y^2\right) \quad\Rightarrow\quad 2rr_y = 2y \;\Rightarrow\; r_y = y/r = \sin\theta$$

Example 4 (cont.)

Using $x = r\cos\theta,\quad y = r\sin\theta,$ we have

$$\tan\theta = y/x$$

Perform derivative w.r.t. x to the eq.

$$\frac{\partial \tan\theta}{\partial x} = \frac{\partial}{\partial x}\left(\frac{y}{x}\right) \quad\Rightarrow\quad \sec^2\theta\,\theta_x = -\frac{y}{x^2}$$

$$\Rightarrow\quad \left(\frac{r}{x}\right)^2 \theta_x = -\frac{y}{x^2} \quad\Rightarrow\quad \theta_x = -\frac{y}{r^2} = -\frac{1}{r}\frac{y}{r} = -\frac{\sin\theta}{r}$$

Perform derivative w.r.t. y to the eq.

$$\frac{\partial \tan\theta}{\partial y} = \frac{\partial}{\partial y}\left(\frac{y}{x}\right) \quad\Rightarrow\quad \sec^2\theta\,\theta_y = \frac{1}{x}$$

$$\Rightarrow\quad \left(\frac{r}{x}\right)^2 \theta_y = \frac{1}{x} \quad\Rightarrow\quad \theta_y = \frac{x}{r^2} = \frac{1}{r}\frac{x}{r} = \frac{\cos\theta}{r}$$

Example 5

Convert the partial differential equation to polar coordinates

$$x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} = x^2 + y^2 . \qquad (1)$$

Solution. In polars: $x=r\cos\theta$, $y=r\sin\theta$. Denoting $f(x, y)=F(r, \theta)$, we have

$$\frac{\partial f}{\partial x} = \frac{\partial F}{\partial r}\frac{\partial r}{\partial x} + \frac{\partial F}{\partial \theta}\frac{\partial \theta}{\partial x}, \qquad \frac{\partial f}{\partial y} = \frac{\partial F}{\partial r}\frac{\partial r}{\partial y} + \frac{\partial F}{\partial \theta}\frac{\partial \theta}{\partial y}$$

Using the results of example 4 yields

$$\frac{\partial f}{\partial x} = F_r\cos\theta - F_\theta\frac{\sin\theta}{r}, \qquad \frac{\partial f}{\partial y} = F_r\sin\theta + F_\theta\frac{\cos\theta}{r},$$

Substitution of $x=r\cos\theta$, $y=r\sin\theta$ and the above two eqs. into (1) yields

$$r\cos\theta\left(F_r\cos\theta - F_\theta\frac{\sin\theta}{r}\right) + r\sin\theta\left(F_r\sin\theta + F_\theta\frac{\cos\theta}{r}\right) = (r\cos\theta)^2 + (r\sin\theta)^2$$

$$r\left(\cos^2\theta + \sin^2\theta\right)F_r = r^2\left(\cos^2\theta + \sin^2\theta\right) \;\Rightarrow\; F_r = r$$

Example 5 (cont.)*

Integrate $F_r = r$ with respect to r

$$F(r,\theta) = \frac{1}{2}r^2 + g(\theta),$$

for some arbitrary differentiable function of $g(\theta)$. Hence in terms of x and y

$$f(x,y) = F(r,\theta) = \frac{1}{2}\left(x^2 + y^2\right) + g\left(\tan^{-1}\left(\frac{y}{x}\right)\right)$$

i.e. $$f(x,y) = \frac{1}{2}\left(x^2 + y^2\right) + h\left(\frac{y}{x}\right),$$

where h is an arbitry differentiable function.

1.4 Vectors and vector geometry

A scalar just has magnitude. The examples are pressure, temperature, density, velocity potential, etc.

Definition 4.1 A vector has both a magnitude and a direction. The examples are force, velocity and vorticity of flows, etc.

