SOMEBODY TRY TO EXPLAIN

MULTIPLE CALCULUS

written by Yang: GuanZhong

CATALOGUE

- Derivative

2) Introduction to Partial Derivative
3) Differential of Binary Function
4) Slope in Different Directions
5) Tangent Plane of Curve
6) Extremum with and without Restriction

- Calculus

1) Introduction to Double Integral
2) Polar Form for Double Integral
3) Surface Area of Binary Function
4) Jacobians Matrix and Jacobians Substitution
5) Triple Integral, Cylindrical Form, Spherical Form

Introduction to Partial Derivative

To begin with, we shall consider the 3-dimensional x-y-z system, in which the value of z is determined by x and y.

After sketching the graph, we want to find out the tiny change in z when there is a tiny change in variables.

However, notice that, compared with unary functions, binary function has two variables, both of which can change, meaning that a sentence like 'a tiny change in variables' will lead to confusion: what is the direction of this tiny change?

In this way, in order to better explain 'the tiny change in z when tiny change in variables', we establish two basic directions of change: parallel to x-axis and parallel to y-axis.

This is reasonable since it is easy for us to use these two directions to get any other directions.

Parallel to x-axis, the tiny change in variables means the tiny change in x and no change in y, so we use $\frac{\partial z}{\partial x}$ to express it.

Similarly, parallel to y-axis, the tiny change in variables means the tiny change in y and no change in x, so we use $\frac{\partial z}{\partial y}$ to express it.

Here, $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ are known as 'partial derivative' since only one of several variables is taken into consideration when finding the tiny change in result.

So now, we know what do $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ mean, but how should we calculate them?

Recall that, when finding partial derivative, only one of several variables is taken into consideration and others remain unchanged, then it is acceptable for us to regard other variables as constants since their job is to express the coefficients of several terms.

If z=f(x,y)=$\sum f(x)g(y)$,

$$\frac{\partial z}{\partial x} = \lim_{\Delta x \to 0} \frac{f(x+\Delta x, y) - f(x,y)}{\Delta x} = \sum f'(x)g(y),$$

$$\frac{\partial z}{\partial y} = \lim_{\Delta y \to 0} \frac{f(x, y+\Delta y) - f(x,y)}{\Delta y} = \sum f(x)g'(y),$$

and basic rules like product rule, quotient rule etc. still holds.

Similar conclusions also hold for functions with more variables.

Just like unitary function, binary function also has secondary derivative, but there is a special conclusion.

$$\frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right) = \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial y}\right).$$

Proof.

Given z=f(x,y),

since

$$\frac{\partial}{\partial y}\left(\frac{\partial f(x,y)}{\partial x}\right)=\lim_{\Delta x \to 0,\ \Delta y \to 0}\frac{\frac{f(x+\Delta x,y+\Delta y)\ -\ f(x,y+\Delta y)}{\Delta x}-\frac{f(x+\Delta x,y)\ -\ f(x,y)}{\Delta x}}{\Delta y}$$

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

$$\frac{\partial}{\partial x}\left(\frac{\partial f(x,y)}{\partial y}\right)=\lim_{\Delta x \to 0,\ \Delta y \to 0}\frac{\frac{f(x+\Delta x,y+\Delta y)\ -\ f(x+\Delta x,y)}{\Delta y}-\frac{f(x,y+\Delta y)\ -\ f(x,y)}{\Delta y}}{\Delta x}$$

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

so

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

Similar conclusions also hold for functions with more variables.

Differential of Binary Function

In order to understand differential, we should firstly understand Rolle's theorem and Lagrange mean value theorem, which are very common in the proof of differential.

Rolle's Theorem

If f(x) be continuous on the closed interval [a,b] and differentiable on the open interval (a,b), f(a)=f(b), then there is at least one number c $\in$ (a,b) such that $f'(c)=0$.

Proof.

We can definitely find a maximum value or minimum value in the interval (a,b). Here, assume we can find a maximum value.

If there exist one number $c\in(a,b)$ such that

$$f(c)\geq f(x),$$

so

$$f(c)\geq f(c+\Delta x).$$

When $\Delta x>0$,

$$\frac{f(c+\Delta x)-f(c)}{\Delta x}\leq 0,$$

When $\Delta x<0$,

$$\frac{f(c+\Delta x)-f(c)}{\Delta x}\geq 0.$$

Thus

$$f'(c)=f'_{+}(c)=\lim_{\Delta x\to 0}\frac{f(c+\Delta x)-f(c)}{\Delta x}\leq 0,$$

$$f'(c)=f'_{-}(c)=\lim_{\Delta x\to 0}\frac{f(c+\Delta x)-f(c)}{\Delta x}\geq 0,$$

so

$$f'(c)=0$$

Lagrange Mean Value Theorem

If f(x) is continuous on the closed interval [a,b] and differentiable on the open interval (a,b), then there exist at least one number c $\in$ (a,b) such that $f'(c)=\frac{f(b)-f(a)}{b-a}$.

Proof.

Line passing through (a,f(a)), (b,f(b)) has expression,

$$y=\frac{f(b)-f(a)}{b-a}(x-a)+f(a).$$

Let g(x) represents the difference between this line and f(x),

$$g(x)=f(x)-\frac{f(b)-f(a)}{b-a}(x-a)-f(a).$$

Since

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

according to Rolle's theorem, there exists at least one number $c\in(a,b)$ such that

$$g'(c)=0,$$

$$f'(c)-\frac{f(b)-f(a)}{b-a}=0,$$

$$f'(c)=\frac{f(b)-f(a)}{b-a}.$$

After knowing these, we can get the expression of the differential of z.

For a binary function z=f(x,y),

$$\mathbf{dz}=\frac{\partial z}{\partial x}\mathbf{dx}+\frac{\partial z}{\partial y}\mathbf{dy}.$$

Proof.

