Chapter 7. Functions of Several Variables

7.1 Introduction

In elementary calculus, we encountered scalar functions of one variable, e.g. $f = f(x)$. However, many physical quantities in engineering and science are described in terms of scalar functions of several variables. For example,

(i) Mass density of a lamina can be described by $\delta(x, y)$, where (x, y) are the coordinates of a point on the lamina.

(ii) Pressure in the atmosphere can be described by $P(x, y, z)$ where (x, y, z) are the coordinates of a point in the atmosphere.

(iii) Temperature distribution of a heated metal ball can be described by $T(x, y, z, t)$ where (x, y, z) are the coordinates of a point in the ball and t is the time.

7.1.1 Functions of Two Variables

A function f of **two variables** is a rule that assigns to each ordered pair of real numbers (x, y) a real number denoted by $f(x, y)$.

We usually write $z = f(x, y)$ to indicate that z is a function of x and y. Moreover, x, y are called the independent variables and z is called the dependent variable. The set of all ordered pairs (x, y) such that $f(x, y)$ can be defined is called the **domain** of f.

7.1.2 Example

(a) $f(x,y) = x^2y^3$.

This is a function of two variables which is defined for any x and y. So the domain of f is the set of all (x,y) with $x, y \in \mathbf{R}$.

(b) $f(x,y) = \sqrt{1 - x^2 - y^2}$.

This function is only defined when $1-x^2-y^2 \geq 0$, or equivalently $x^2 + y^2 \leq 1$.

So the domain of f is the set

$$D = \{(x,y) : x^2 + y^2 \leq 1\}.$$

Note that D represents all the points in the xy plane lying within (and on) the unit circle.

(c) We can also define f in "pieces" as a **compound function**. For example

$$f(x,y) = \begin{cases} \sqrt{x-y} & \text{if } x > y, \\ \sqrt{y-x} & \text{if } x < y, \\ 1 & \text{if } x = y\,. \end{cases}$$

7.1.3 Functions of Three or More Variables

We can define functions of **three** variables $f(x, y, z)$, **four** variables $f(x, y, z, w)$, etc in a similar way.

7.2 Geometric Representation

7.2.1 Graphs of functions of two variables

The graph of a function $f(x)$ of one variable is a curve in the xy-plane, which can be regarded as the set of all points (x, y) in the xy-plane such that $y = f(x)$. By analogy, we have the **graph** of a function $f(x, y)$

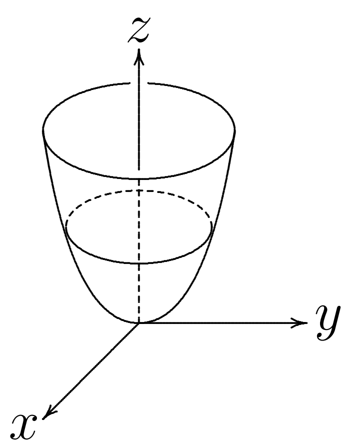
of two variables is the set of all points (x, y, z) in the three dimensional xyz-space such that $z = f(x, y)$. This set represents a surface in the xyz-space.

7.2.2 Example

The graph of $f(x, y) = 5 - 3x - 2y$ is the plane with equation $z = 5 - 3x - 2y$ (or $3x + 2y + z = 5$).

7.2.3 Example

The graph of $g(x, y) = 8x^2 + 2y^2$ is the paraboloid (see diagram below) with equation $z = 8x^2 + 2y^2$.



7.3 Partial Derivatives

Let $f = f(x, y)$ be a function of two variables. Either a change of x or a change of y can cause a change of f. In order to measure the rate of change of f with respect to the variable x, we need to fix the variable y, and vice versa. Note that when we fix one of the variables of $f(x, y)$, then it becomes a function of one variable.

7.3.1 Example

Let $f(x, y) = x^2 - 2xy + 3y^3$. If we fix $y = 2$, say, then

$$f(x, 2) = x^2 - 4x + 24$$

is a function in x alone.

Similarly, if we fix $x = -1$, then

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

is a function in y alone.