A vector $\boldsymbol{v}$ is noted as

$\vec{v}$ or $\boldsymbol{v}$ in printing, $\vec{v}$ or $\underline{v}$ in writing

$|\boldsymbol{v}|$ or v denotes the magnitude of $\boldsymbol{v}$. If $\boldsymbol{v} = (v_1, v_2, v_3)$, then

$$v = |\boldsymbol{v}| = \sqrt{v_1^2 + v_2^2 + v_3^2}$$

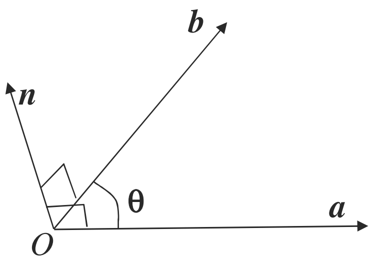
A unit vector is a vector with its magnitude being 1.

In the Cartesian coordinate system $\boldsymbol{i}, \boldsymbol{j}$ and $\boldsymbol{k}$ denote the unit coordinate vectors along the x, y, z-axes respectively.

Dot product and cross-product of two vectors

Definition 4.2. Consider two vectors

$$\boldsymbol{a} = a_1\boldsymbol{i} + a_2\boldsymbol{j} + a_3\boldsymbol{k}, \quad \boldsymbol{b} = b_1\boldsymbol{i} + b_2\boldsymbol{j} + b_3\boldsymbol{k} .$$



The scalar or dot product of the two vectors is

$$\boxed{\boldsymbol{a}\cdot\boldsymbol{b} = |\boldsymbol{a}||\boldsymbol{b}|\cos\theta = a_1b_1 + a_2b_2 + a_3b_3 ,}$$

where θ is the angle between $\boldsymbol{a}$ and $\boldsymbol{b}$.

The vector or cross-product of the two vectors is

$$\boxed{\boldsymbol{v} = \boldsymbol{a}\times\boldsymbol{b} = |\boldsymbol{a}||\boldsymbol{b}|\sin\theta\boldsymbol{n} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = (a_2b_3 - a_3b_2)\boldsymbol{i} + (a_3b_1 - a_1b_3)\boldsymbol{j} + (a_1b_2 - a_2b_1)\boldsymbol{k}}$$

where $\boldsymbol{n}$ is the unit vector normal to both $\boldsymbol{a}$ and $\boldsymbol{b}$ such that ($\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{n}$) forms a right handed system.

Equation of a line

If $\boldsymbol{a} \neq 0$, $\boldsymbol{b} \neq 0$, then

$\boldsymbol{a}$ and $\boldsymbol{b}$ are perpendicular if and only if $\boldsymbol{a} \cdot \boldsymbol{b} = 0$

$\boldsymbol{a}$ and $\boldsymbol{b}$ are parallel if and only if $\boldsymbol{a} \times \boldsymbol{b} = 0$

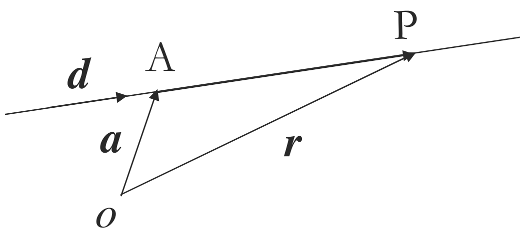
The vector equation of a straight line is

$$\boxed{\boldsymbol{r} = \boldsymbol{a} + t\boldsymbol{d} \quad \text{for } t \in \mathbb{R}}$$

where $\boldsymbol{a}$ is the position vector of a point on the line, $\boldsymbol{d}$ is a vector parallel to the line and $\boldsymbol{r}$ is the position vector of a general point on the line.

$AP = \boldsymbol{r} - \boldsymbol{a}$ and $AP \parallel \boldsymbol{d}$

So that $\boldsymbol{r} - \boldsymbol{a} \parallel \boldsymbol{d} \quad \Rightarrow \quad \boldsymbol{r} - \boldsymbol{a} = t\boldsymbol{d}$



Equation of a plane

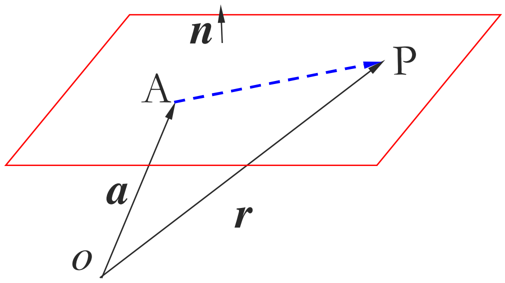
Fact 4.3 The vector equation of a *plane* is

$$\boxed{\boldsymbol{r} \cdot \boldsymbol{n} = \boldsymbol{a} \cdot \boldsymbol{n}},$$

where $\boldsymbol{a}$ is the position vector of a point in the plane, $\boldsymbol{n}$ is a vector normal to the plane and $\boldsymbol{r}$ is the position vector of a general point in the plane.