Given z=f(x,y).

$$dz=\lim_{\Delta x\to 0,\ \Delta y\to 0} f(x+\Delta x\ ,y+\Delta y)-f(x,y)$$
$$=\lim_{\Delta x\to 0,\ \Delta y\to 0} f(x+\Delta x\ ,y+\Delta y-f(x+\Delta x,y)+f(x+\Delta x,y)-f(x,y).$$

According to Lagrange Mean Value Theorem,

$$dz=\lim_{\Delta x\to 0,\ \Delta y\to 0}\frac{\partial}{\partial x}[f(x+\theta_1\Delta x\ ,y+\Delta y)]\Delta x+\frac{\partial}{\partial y}[f(x\ ,y+\theta_2\Delta y)]\Delta y$$

where $0<\theta_1,\theta_2<1$.

Since $\frac{\partial z}{\partial x},\frac{\partial z}{\partial y}$ exist and are continuous in the interval, we can get

$$\frac{\partial}{\partial x}[f(x+\theta_1\Delta x\ ,y+\Delta y)]=\frac{\partial z}{\partial x}+\alpha,$$

$$\frac{\partial}{\partial y}[f(x\ ,y+\theta_2\Delta y)]=\frac{\partial z}{\partial y}+\beta,$$

where α and β are higher order infinitesimals. You may consider them as infinitesimals which are even significant smaller than $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$.

So

$$dz=\lim_{\Delta x\to 0,\ \Delta y\to 0}\left(\frac{\partial f(x,y)}{\partial x}+\alpha\right)\Delta x+\left(\frac{\partial f(x,y)}{\partial y}+\beta\right)\Delta y$$

$$=\lim_{\Delta x\to 0,\ \Delta y\to 0}\frac{\partial f(x,y)}{\partial x}\Delta x+\frac{\partial f(x,y)}{\partial y}\Delta y+\alpha\Delta x+\beta\Delta y$$

$$=\frac{\partial f(x,y)}{\partial x}dx+\frac{\partial f(x,y)}{\partial y}dy$$

since α and β are too small, we can ignore them.

Similar conclusions also hold for functions with more variables.

According to the conclusion we get before, we can also find $\frac{dz}{dx}$ or $\frac{dz}{dy}$ by dividing dx or dy on the both sides.

By applying this knowledge, we can find an easy way to find implicit function.

Given f(x,y)=0, assume that f(x,y)=z, and if $\frac{\partial z}{\partial x},\frac{\partial z}{\partial y}$ exist and are continuous in the interval, we can get

$$\frac{dy}{dx}=-\frac{\frac{\partial z}{\partial x}}{\frac{\partial z}{\partial y}}$$

if $\frac{\partial z}{\partial y}\neq 0$.

Proof.

Given f(x,y)=0 and assume z=f(x,y),

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

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

Since $\frac{dz}{dx}=0$, $\frac{dx}{dx}=1$,

$$\frac{\partial z}{\partial x}+\frac{\partial z}{\partial y}\frac{dy}{dx}=0,$$

$$\frac{dy}{dx}=-\frac{\frac{\partial z}{\partial x}}{\frac{\partial z}{\partial y}}$$

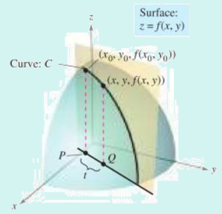
if $\frac{\partial z}{\partial y} \neq 0$.

Slope in Different Directions

Given z=f(x,y), we now can get the slope of the curves parallel to either x-axis or y-aixs by doing partial derivative.

But this is not enough --- there are many other curves which are not parallel to either x-axis or y-aixs and what if we want to find the slope of these curves?

If the curve passes through (x_0,y_0,z_0) and the direction of the plane, which consisting of the curve, is u=icos θ +jsin θ , then the slope of the curve, D_uf(x,y), is



$$D_u f(x,y)=\frac{\partial z}{\partial x}\cos\theta+\frac{\partial z}{\partial y}\sin\theta$$

Proof.

In order to find the slope in any curves, we should firstly express the curve we want.

If the curve passes through (x_0,y_0,z_0) and the direction of the plane, which consisting of the curve, is u=icos θ +jsin θ , then

$$x=x_0+t\cos\theta,\ y=y_0+t\sin\theta.$$

Recall

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

so

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

$$D_u f(x,y)=\frac{\partial z}{\partial x}\cos\theta+\frac{\partial z}{\partial y}\sin\theta$$

by applying the equality we get.

You may notice that

$$D_u f(x,y)=\begin{pmatrix}\frac{\partial z}{\partial x}\\ \frac{\partial z}{\partial y}\end{pmatrix}\bullet\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix},$$

and the latter vector is just the direction vector of the plane.

You may notice that the first vector is very unique --- it tells us the partial derivative of f(x,y) and by knowing it and the direction of the plane which consisting of the curve, we can get the slope of the curve.

In fact, the first vector is a special vector that we called ‘gradient’.

Gradient of f(x,y), known as ∇f(x,y), is

$$\nabla f(x,y)=\begin{pmatrix}\frac{\partial z}{\partial x}\\ \frac{\partial z}{\partial y}\end{pmatrix}.$$

The direction of gradient always faces to the direction of extremum slope.

Proof.

$$D_u f(x,y)=\begin{pmatrix}\frac{\partial z}{\partial x}\\ \frac{\partial z}{\partial y}\end{pmatrix} \bullet \begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}=\left|\begin{pmatrix}\frac{\partial z}{\partial x}\\ \frac{\partial z}{\partial y}\end{pmatrix}\right|\left|\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}\right|\cos\alpha = \left|\begin{pmatrix}\frac{\partial z}{\partial x}\\ \frac{\partial z}{\partial y}\end{pmatrix}\right|\cos\alpha$$

where α is the angle between the direction of gradient and the direction of curve, whose slope is going to be found.

Notice that, $D_u f(x,y)$ is maximum when $\alpha=0$, and minimum when $\alpha=\pi$. In other words, $D_u f(x,y)$ is maximum when it has the same direction with the gradient, and it is minimum when it has the opposite direction with the gradient.