7.3.2 First order partial derivatives

Let $f(x, y)$ be a function of two variables. Then the (first order) **partial derivative of f with respect to x** at the point (a, b) is

$$\left.\frac{d}{dx} f(x, b)\right|_{x=a} = \lim_{h \to 0} \frac{f(a+h, b) - f(a, b)}{h}$$

We say that the partial derivative does not exist at (a, b) if the limit on the RHS above does not exist.

When the above partial derivative exists, we denote it by $\left.\dfrac{\partial f}{\partial x}\right|_{(a,b)}$ or $f_x(a, b)$.

Similarly, the (first order) **partial derivative of f with respect to y** (instead of x) at the point (a,b) is:

$$\frac{d}{dy}f(a,y)\Big|_{y=b} = \lim_{h\to 0}\frac{f(a,b+h)-f(a,b)}{h}$$

and is denoted by

$$\frac{\partial f}{\partial y}\Big|_{(a,b)} \quad \text{or} \quad f_y(a,b).$$

If we let $z=f(x,y)$, we also write

$$f_x=\frac{\partial z}{\partial x}, \quad \text{and} \quad f_y=\frac{\partial z}{\partial y}.$$

In practice, when we compute $f_x(a,b)$ (resp. $f_y(a,b)$), we simply treat the y (resp. x) variable of $f(x,y)$ as constant and differentiate f with respect to x (respectively y) before substituting $x=a$ and $y=b$.

7.3.3 Example

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

Find $f_x(2, 0)$, and $f_y(2, 0)$.

Solution: Treat y as a constant and compute

$$\begin{aligned} f_x &= \frac{d}{dx} f(x, y) = \frac{d}{dx}(x^3 + y)\cos(y^2) \\ &= 3x^2 \cos(y^2) \end{aligned}$$

Then $f_x(2, 0) = 3(2)^2 \cos(0^2) = 12$.

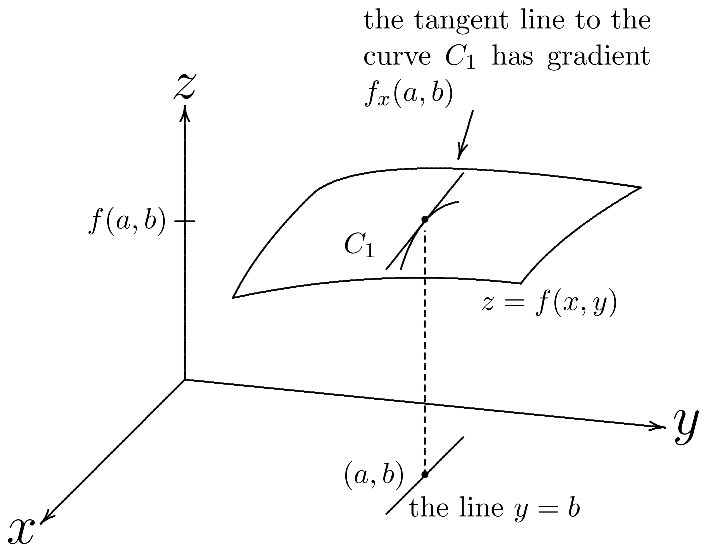
Treat x as a constant and compute

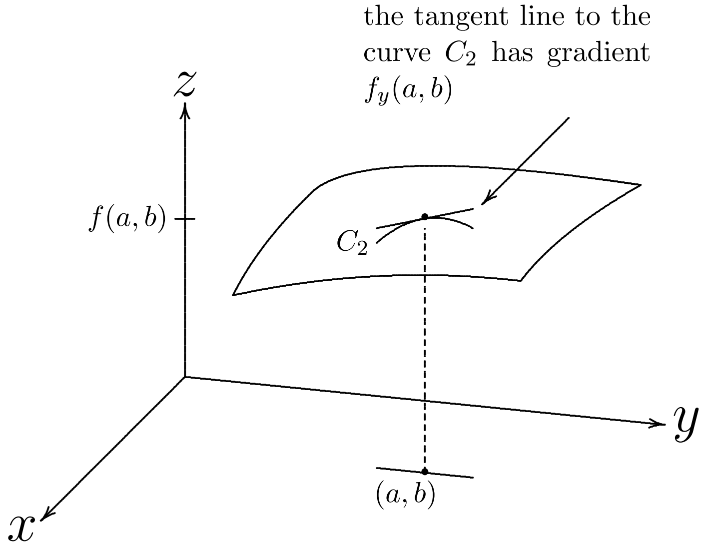
$$\begin{aligned} f_y &= \frac{d}{dy} f(x, y) = \frac{d}{dy}(x^3 + y)\cos(y^2) \\ &= \cos(y^2) - (x^3 + y)\sin(y^2)\ 2y. \end{aligned}$$