$AP = \boldsymbol{r} - \boldsymbol{a}$ and $AP \perp \boldsymbol{n}$

So that $\boldsymbol{r} - \boldsymbol{a} \perp \boldsymbol{n} \quad \Rightarrow \quad (\boldsymbol{r} - \boldsymbol{a}) \cdot \boldsymbol{n} = 0$



Fact 3.4. If $\boldsymbol{n}$ is a normal vector to the surface S at the point A with position vector $\boldsymbol{a}$, then the vector equation of the tangent plane to S at A is, again, given by

$$\boldsymbol{r} \cdot \boldsymbol{n} = \boldsymbol{a} \cdot \boldsymbol{n},$$

where $\boldsymbol{r}$ is the position vector of a general point in the plane.

Example 1

Show that a straight line can be presented as an intersection of two planes.

Proof. If we have a straight line

$$\boldsymbol{r}=\boldsymbol{a}+t\boldsymbol{b},$$

then by multiplying this equation by two vectors $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ normal to $\boldsymbol{b}$ we get

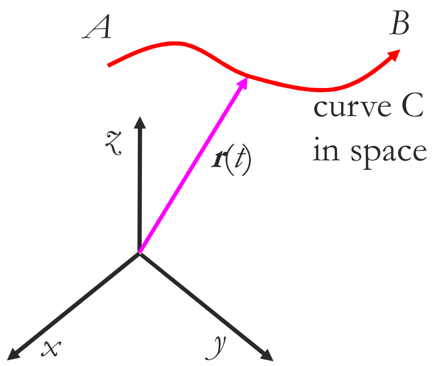
$$\boldsymbol{r}\cdot\boldsymbol{n}_1=\boldsymbol{a}\cdot\boldsymbol{n}_1 \quad \text{and} \quad \boldsymbol{r}\cdot\boldsymbol{n}_2=\boldsymbol{a}\cdot\boldsymbol{n}_2 ,$$

which are equations for two planes.

Representation of curves

A curve C can be expressed using parametric form

$$\boldsymbol{r}=\boldsymbol{r}(t)=x(t)\boldsymbol{i}+y(t)\boldsymbol{j}+z(t)\boldsymbol{k}, \quad a\le t\le b. \qquad (1)$$



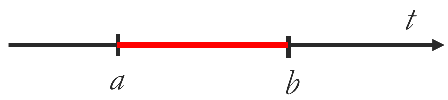
Thus, the initial point A at

$$(x(a), y(a), z(a))$$

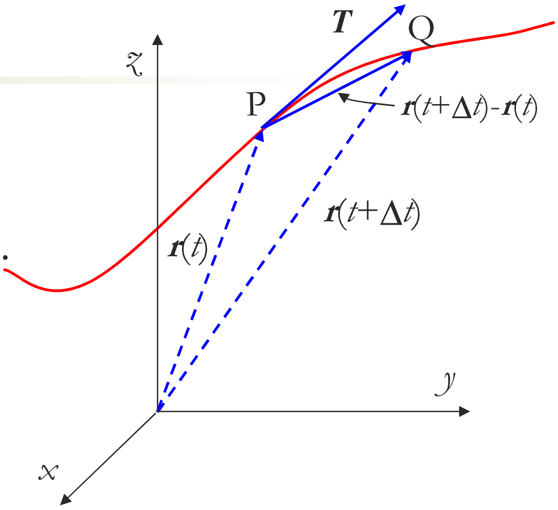
and the terminal point B at

$$(x(b), y(b), z(b))$$

$\boldsymbol{r}(t)$ is a mapping of the interval $[a, b]$ on the t-axis to the curve in xyz-space.



The tangent vector of a curve



The parametric expression of a curve is

$\boldsymbol{r} = x(t)\boldsymbol{i} + y(t)\boldsymbol{j} + z(t)\boldsymbol{k}, \quad a \le t \le b$.