In reality, sometimes we are asked to find a path, which always changes at the greatest rate. That be said, we are asked to find a path, which always faces to the direction of gradient no matter where we are.

When finding a path which continuously moves in a direction of maximum increase, equality

$$\frac{dx}{\frac{\partial z}{\partial x}}=\frac{dy}{\frac{\partial z}{\partial y}}$$

is satisfied.

Proof.

Assume that path is

$$\mathbf{r}(t) = \begin{pmatrix}x(t)\\ y(t)\end{pmatrix}.$$

Since it continuously moves in a direction of maximum increase, then the direction of the rate of change of path is always the same as the direction of gradient at that point.

Recall that, if two vectors are in the same directions, then one can be expressed as the scalar times of another.

Thus,

$$\mathbf{r}'(t)=\begin{pmatrix}\frac{dx}{dt}\\ \frac{dy}{dt}\end{pmatrix}=k\begin{pmatrix}\frac{\partial z}{\partial x}\\ \frac{\partial z}{\partial y}\end{pmatrix},$$

$$\frac{dx}{\frac{\partial z}{\partial x}}=\frac{dy}{\frac{\partial z}{\partial y}}.$$

Tangent Plane of Curve

What if we want to find the tangent plane of the curve?

Let S be a surface given by z=f(x,y), and we can expressed it as F(x,y,z)=0.

Assume that P (x_0,y_0,z_0) be the point on S, then the tangent plane passing through P has normal vector $\nabla f(x_0, y_0, z_0)$.

So the expression of tangent plane is

$$\nabla f(x_0, y_0, z_0) \bullet (x_0-x, y_0-y, z_0-z)=0.$$

Proof.

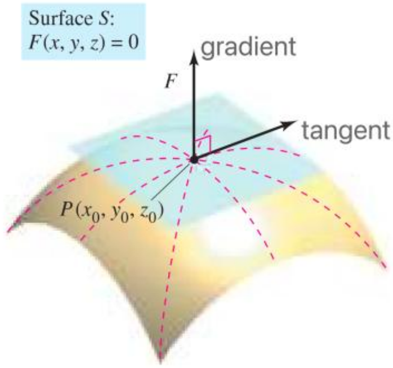
There are two possible explanations.

(1) Prove by Calculation

Let S be a surface given by z=f(x,y), and we can expressed it as F(x,y,z)=0.

Assume that P (x_0,y_0,z_0) be the point on S, and C be a curve on S passing through P with expression

$$\mathbf{r}(t)=\begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}.$$



So

$$F(x(t),y(t),z(t))=0,$$

$$\frac{\partial F}{\partial x}\frac{dx}{dt}+\frac{\partial F}{\partial y}\frac{dy}{dt}+\frac{\partial F}{\partial z}\frac{dz}{dt}=0.$$

Hence for P,

$$\nabla f(x_0, y_0, z_0) \bullet \mathbf{r}'(t_0)=0$$

Here, we may find that $\mathbf{r}'(t_0)$ represents the direction of curve.

Notice that the result of dot product is 0 and neither $\mathbf{r}'(t_0)$ nor $\nabla f(x_0, y_0, z_0)$ are zero, so the gradient at P is orthogonal to the tangent vector.

This statement holds for every curve on S through P, so there is only one direction to which the gradient can face, and that is the direction of the normal vector of the tangent plane.

Thus, the tangent plane passing through P has normal vector $\nabla f(x_0, y_0, z_0)$.

(2) Prove by Using Contour Surface

We may notice that, S is a contour surface since

$$F(x(t),y(t),z(t))=0.$$

Recall that the direction of gradient always faces to the direction of extremum change, so the gradient must be in the direction of the normal vector of the tangent plane, or otherwise the gradient will have a zero component.

Extremum with and without Restriction

Just like unitary function, sometimes we are asked to find out the extremum of the binary function. Just like finding the extremum of unitary function, the first thing we do to find out extremum of a binary function is to let the first partial derivatives be either 0 or not defined in order to narrow the scope.

Critical point is a point which satisfies one of these two conditions:

① $\frac{\partial z}{\partial x}=0$ **and** $\frac{\partial z}{\partial y}=0$. ② $\frac{\partial z}{\partial x}$ **or** $\frac{\partial z}{\partial y}$ **not exist.**

Relative maximum and minimum point only occurs at critical point, but critical point may not be relative maximum or minimum point, it can be saddle point(neither maximum nor minimum).

Just like finding the extremum of unitary function, after finding out those critical points, the next thing we need to do is to use second partial derivative to judge the feature of the points we found.

Let z=f(x,y) have continuous second partial derivatives on an open region containing a point (a,b) for which

$$\frac{\partial z}{\partial x}=0 \text{ and } \frac{\partial z}{\partial y}=0.$$

To test for the relative extremum of f(x,y), consider the quantity

$$d=\frac{\partial^2 z}{\partial x^2}\frac{\partial^2 z}{\partial y^2}-[\frac{\partial^2 z}{\partial x \partial y}]^2.$$

(1) If d>0 and $\frac{\partial^2 z}{\partial x^2}>0$, then f(x,y) has a relative minimum at (a,b).

(2) If d>0 and $\frac{\partial^2 z}{\partial x^2}<0$, then f(x,y) has a relative maximum at (a,b).

(3) If d<0, then f(x,y) has a saddle point at (a,b).

(4) The test is inconclusive if d=0.

Proof.

In order to judge the feature of the point we found, we should find out the quadratic approximation at that point, since in this way we can easily judge whether or not it is maximum or minimum:

(1) If the quadratic approximation opens upwards, then understandably the point is a minimum point.

(2) If the quadratic approximation opens downwards, then understandably the point is a maximum point.

(3) If the quadratic approximation opens neither upwards nor downwards, then understandably the point is a saddle point.

So we need to firstly derive the expression of quadratic approximation at a point, say, (x_0, y_0).