Then $f_y(2, 0) = \cos(0^2) - (2^3 + 0)\sin(0^2)\ 2(0) = 1$.

7.3.4 Geometric interpretation

Geometrically, $f_x(a,b)$ measures the rate of change of f in the direction of vector $\mathbf{i}$ at the point (a,b). If we consider the line $y = b$ on the xy-plane parallel to the x-axis and passing through the point (a,b), the image of this line under f is a curve C_1 on the surface $z = f(x,y)$. Then $f_x(a,b)$ is just the gradient of the tangent line to C_1 at (a,b).





Similarly, $f_y(a, b)$ is just the gradient of the tangent line at (a, b) of the curve C_2 traced out as the image of the line $x = a$ under f.

7.3.5 Higher order partial derivatives

We have seen that the partial derivatives f_x and f_y of a function of two variables f are also functions of two variables. Hence, we can study the partial derivatives of f_x and f_y.

The **second order partial derivatives** of f are:

$$f_{xx} = (f_x)_x = \frac{\partial^2 f}{\partial x^2} \qquad f_{xy} = (f_x)_y = \frac{\partial^2 f}{\partial y \partial x}$$
$$f_{yx} = (f_y)_x = \frac{\partial^2 f}{\partial x \partial y} \qquad f_{yy} = (f_y)_y = \frac{\partial^2 f}{\partial y^2}.$$

If $z = f(x, y)$, we also have the following notation:

$$f_{xx} = \frac{\partial^2 z}{\partial x^2} \qquad f_{xy} = \frac{\partial^2 z}{\partial y \partial x} \qquad f_{yx} = \frac{\partial^2 z}{\partial x \partial y} \qquad f_{yy} = \frac{\partial^2 z}{\partial y^2}.$$

7.3.6 Example

Find the second partial derivatives of

$$f(x, y) = 4x^3 + x^2y^3 - 6y^2.$$

Solution: We have $f_x = 12x^2 + 2xy^3$, so

$$f_{xx} = 24x + 2y^3, \quad f_{xy} = 6xy^2.$$

We have $f_y = 3x^2y^2 - 12y$, so

$$f_{yx} = 6xy^2, \quad f_{yy} = 6x^2y - 12.$$

7.3.7 Mixed Derivatives

For most functions in practice, we have

$$f_{xy}(a,b) = f_{yx}(a,b). \tag{1}$$

7.3.8 Example

Let $f(x,y) = xy + \dfrac{e^y}{y^2+1}$. Find f_{yx}.

Solution: The notation f_{yx} means differentiating first with respect to y and then with respect to x.

This is the same as f_{xy}, i.e. we can postpone the differentiation with respect to y and differentiate first with respect to x:

$$f_x = y \implies f_{xy} = (f_x)_y = 1.$$

This is much easier than to first differentiate with respect to y!

7.3.9 Functions of Three or More Variables

For functions of three or more variables, we have similar definitions and notations for partial derivatives. For example, for functions of three variables $f(x, y, z)$, we fix two of the variables and differentiate with respect to the third one.

$$f_x, \quad f_y, \quad f_z, \qquad \text{or} \qquad \frac{\partial f}{\partial x}, \; \frac{\partial f}{\partial y}, \; \frac{\partial f}{\partial z}.$$

7.4 Chain Rule

Suppose the length ℓ, width w and height h of a box change with time. At time t_0, the dimensions of the box are $\ell = 2$ m, $w = 3$ m, $h = 4$ m, and ℓ and w are increasing at a rate of 5 ms^{-1} while h is decreasing at a rate of 6 ms^{-1}. What is the rate of change of the volume of the box at time t_0?