To each t, there a point $\boldsymbol{r}(t)$ on the curve. The tangent vector $\boldsymbol{T}$ of the curve can be found by

$$\boldsymbol{T} = \lim_{\Delta t \to 0} \frac{\boldsymbol{r}(t+\Delta t) - \boldsymbol{r}(t)}{\Delta t}$$

$$= \lim_{\Delta t \to 0} \frac{x(t+\Delta t) - x(t)}{\Delta t}\boldsymbol{i} + \lim_{\Delta t \to 0} \frac{y(t+\Delta t) - y(t)}{\Delta t}\boldsymbol{j} + \lim_{\Delta t \to 0} \frac{z(t+\Delta t) - z(t)}{\Delta t}\boldsymbol{k}$$

$$= \frac{dx}{dt}\boldsymbol{i} + \frac{dy}{dt}\boldsymbol{j} + \frac{dz}{dt}\boldsymbol{k} = \frac{d\boldsymbol{r}}{dt} \qquad \boxed{\boldsymbol{T} = \frac{d\boldsymbol{r}}{dt}}$$

In general, for a sufficiently smooth curve one can always have parameterization for which $\boldsymbol{r}' \neq 0$

1.5 Gradient vector and directional derivative

Definition 5.1 Let function $f : \mathbb{R}^3 \to \mathbb{R}$ be a function of three variables x, y, z. If the partial derivatives of f with respect to x, y and z exist at the point (a, b, c) then the gradient vector of f at (a, b, c) is defined to be the vector

$$\boxed{\nabla f(a,b,c) = \mathrm{grad} f(a,b,c) \equiv \frac{\partial f(a,b,c)}{\partial x}\boldsymbol{i} + \frac{\partial f(a,b,c)}{\partial y}\boldsymbol{j} + \frac{\partial f(a,b,c)}{\partial z}\boldsymbol{k}}$$

∇f is read as del f or nabla f.

Fact 5.2 The value of ∇f at a point (a, b, c) is independent of the co-ordinate system (and basis vectors), i.e. ∇f is well defined.

This actually means that

$$\nabla f = \frac{\partial f}{\partial x}\boldsymbol{i} + \frac{\partial f}{\partial y}\boldsymbol{j} + \frac{\partial f}{\partial z}\boldsymbol{k} = \frac{\partial f}{\partial x_1}\boldsymbol{i}_1 + \frac{\partial f}{\partial y_1}\boldsymbol{j}_1 + \frac{\partial f}{\partial z_1}\boldsymbol{k}_1$$

One can prove this using the chain rule and the definition of basic vectors for arbitrary two Cartesian coordinates.

Key points

The gradient vector ∇f of a n-variable function is a *vector function* in n dimensions.

For example if $f : \mathbb{R}^2 \to \mathbb{R}$, then $\nabla f = f_x \boldsymbol{i} + f_y \boldsymbol{j}$.

We can think of differentiation w.r.t. x as an operator d/dx, that acts on functions of x, so that

$$\frac{d}{dx} f = \frac{df}{dx}$$

It is often useful to consider ∇ as an operator, i.e. 'something that you do to functions',

$$\nabla f = \left(\boldsymbol{i} \frac{\partial}{\partial x} + \boldsymbol{j} \frac{\partial}{\partial y} + \boldsymbol{k} \frac{\partial}{\partial z} \right) f = \frac{\partial f}{\partial x} \boldsymbol{i} + \frac{\partial f}{\partial y} \boldsymbol{j} + \frac{\partial f}{\partial z} \boldsymbol{k}$$

Like usual differentiation, ∇ is a linear operator, i.e. for any functions f and g and any constants λ and μ,

$$\nabla (\lambda f + \mu g) = \lambda \nabla f + \mu \nabla g \,.$$

Grad as a normal

Theorem 5.3 Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a sufficiently smooth function and let (a, b) be a point on the level curve $f(x, y) = k$ for some constant k. If $\nabla f(a, b) \neq 0$, then the vector $\nabla f(a, b)$ is normal to the level curve $f(x, y) = k$ at the point (a, b).

Let $f : \mathbb{R}^3 \to \mathbb{R}$ be a sufficiently smooth function and let (a, b, c) be a point on the level surface $f(x, y, z) = k$ for some constant k.

If $\nabla f(a, b, c) \neq 0$, then the vector $\nabla f(a, b, c)$ is normal to the level surface $f(x, y, z) = k$ at the point (a, b, c).