Notice that, our requirement, or expectation, of a quadratic approximation Q(x,y), is:

$$Q(x_0, y_0)=z_0, \frac{\partial Q(x_0,y_0)}{\partial x}=\frac{\partial z}{\partial x}, \frac{\partial Q(x_0,y_0)}{\partial y}=\frac{\partial z}{\partial y}, \frac{\partial^2 Q(x_0,y_0)}{\partial x^2}=\frac{\partial^2 z}{\partial x^2}, \frac{\partial^2 Q(x_0,y_0)}{\partial y^2}=\frac{\partial^2 z}{\partial y^2}, \frac{\partial^2 Q(x_0,y_0)}{\partial x\partial y}=\frac{\partial^2 z}{\partial x\partial y}.$$

In this way, we can easily find out the expression of quadratic approximation:

$$Q(x,y)=z_0+\frac{\partial z}{\partial x}(x-x_0)+\frac{\partial z}{\partial y}(y-y_0)+\frac{1}{2}\frac{\partial^2 z}{\partial x^2}(x-x_0)^2+\frac{\partial^2 z}{\partial x\partial y}(x-x_0)(y-y_0)+\frac{1}{2}\frac{\partial^2 z}{\partial y^2}(y-y_0)^2.$$

Assume that $X=x-x_0$, $Y=y-y_0$, recall that $\frac{\partial z}{\partial x}=\frac{\partial z}{\partial y}=0$, re-arrange,

$$Q(x,y)=z_0+\frac{1}{2}\frac{\partial^2 z}{\partial x^2}X^2+\frac{\partial^2 z}{\partial x\partial y}XY+\frac{1}{2}\frac{\partial^2 z}{\partial y^2}Y^2.$$

In order to judge the opening direction of the surface, we only focus on the last three terms. What we want to do is to prove that the sum of these three terms is always positive or negative. If so, recall the fact that critical point is the critical point of the surface and $\frac{\partial z}{\partial x}=\frac{\partial z}{\partial y}=0$, then the opening direction is unique and we can judge it:

(1)If the sum of these three terms is always positive, then the surface opens upwards and critical point is the minimum point of the quadratic approximation, so it is the minimum point of the original surface.

(2)If the sum of these three terms is always negative, then the surface opens downwards and critical point is the maximum point of the quadratic approximation, so it is the maximum point of the original surface.

In order to show that Q(x,y) is always positive or negative, notice that it is a quadratic equation of x, so we just need to focus on its determination and the coefficient of highest term.

We need to prove $\Delta<0$,

$$\left(\frac{\partial^2 z}{\partial x\partial y}\right)^2 - \frac{\partial^2 z}{\partial x^2}\frac{\partial^2 z}{\partial y^2} < 0.$$

If so, critical point is either maximum or minimum, and if $\frac{\partial^2 z}{\partial x^2} > 0$, according to graph, critical point is minimum, similar for maximum.

And similarly, when

$$\left(\frac{\partial^2 z}{\partial x\partial y}\right)^2 - \frac{\partial^2 z}{\partial x^2}\frac{\partial^2 z}{\partial y^2} > 0,$$

critical point is a saddle point.

One application of the extremum of binary function is the least squares regression line, which is commonly used in statistics to model a group of data that has a linear relationship.

The least squares regression line for $\{(x_1,y_1),\ldots,(x_n,y_n)\}$ is given by f(x)=ax+b, where

$$a=\frac{n\sum_{i=1}^{n}x_iy_i-\sum_{i=1}^{n}x_i\sum_{i=1}^{n}y_i}{n\sum_{i=1}^{n}x_i^2-(\sum_{i=1}^{n}x_i)^2},$$

$$b=\frac{1}{n}\left(\sum_{i=1}^{n}y_i-a\sum_{i=1}^{n}x_i\right).$$

Proof.

Assume that we have a group of data {(x_1,y_1),...,(x_n,y_n)}, and there is a linear relationship between x and y, and we want to model this relationship by using f(x)=ax+b.

In order to better achieve this, the difference between the f(x_i) and y_i should be minimised.

In order to avoid the influence of the sign of the difference, we use square error, S,

$$S=\sum_{i=1}^{n}[f(x_i)-y_i]^2=\sum_{i=1}^{n}[ax_i+b-y_i]^2.$$

By minimising S, we are finding the minimum of S.

Notice that, here, a and b are variables,

$$\frac{\partial S}{\partial a}=\sum_{i=1}^{n}2x_i[ax_i+b-y_i]=2a\sum_{i=1}^{n}x_i^2+2b\sum_{i=1}^{n}x_i-2\sum_{i=1}^{n}x_iy_i,$$

$$\frac{\partial S}{\partial b}=\sum_{i=1}^{n}2x_i[ax_i+b-y_i]=2a\sum_{i=1}^{n}x_i+2nb-2\sum_{i=1}^{n}y_i.$$

By letting $\frac{\partial S}{\partial a}=\frac{\partial S}{\partial b}=0$, we can get the result.

In the aforementioned situation, there is no restriction when we find the extremum. However, sometimes, we have to find the extremum of z=f(x,y) with the restriction g(x,y)=c, where c is a constant.

Lagrange's Theorem

Let f(x,y) and g(x,y) have continuous first partial derivatives such that f(x,y) has an extremum at a point (x_0,y_0) on the smooth constraint curve g(x,y)=c, where c is a constant. If $\nabla g(x_0,y_0)\neq 0$, then there is a real number λ such that

$$\nabla f(x_0,y_0)=\lambda\nabla g(x_0,y_0).$$

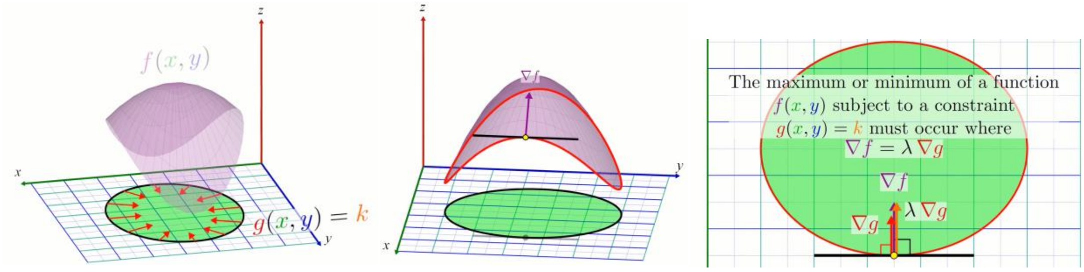
Proof.

Recall that the direction of gradient always faces to the direction of extremum slope, we then consider g(x,y)=c as a level curve of g(x,y), so the $\nabla g(x,y)$ is always perpendicular to constrain curve, or otherwise the gradient will have a zero component.

If (x_0,y_0) is maxima, then tangent line is parallel to x-y plane, meaning that the slope in the direction of this tangent line is 0. In this way, $\nabla f(x_0,y_0)$ will face to the same direction as $\nabla g(x_0,y_0)$ since otherwise, $\nabla f(x_0,y_0)$ will have a zero component.

Recall that, if two vectors are in the same directions, then one can be expressed as the scalar times of another, so there is a real number λ such that

$$\nabla f(x_0,y_0)=\lambda\nabla g(x_0,y_0).$$



That be said, we can find a method of finding extremum with restriction.

Let f(x,y) and g(x,y) satisfy the hypothesis of Lagrange's Theorem, and let f(x,y) have a minimum or maximum subject to the constraint g(x,y)=c, where c is a constant. To find the minimum or maximum of f(x,y), use the following steps.

(1) Solve the following system of equations

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

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

$$g(x,y)=c.$$

(2) Evaluate f(x,y) at each solution point obtained in the first step. The largest value yields the maximum and the smallest value yields the minimum.

Introduction to Double Integral

Just like what we do when doing derivative, when doing double integral, we only take one variable into consideration while consider others as constants.

$$\int_{h_1(y)}^{h_2(y)} \frac{\partial f(x,y)}{\partial x}\, dx = f(x,y)\Big]_{h_1(y)}^{h_2(y)} = f(h_2(y),y) - f(h_1(y),y),$$

$$\int_{h_1(x)}^{h_2(x)} \frac{\partial f(x,y)}{\partial y}\, dy = f(x,y)\Big]_{h_1(x)}^{h_2(x)} = f(x,h_2(x)) - f(x,h_1(x)).$$

But notice that, integration variable cannot also be the limit.

Sometimes, we need to consider one variable as integration variable, and then regard another variable as integration variable, which is known as 'iterated integration', like

$$\int_a^b \int_{h_1(y)}^{h_2(y)} f_x(x,y)\ dx\ dy \text{ or } \int_c^d \int_{h_1(x)}^{h_2(x)} f_y(x,y)\ dy\ \ dx.$$

The inside limits of integration can be variable with the respect to outer variable of integration. However, the outer limits of integration must be constant with the respect to both variables of integration.

Given z=f(x,y).

If R is defined by x∈[a,b], y∈[g_1(x),g_2(x)], where g_1(x) and g_2(x) are continuous on [a,b], then the area of R is given by

$$A=\int_a^b \int_{g_1(x)}^{g_2(x)} f(x,y)\ dydx.$$

If R is defined by y∈[c,d], y∈[h_1(y),h_2(y)], where h_1(y) and h_2(y) are continuous on [c,d], then the area of R is given by

$$A=\int_c^d \int_{h_1(y)}^{h_2(y)} f(x,y)\ dxdy.$$

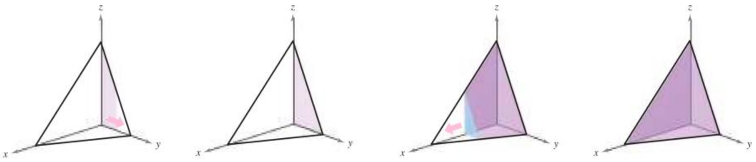
In order to better understand what these integrals mean, it may be helpful for us to consider the aforementioned two integral as different ways to find out the volume under the surface.

$$A=\int_a^b \int_{g_1(x)}^{g_2(x)} f(x,y)\ dydx$$

means that firstly, for each value of x_0, the area under the surface and on the plane of x=x_0 can be expressed by x_0, thus, we add all these areas together to get the volume.
Similarly,

$$A=\int_c^d \int_{h_1(y)}^{h_2(y)} f(x,y)\ dxdy$$

means that firstly, for each value of y_0, the area under the surface and on the plane of y=y_0 can be expressed by y_0, thus, we add all these areas together to get the volume.



Integrate with respect to y to obtain the area of the cross section. Integrate with respect to x to obtain the volume of the solid.
Figure 14.17

This kind of idea is very important for the learning of double integrals.

Actually, as we can see, ‘the tiny change in x’ multiplying ‘the tiny change in y’ will result in ‘the tiny change in area’.

$$\lim_{\|\Delta x\|\to 0,\|\Delta y\|\to 0} \Delta x_i \Delta y_i = \lim_{\|\Delta x\|\to 0,\|\Delta y\|\to 0} \Delta A_i$$

where $\|\Delta x\|$ be the largest absolute value of change in x and $\|\Delta y\|$ be the largest largest absolute value of change in y.

So

$$dxdy=dA.$$

Notice that, the underlying idea of substitution is that the values of two expression are the same, and we replace one by another to simplify the calculation.

By using this strategy, we may re-write the conclusion we got just now.

Fubini’s Theorem

If R is defined by $x\in[a,b]$, $y\in[g_1(x),g_2(x)]$, where $g_1(x)$ and $g_2(x)$ are continuous on [a,b], then the area of R is given by

$$A=\int_a^b \int_{g_1(x)}^{g_2(x)} f(x,y)\ dydx=\int_R \int f(x,y)\ dA.$$

If R is defined by $y\in[c,d]$, $y\in[h_1(y),h_2(y)]$, where $h_1(y)$ and $h_2(y)$ are continuous on [c,d], then the area of R is given by

$$A=\int_c^d \int_{h_1(y)}^{h_2(y)} f(x,y)\ dxdy=\int_R \int f(x,y)\ dA.$$

Actually, this is a very useful conclusion, and in Physics we use it to calculate the center of mass of an object and its polar moment of inertia.