Outline of Proof:

Let the level curve $f(x, y) = k$ have the vector equation

$$\boldsymbol{r}(t) = x(t)\boldsymbol{i} + y(t)\boldsymbol{j} \ ,$$

then $f(x(t), y(t)) = k$ for all t.

Thus: $$\frac{d}{dt} f(x(t), y(t)) = \frac{dk}{dt} = 0$$

By the Chain Rule, we have

$$x' f_x + y' f_y = \nabla f \cdot \boldsymbol{r}' = 0 \quad \Rightarrow \quad \nabla f \perp \boldsymbol{r}'$$

Assume here the parameter t is chosen suitably so that $\boldsymbol{r}' \neq 0$.

The proof for the 3D case is almost identical: the same argument shows that ∇f is normal to any curve in the level surface passing through the point (a, b, c) and hence normal to the surface.

Finding the tangent plane to a surface

To calculate the tangent plane to a surface

$$z = F(x, y) \quad \text{or} \quad f(x, y, z) = k$$

at the point $\boldsymbol{a} = (a, b, c)$:

(1) Write the surface in the form of the level surface

$$f(x, y, z) = F(x, y) - z = 0$$

or $$f(x, y, z) = k \ ,$$

(2) Calculate the normal vector to the surface at the point $\boldsymbol{a}$

$$\boldsymbol{n} = \nabla f(\boldsymbol{a})$$

(3) The tangent plane is given by

$$\boldsymbol{r} \cdot \boldsymbol{n} = \boldsymbol{a} \cdot \boldsymbol{n}$$

Example 1

Find the tangent plane for the cone $z^2=4(x^2+y^2)$ at point P: (1, 0, 2).

Solution. Express the cone in the format of $f(x, y, z)=0$:

$$f=4(x^2+y^2)-z^2=0 .$$

Calculate the gradient

$$\text{grad}\, f = 8x\boldsymbol{i}+8y\boldsymbol{j}-2z\boldsymbol{k}$$

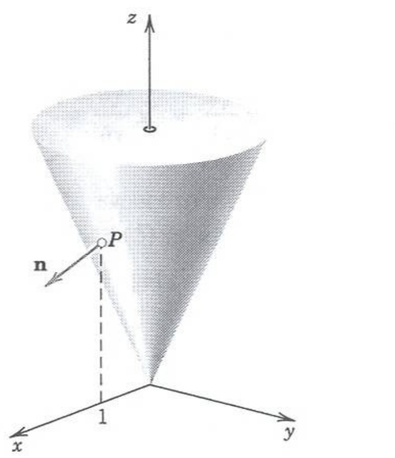
and at P, $\quad \boldsymbol{n}=\text{grad}\, f=8\boldsymbol{i}-4\boldsymbol{k}$.

The tangent plane

$$\boldsymbol{r}\cdot\boldsymbol{n} = \boldsymbol{a}\cdot\boldsymbol{n}$$

$\Rightarrow$ $(x, y, z)\cdot(8, 0, -4) = (1, 0, 2)\cdot(8, 0, -4)$

i.e. $8x-4z=0 \Rightarrow 2x-z=0$



Cone and unit normal vector **n**

Example 2

Find the equation of the tangent plane to the surface

$$z = x^2 + y^2 + x\sin y$$

at the point (1, 0, 1) . (Exam paper 2011)

Solution. Write the surface in the level surface format

$$f = x^2 + y^2 + x\sin y - z = 0$$

Its normal can be calculated as

$$\boldsymbol{n} = \nabla f = (2x+\sin y)\boldsymbol{i} + (2y + x\cos y)\boldsymbol{j} - \boldsymbol{k}$$

at the point (1, 0, 1) $\qquad \boldsymbol{n} = 2\boldsymbol{i} + \boldsymbol{j} - \boldsymbol{k}$

The equation for the tangent plane is then given

$$\boldsymbol{r}\cdot\boldsymbol{n} = (1,0,1)\cdot\boldsymbol{n}$$

i.e. $(x, y, z)\cdot(2,\ 1,\ -1) = (1,0,1)\cdot(2,\ 1,\ -1)$

or $\quad 2x + y - z = 1$

Example 3

Find a vector tangent to the curve of intersection of the two surfaces

$z=x^2-y^2$ and $xyz=-30$

at the point (-3, 2, 5).