Let $\rho(x,y)$ be a continuous density function on the planar lamina R. Then the moments of mass with respect to x- and y-axes are

$$M_x=\int_R \int y\rho(x,y)\ dA,\ M_y=\int_R \int x\rho(x,y)\ dA.$$

If m is the mass of the lamina, then the center of mass is

$$(\bar{x},\bar{y})=\left(\frac{M_y}{m},\frac{M_x}{m}\right).$$

If R represents a simple plane region rather than a lamina, then point $(\bar{x},\bar{y})$ is called the centroid of the region.

Proof.

We should firstly know what M_x and M_y are.

$$M_x=\iint_R \underbrace{\boxed{y}}_{\substack{\text{The distance of}\\ \text{each slice}\\ \text{to the x-axis.}}}\underbrace{\boxed{\rho(x,y)}}_{\substack{\text{Mass of each slice.}\\ \text{Since uniform mass,}\\ \text{use area to represent}\\ \text{mass.}}}\ dA$$

Similarly for M_y. So M_x and M_y are really summations of ‘distance’ multiplying ‘weight’ of each slice.

So how should we understand $\bar{x}$ and $\bar{y}$? Recall that there are two ways for us to calculate the moments of mass with respect to x- and y-axes: one is to find out the summations of ‘distance’ multiplying ‘weight’ of each slice, another is to directly multiple ‘total weight, centered at center of mass’ with ‘the distance from center of mass to axes’.

Thus, by applying this, we can get what we want.

Let ρ(x,y) be a continuous density function on the planar lamina R. Then the inertia with respect to x- and y-axes are

$$I_x=\int_R \int y^2\rho(x,y)\ dA,\ I_y=\int_R \int x^2\rho(x,y)\ dA.$$

The sum of the moments I_x and I_y is called the polar moment of inertia and is denoted by I_0.

Proof.

We should firstly know what I_x and I_y are.

$$I_x=\iint_R \underbrace{\boxed{y^2}}_{\text{Square of distance to x-axis for each slice}}\underbrace{\boxed{\rho(x,y)}}_{\text{Mass for each slice}}\ dA$$

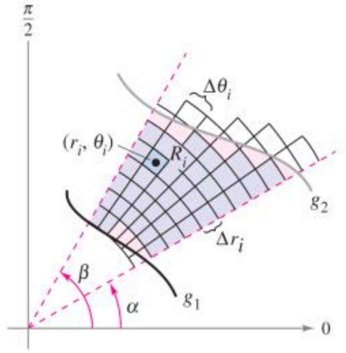
Similarly for I_y. So I_x and I_y are really summations of ‘distance^2’ multiplying ‘weight’ of each slice.

Polar Form for Double Integral

By using the graph on the right hand side, we can derive

$$\lim_{\|\Delta\|_1 \to 0} \Delta A_i = \lim_{\|\Delta\|_2 \to 0} r\Delta r_i \Delta\theta_i$$

where $\|\Delta\|_1$ be the largest diagonal of rectangles and $\|\Delta\|_2$ be the largest diagonal of polar sectors.



Polar grid superimposed over region R

Thus

$$dA = r\, dr d\theta,$$

so

$$dxdy = r\, dr d\theta.$$

So by using this equality and tap the strategy we mentioned before, we can do a substitution and express integration by using polar form.

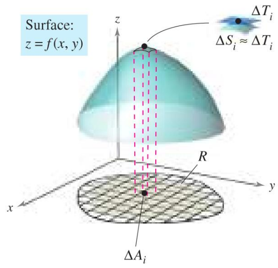
Let R be a plane region consisting of all points (x,y)=(r$\cos\theta$,r$\sin\theta$) satisfying the conditions 0≤ $g_1(\theta)$≤r≤ $g_2(\theta)$,α ≤r≤ β, where 0≤ β-α ≤2π. If $g_1(\theta)$ and $g_2(\theta)$ are continuous on [α,β] and f(x,y) is continuous on R, then

$$\int_R \int f(x,y)\ dA = \int_\alpha^\beta \int_{g_1(\theta)}^{g_2(\theta)} f(r\cos\theta, r\sin\theta) r\, \mathbf{drd\theta}.$$

Surface Area of Binary Function

Now we consider the process of calculate the area of the surface.

In order to do this, for z=f(x,y), at (x_0,y_0,z_0), we consider a small rectangle, centered at (x_0,y_0,z_0) and lying on the tangent plane of (x_0,y_0,z_0). Notice that, when the length of the largest diagonal of this rectangle is small enough, the area of this small rectangle, centered at (x_0,y_0,z_0) and lying on the tangent plane of (x_0,y_0,z_0), is approximately equal to the area of the surface centered at (x_0,y_0,z_0) and lying under the small rectangle.



That be said,

$$\lim_{\|\Delta\|\to 0} \Delta S_i = \lim_{\|\Delta\|\to 0} \Delta T_i$$

where ΔT_i represents the area of a small rectangle centered at (x_i,y_i,z_i) and lying on the tangent plane of (x_i,y_i,z_i), $\|\Delta\|$ be the largest diagonal of rectangles.

In order to find out the area of small rectangle, we can use cross product.

Notice that its two sides of rectangle can be expressed as

$$\mathbf{a} = \lim_{\|\Delta\|\to 0}(\Delta x_i \mathbf{i} + \frac{\partial f(x_i,y_i)}{\partial x}\Delta x_i \mathbf{k}) \text{ , } \mathbf{b} = \lim_{\|\Delta\|\to 0}(\Delta y_i \mathbf{j} + \frac{\partial f(x_i,y_i)}{\partial y}\Delta y_i \mathbf{k}).$$

Recall that

$$\lim_{\|\Delta x\|\to 0, \|\Delta y\|\to 0} \Delta x_i \Delta y_i = \lim_{\|\Delta x\|\to 0, \|\Delta y\|\to 0} \Delta A_i$$

where $\|\Delta x\|$ be the largest absolute value of change in x and $\|\Delta y\|$ be the largest largest absolute value of change in y.