Solution. The tangent to this curve at the point will be perpendicular to the normals $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$ of the two surfaces.

$$\boldsymbol{n}_1=\nabla\left(x^2-y^2-z\right)\Big|_{(-3,2,5)}=\left(2x\boldsymbol{i}-2y\boldsymbol{j}-\boldsymbol{k}\right)\Big|_{(-3,2,5)}=-6\boldsymbol{i}-4\boldsymbol{j}-\boldsymbol{k}$$

$$\boldsymbol{n}_2=\nabla\left(xyz+30\right)\Big|_{(-3,2,5)}=\left(yz\boldsymbol{i}+xz\boldsymbol{j}+xy\boldsymbol{k}\right)\Big|_{(-3,2,5)}=10\boldsymbol{i}-15\boldsymbol{j}-6\boldsymbol{k}$$

The tangent $\boldsymbol{T}$ is given as

$$\boldsymbol{T}=\boldsymbol{n}_1\times\boldsymbol{n}_2=\begin{vmatrix}\boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k}\\ -6 & -4 & -1\\ 10 & -15 & -6\end{vmatrix}=9\boldsymbol{i}-46\boldsymbol{j}+130\boldsymbol{k}$$

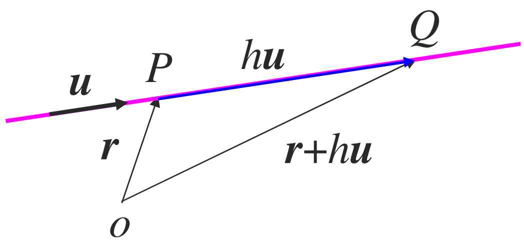
Directional derivatives

Consider the function $f(x, y, z)$ at point $P(x, y, z)$.

Draw a line PQ along the direction of $\boldsymbol{u}=u_1\boldsymbol{i}+u_2\boldsymbol{j}+u_3\boldsymbol{k}$ which is a unit vector .

Denote the distance of PQ as h

$$|PQ|=h$$



One can see from the figure

Point P is at $\boldsymbol{r}=(x, y, z)$

Point Q is at $\boldsymbol{r}+h\boldsymbol{u}=(x+hu_1, y+hu_2, z+hu_3)$

The directional derivative of f at (x, y, z) in the direction of $\boldsymbol{u}$ is defined

$$D_{\boldsymbol{u}}f(x,y,z)=\lim_{Q\to P}\frac{f(Q)-f(P)}{|PQ|}$$

$$=\lim_{h\to 0^+}\frac{f(x+hu_1, y+hu_2, z+hu_3)-f(x,y,z)}{h}$$

Directional derivatives

Definition 5.4 Suppose that $f: \mathbb{R}^3 \to \mathbb{R}$ and $\boldsymbol{u} = u_1\boldsymbol{i} + u_2\boldsymbol{j} + u_3\boldsymbol{k}$ is a unit vector in $\mathbb{R}^3$. The directional derivative of f at (x, y, z) in the direction of $\boldsymbol{u}$ is defined to be

$$D_{\boldsymbol{u}} f(x, y, z) = \lim_{h\to 0^+} \frac{f(x+hu_1, y+hu_2, z+hu_3) - f(x, y, z)}{h}$$

Example 4: what does D_{-i} equal to

$$D_{-i} f(x, y, z)$$

$$= \lim_{h\to 0^+} \frac{f(x-h, y, z) - f(x, y, z)}{h} \qquad (\text{since } \boldsymbol{u} = -\boldsymbol{i})$$

$$= -\lim_{\delta\to 0} \frac{f(x+\delta, y, z) - f(x, y, z)}{\delta} \qquad (\delta = -h)$$

$$= -f_x(x,y,z)$$

Directional derivatives (cont.)