So area is

$$|a \times b| = \lim_{\|\Delta\|\to 0}([\frac{\partial f(x_i,y_i)}{\partial x}]^2 + [\frac{\partial f(x_i,y_i)}{\partial y}]^2 + 1)^{\frac{1}{2}}\Delta A_i.$$

Thus

$$\lim_{\|\Delta\|\to 0} \Delta S_i = \lim_{\|\Delta\|\to 0}([\frac{\partial f(x_i,y_i)}{\partial x}]^2 + [\frac{\partial f(x_i,y_i)}{\partial y}]^2 + 1)^{\frac{1}{2}}\Delta A_i.$$

If f(x,y) and its first partial derivatives are continuous on the closed region R in the xy-plane, then the area of the surface S given by z=f(x,y) over R is defined as

$$\textbf{Area=}\int_R \int dS \textbf{=} \int_R \int ([\frac{\partial z}{\partial x}]^2 + [\frac{\partial z}{\partial y}]^2 + 1)^{\frac{1}{2}} dA.$$

Jacobians Matrix and Jacobians Substitution

Well, sometimes, we may want to use substitution with any arbitrary variables to simplify calculation. To replace original integration variables with any variables we want, we should use Jacobians substitution.

In order to understand what is Jacobians substitution, we should firstly understand Jacobians matrix.

We firstly consider a plane.

For a non-linear transformation, how should we express the corresponding matrix?

Well, since we know how to express a linear transformation by using matrix, the question can be solved easily if we can convert a non-linear transformation into a linear transformation.

In order to do this, we need to learn the concept of 'local linear', which means that, if we magnify a small rectangle centered at a point on a plane by a sufficient multiple, then the nonlinear transformation centered at that point can be approximated as a linear transformation centered at that point.

Express in a mathematical way:

If

$$\mathbf{M}\binom{x}{y} = \binom{u}{v}$$

where **M** is a non-linear transformation,

$$\mathbf{M}\binom{\partial x}{0} = \binom{\partial u_x}{\partial v_x},\ \mathbf{M}\binom{0}{\partial y} = \binom{\partial u_y}{\partial v_y}.$$

Notice that $\binom{\partial u_x}{\partial v_x}$ is not necessarily equal to $\binom{\partial u_y}{\partial v)_y}$, but there are all 'the tiny change in u or v' or '∂u or ∂v'.

We want to know the value of basic vectors of this matrix, so

$$\frac{1}{\partial x}\mathbf{M}\binom{\partial x}{0} = \frac{1}{\partial x}\binom{\partial u_x}{\partial v_x},\ \mathbf{M}\binom{1}{0} = \begin{pmatrix}\frac{\partial u}{\partial x}\\ \frac{\partial v}{\partial x}\end{pmatrix},$$

$$\frac{1}{\partial y}\mathbf{M}\binom{0}{\partial y} = \frac{1}{\partial y}\binom{\partial u_y}{\partial v_y},\ \mathbf{M}\binom{0}{1} = \begin{pmatrix}\frac{\partial u}{\partial y}\\ \frac{\partial v}{\partial y}\end{pmatrix},$$

$$\mathbf{M} = \begin{pmatrix}\frac{\partial u}{\partial x} & \frac{\partial u}{\partial y}\\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y}\end{pmatrix},$$

$$|\det\mathbf{M}| = \left|\frac{\partial u}{\partial x}\frac{\partial v}{\partial y} - \frac{\partial u}{\partial y}\frac{\partial v}{\partial x}\right|.$$

Jacobians Matrix

For a non-linear transformation, M, if $\mathbf{M}\binom{x}{y} = \binom{u}{v}$, then

$$\mathbf{M} = \begin{pmatrix}\frac{\partial u}{\partial x} & \frac{\partial u}{\partial y}\\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y}\end{pmatrix}.$$

Now we can tap Jacobians matrix into substitution of integration.

We firstly learn what the determination of Jacobians matrix tells us: if we firstly find out the area of a shape in the original xy-plane and the area of the transformed shape in the plane with new basic vectors, then the ratio of the area of transformed shape to the area of original shape is just the value of determination.

So

$$\lim_{\|\Delta\|_1 \to 0} \Delta A_i = \lim_{\|\Delta\|_2 \to 0} \Delta T_i \, (\det \mathbf{M}_i)^{-1}$$

where ΔA_i represents the area of small rectangle centered at (x_i,y_i); ΔT_i represents the area of small quadrilateral centered at (x_i,y_i), transformed by the small rectangle; $\|\Delta\|_1$ represents the length of largest diagnose of all the rectangles; $\|\Delta\|_2$ represents the length of largest diagnose of all the quadrilaterals; **M** represents the Jacobians matrix at (x_i,y_i).

Notice that, the underlying idea of substitution is that the values of two expression are the same, and we replace one by another to simplify the calculation.

Jacobians Substitution

Let R be a vertically or horizontally simple region in the xy-plane, and let S be a vertically or horizontally simple region in the uv-plane. Let T from S to R be given by T(u,v)=(x,y)=(g(u,v),h(u,v)), where g(u,v) and h(u,v) have continuous first partial derivatives. Assume that T is one-to-one except possibly on the boundary of S. If f(x,y) is continuous on R, and $\left|\frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial y}{\partial u}\frac{\partial x}{\partial v}\right|$ **is non-zero on S, then**

$$\int_R \int f(x,y)\, dA = \int_S \int f(\mathbf{g(u,v)}, \mathbf{h(u,v)}) \left|\frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial y}{\partial u}\frac{\partial x}{\partial v}\right| dudv.$$

In fact, you may notice that,

$$\frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial y}{\partial u}\frac{\partial x}{\partial v},$$

so

$$\int_R \int f(x,y)\, dA = \int_S \int f(g(u,v), h(u,v)) \left|\frac{\partial(x,y)}{\partial(u,v)}\right| dudv.$$

Triple Integral, Cylindrical Form, Spherical Form

Similar with double integral, when doing triple integral, we only take one variable into consideration while consider others as constants.

If f(x,y,z) is continuous over a bounded solid region Q, then the triple integral of f(x,y,z) over Q is

$$\int\int_Q\int f(x,y,z)\ dxdydz = \int\int_Q\int f(x,y,z)\ dV.$$

We can use multiple integral to find out the center of mass and inertia of a 3D object.

Let ρ(x,y,z) be a continuous density function on the solid region Q. Then the moments of mass with respect to yz-, xz- and xy- planes are

$$M_{yz} = \int\int_Q\int xf(x,y,z)\ dV,\ M_{xz} = \int\int_Q\int yf(x,y,z)\ dV,\ M_{xy} = \int\int_Q\int zf(x,y,z)\ dV.$$

If m is the mass of the object, then the center of mass is

$$(\bar{x},\bar{y},\bar{z}) = \left(\frac{M_{yz}}{m},\frac{M_{xz}}{m},\frac{M_{xy}}{m}\right).$$

Let ρ(x,y,z) be a continuous density function on the solid region Q. Then the inertia with respect to x-, y- and z-axes, plus yz-, xz- and xy- planes are

$$I_x = \int\int_Q\int (y^2+z^2)f(x,y,z)\ dV,$$

$$I_y = \int\int_Q\int (x^2+z^2)f(x,y,z)\ dV,$$

$$I_z = \int\int_Q\int (x^2+y^2)f(x,y,z)\ dV,$$

$$I_{yz} = \int\int_Q\int x^2 f(x,y,z)\ dV,$$

$$I_{xz} = \int\int_Q\int y^2 f(x,y,z)\ dV,$$

$$I_{xy} = \int\int_Q\int z^2 f(x,y,z)\ dV.$$

Similarly, polar form also exists for triple integral.

Let Q be a solid region whose projection R onto the xy-plane can be describe in polar coordinates,

$$Q = \{(x,y,z):\ (x,y) \text{ is in } R,\ h_1(x,y) \le z \le h_2(x,y)\}$$

and

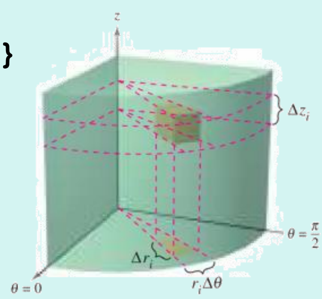
$$R = \{(r,\theta):\ \theta_1 \le \theta \le \theta_2,\ g_1(\theta) \le r \le g_2(\theta)\}.$$

If f(x,y,z) is a continuous function on the solid Q, then

$$\int\int_Q\int f(x,y,z)\ dV$$

$$= \int_R\int \left[\int_{h_1(x,y)}^{h_2(x,y)} f(x,y,z)\ dz\right] dA$$

$$= \int_{\theta_1}^{\theta_2}\int_{g_1(\theta)}^{g_2(\theta)}\int_{h_1(r\cos\theta, r\sin\theta)}^{h_2(r\cos\theta, r\sin\theta)} f(r\cos\theta, r\sin\theta, z)\ dzdrd\theta.$$



Volume of cylindrical block: $\Delta V_i = r_i \Delta r_i \Delta\theta_i \Delta z_i$

Apart from this, there is another form of integration for triple integral, namely spherical form.

In the spherical coordinate system, each point is represented by an ordered triple: the first coordinate ρ is the distance from origin and $\rho \geq 0$; the second coordinate θ is the same angle used in cylindrical coordinates for $r \geq 0$; the third coordinate φ is the angle between the positive z-axis and the line segment $\vec{OP}$, $0 \leq \varphi \leq \pi$.

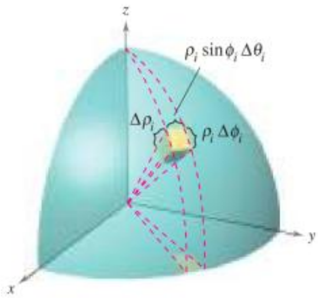
In this way,

$$x=\rho\sin\varphi\cos\theta,$$
$$y=\rho\sin\varphi\sin\theta,$$
$$z=\rho\cos\varphi.$$

By using the graph on the right-hand side, we may notice that:

$$\lim_{\|\Delta\|_1 \to 0} \Delta V_i = \lim_{\|\Delta\|_2 \to 0} \rho_i^2 \sin\varphi \Delta p_i \Delta\varphi_i \Delta\theta_i$$

where $\|\Delta\|_1$ be the largest diagonal of boxes and $\|\Delta\|_2$ be the largest diagonal of polar sectors.



Spherical block:
$\Delta V_i \approx \rho_i^2 \sin\phi_i \Delta\rho_i \Delta\phi_i \Delta\theta_i$

Thus, for continuous function f(x,y,z) defined on the solid region Q, if it can be expressed in spherical coordinates with

$$\{(\rho, \theta, \varphi): \rho_1 \leq \rho \leq \rho_2, \theta_1 \leq \theta \leq \theta_2, \varphi_1 \leq \varphi \leq \varphi_2\}$$

where $\rho_1 > 0$, $\theta_2 - \theta_1 \leq 2\pi$, $0 \leq \varphi_1 \leq \varphi_2 \leq \pi$, we can get

$$\int\int_Q \int f(x,y,z)\, dV = \int_{\theta_1}^{\theta_2} \int_{\varphi_1}^{\varphi_2} \int_{\rho_1}^{\rho_2} f(\rho, \theta, \varphi)\, \rho^2 \sin\varphi \, d\rho d\varphi d\theta.$$