Lemma 5.5 Let f : $\mathbb{R}^3 \to \mathbb{R}$ be sufficiently smooth and $\boldsymbol{u} = u_1\boldsymbol{i} + u_2\boldsymbol{j} + u_3\boldsymbol{k}$ be a unit vector in $\mathbb{R}^3$. If it exists, then

$$\frac{d}{dt} f(a+tu_1, b+tu_2, c+tu_3)\Big|_{t=0} = D_u f(a,b,c)$$

Proof Denote $F(t) = f(a+tu_1, b+tu_2, c+tu_3)$

$$\text{LHS} = \frac{dF}{dt}\Big|_{t=0} = F'(0)$$

$$= \lim_{h\to 0} \frac{F(h) - F(0)}{h}$$

$$= \lim_{h\to 0} \frac{1}{h}\left[f(a+hu_1, b+hu_2, c+hu_3) - f(a,b,c)\right]$$

$$= D_u f(a,b,c)$$

Relation between gradient and directional derivative

Theorem 5.6 Let $f: \mathbb{R}^3 \to \mathbb{R}$ be sufficiently smooth, $\boldsymbol{u} = u_1\boldsymbol{i} + u_2\boldsymbol{j} + u_3\boldsymbol{k}$ be a unit vector in $\mathbb{R}^3$. Then

$$\boxed{D_u f(a,b,c) = \boldsymbol{u}\cdot\nabla f(a,b,c)}$$

Proof.

$$D_u f(a,b,c) = \frac{d}{dt} f(a+tu_1, b+tu_2, c+tu_3)\Big|_{t=0} \qquad \text{(Using Lemma 4.5)}$$

$$= \left\{ \begin{array}{l} \frac{\partial}{\partial x} f(a+tu_1, b+tu_2, c+tu_3)\frac{d}{dt}(a+tu_1) + \\ \frac{\partial}{\partial y} f(a+tu_1, b+tu_2, c+tu_3)\frac{d}{dt}(b+tu_2) + \\ \frac{\partial}{\partial z} f(a+tu_1, b+tu_2, c+tu_3)\frac{d}{dt}(c+tu_3) \end{array} \right\}_{t=0} \qquad \text{Using the chain rule}$$

$$= u_1\frac{\partial}{\partial x} f(a,b,c) + u_2\frac{\partial}{\partial y} f(a,b,c) + u_3\frac{\partial}{\partial x} f(a,b,c) = \boldsymbol{u}\cdot\nabla f(a,b,c)$$

Key Points

The directional derivative is a *scalar function* not a vector.

D_u is linear, i.e. for any functions f and g and any scalar constants λ and μ,

$$D_u(\lambda f + \mu g) = \lambda D_u f + \mu D_u g\ .$$

To calculate the directional derivative of a function f in the direction of the vector $\boldsymbol{U}$:

(1) Calculate the unit vector $\boldsymbol{u} = \boldsymbol{U}/|\boldsymbol{U}|$;

(2) Calculate ∇f ;

(3) Calculate $D_u f = \nabla f \cdot \boldsymbol{u}$

To calculate the second directional derivative of the function f in the direction of $\boldsymbol{u}$:

First calculate $g = D_u f$. As g is a scalar function, we can the calculate the directional derivative of g.

Example 5

Let $f(x, y, z) = xy + yz + zx$

Find the directional derivative of f at the point (1, 1, 1) in the direction to the point (2, 2, 3). (Exam paper 2011)

Solution: The gradient of f is given by

$$\nabla f = (y+z)\boldsymbol{i} + (z+x)\boldsymbol{j} + (x+y)\boldsymbol{k}$$

at the point (1, 1, 1) $\quad \nabla f(1,1,1) = 2\boldsymbol{i} + 2\boldsymbol{j} + 2\boldsymbol{k}$

The unit vector from the point (1, 1, 1) to the point (2, 2, 3) is

$$\boldsymbol{u} = \frac{\boldsymbol{i}+\boldsymbol{j}+2\boldsymbol{k}}{\sqrt{1^2+1^2+2^2}} = \frac{\boldsymbol{i}+\boldsymbol{j}+2\boldsymbol{k}}{\sqrt{6}}$$

Finally we have the directional derivative

$$D_u f = \boldsymbol{u}\cdot\nabla f(1,1,1) = \frac{\boldsymbol{i}+\boldsymbol{j}+2\boldsymbol{k}}{\sqrt{6}}\cdot(2\boldsymbol{i}+2\boldsymbol{j}+2\boldsymbol{k}) = \frac{2+2+4}{\sqrt{6}} = \frac{8}{\sqrt{6}}$$

Steepest ascent theorem

Theorem 5.7 Let f be a function of several real variables. Suppose that ∇f exists and is non-zero at the point $\boldsymbol{a}$.

Then the direction of the vector $\nabla f(\boldsymbol{a})$ is always the direction of maximum increase of the function f at the point $\boldsymbol{a}$.

Proof: Let $\boldsymbol{u}$ be any unit vector. By Theorem 4.6,

$$D_u f(\boldsymbol{a}) = \boldsymbol{u}\cdot\nabla f(\boldsymbol{a}) = |\boldsymbol{u}||\nabla f(\boldsymbol{a})|\cos\gamma = |\nabla f(\boldsymbol{a})|\cos\gamma$$

where γ is the angle between $\boldsymbol{u}$ and ∇f.

For a given function f and given point $\boldsymbol{a}$, $|\nabla f(\boldsymbol{a})|$ is fixed.

$D_u f(\boldsymbol{a})$ is maximal when $\gamma = 0$, i.e. when $\boldsymbol{u}$ and $\nabla f(\boldsymbol{a})$ are in the same direction.

Hence the maximum directional derivative is in the direction of the gradient or, in other words, the gradient is the direction of the steepest ascent of the function.

Example 6

Let $\varphi(x, y, z)=xyz$ and the point P at (2, 1, 1).

(a) In what direction from the point P does φ have its maximum rate of change? What is this rate of change?

(b) In what direction from the point P does φ have its minimum rate of change? What is this rate of change?

Solution. $\text{grad}\varphi = yz\boldsymbol{i} + zx\boldsymbol{j} + xy\boldsymbol{k}$

at P, grad $\varphi(2, 1, 1)=\boldsymbol{i}+2\boldsymbol{j}+2\boldsymbol{k}$.

(a) Along direction $\boldsymbol{u}$, the directional derivative is

$$D_{\boldsymbol{u}}\varphi=|\nabla\varphi|\cos\gamma = \sqrt{1+2^2+2^2}\cos\gamma = 3\cos\gamma$$

When $\gamma=0$, $D_{\boldsymbol{u}}\varphi$ has its greatest rate of change along the direction grad $\varphi(2, 1, 1)= \boldsymbol{i}+2\boldsymbol{j}+2\boldsymbol{k}$. This maximum rate is 3.

(b) When $\gamma=\pi$, $D_{\boldsymbol{u}}\varphi$ has its least rate of change along the direction $-\boldsymbol{i}-2\boldsymbol{j}-2\boldsymbol{k}$. This minimum rate is -3.

Example 7

The height of a hill is given by $f(x, y) = xe^{-y}$. Assuming that a wolf always follows the path of steepest ascent, find the path the wolf follows if it starts at (2, 1).

Solution. Let the path be $\boldsymbol{r}(t) = x(t)\boldsymbol{i} + y(t)\boldsymbol{j}$. The tangent vector to the path, $\boldsymbol{r}'(t) = x'(t)\boldsymbol{i}+y'(t)\boldsymbol{j}$, is in the direction of the maximum increase of the function f at (x, y).

∇f at a point (x, y) is always in the direction of the maximum increase of the function f at (x, y)

Thus $\boldsymbol{r}'(t) = x'(t)\boldsymbol{i}+y'(t)\boldsymbol{j}$ is parallel to $\nabla f = f_x\boldsymbol{i} + f_y\boldsymbol{j}$.

$$x'(t)\boldsymbol{i}+y'(t)\boldsymbol{j} = \lambda\,(f_x\,\boldsymbol{i} + f_y\boldsymbol{j}\,),\ \ \lambda\neq 0 \ \Rightarrow\ \frac{y'(t)}{x'(t)} = \frac{f_y}{f_x}$$

Hence $$\frac{y'(t)}{x'(t)} = \frac{f_y}{f_x}$$

Example 7 (cont.)

As $f(x, y) = xe^{-y}$, we have

$$\frac{dy}{dx} = \frac{y'(t)}{x'(t)} = \frac{s_y}{s_x} = \frac{-xe^{-y}}{e^{-y}} = -x \quad \Rightarrow \quad \frac{dy}{dx} = -x$$

As $$\frac{dy}{dx} = \frac{y'(t)}{x'(t)} \quad \Rightarrow \quad \frac{dy}{dx} = -x$$

So that $y = -x^2/2 + c$.

But the path starts at (2, 1), so when $x = 2$, $y = 1$, i.e. the path of the wolf is $y = -x^2/2 + 3$.