In the above problem, the volume V of the box is a function of three variables in ℓ, w, h:

$$V = V(\ell, w, h)$$

while these three variables are in turn functions of time t: $\ell = \ell(t), \quad w = w(t), \quad h = h(t)$. Clearly, a change of t will cause a change of V.

We say that V is a *composite* function of t and write

$$V(t) = V(\ell(t),\ w(t),\ h(t)).$$

The rate of change of V is given by $\dfrac{dV}{dt}$.

Can we express $\dfrac{dV}{dt}$ in terms of $\dfrac{d\ell}{dt}$, $\dfrac{dw}{dt}$, and $\dfrac{dh}{dt}$?

The answer is given by Chain rule.

7.4.1 Chain rule for one independent variable on $f(x, y)$

Suppose $z = f(x, y)$ is a function of two variables x and y, and $x = x(t)$, $y = y(t)$ are both functions of t. Then z is a function of t: $z(t) = f(x(t),\ y(t))$ and

$$\frac{dz}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}.$$

7.4.2 Example

Let $z = 3xy^2 + x^4 y$, where $x = \sin 2t$, $y = \cos t$. Find $\dfrac{dz}{dt}$.

Solution:

$$\begin{aligned}\frac{dz}{dt} &= \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt} \\ &= (3y^2 + 4x^3 y)(2\cos 2t) + (6xy + x^4)(-\sin t).\end{aligned}$$

7.4.3 Chain rule for two independent variables on $f(x, y)$

Suppose $z = f(x, y)$ is a function of two variables x and y, and $x = x(s, t)$, $y = y(s, t)$ are both functions of two variables s and t. Then z is a function of s and t: $z(s,\ t) = f(x(s, t),\ y(s, t))$ and

$$\frac{\partial z}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s} \quad \text{and} \quad \frac{\partial z}{\partial t} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t}$$

7.4.4 Example

Let $z = e^{2x}\cos 3y$, where $x = st^2$, $y = s^2 t$.

$$\begin{aligned}\frac{\partial z}{\partial s} &= \frac{\partial z}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial s}\\ &= (2e^{2x}\cos 3y)t^2 + (-3e^{2x}\sin 3y)(2st).\\ \frac{\partial z}{\partial t} &= \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial t}\\ &= (2e^{2x}\cos 3y)(2st) + (-3e^{2x}\sin 3y)(s^2).\end{aligned}$$

7.4.5 Chain rule for one independent variable on $f(x, y, z)$

Chain rules can be extended in a similar way for functions of three or more variables.

For example, suppose $w = f(x, y, z)$ is a function of three variables x, y and z, and $x = x(t)$, $y = y(t)$, $z = z(t)$ are functions of t. Then w is a function of t: $\quad w(t) = f(x(t),\ y(t),\ z(t))$ and we have

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

7.4.6 Example

Back to the problem at the beginning of this section.

The volume

$$V(\ell, w, h) = \ell \times w \times h.$$

Hence

$$\frac{\partial V}{\partial \ell} = wh, \quad \frac{\partial V}{\partial w} = \ell h, \quad \frac{\partial V}{\partial h} = \ell w.$$

By chain rule,

$$\begin{aligned} \frac{dV}{dt} &= \frac{\partial V}{\partial \ell}\frac{d\ell}{dt} + \frac{\partial V}{\partial w}\frac{dw}{dt} + \frac{\partial V}{\partial h}\frac{dh}{dt} \\ &= wh\frac{d\ell}{dt} + \ell h\frac{dw}{dt} + \ell w\frac{dh}{dt} \end{aligned}$$

We are given, at time t_0,

$$\ell = 2 \text{ m}, \ w = 3 \text{ m}, \ h = 4 \text{ m},$$

while $\dfrac{d\ell}{dt} = 5\,\text{ms}^{-1}$, $\dfrac{dw}{dt} = 5\,\text{ms}^{-1}$ and $\dfrac{dh}{dt} = -6\,\text{ms}^{-1}$.

Hence

$$\frac{dV}{dt} = (3)(4)(5) + (2)(4)(5) + (2)(3)(-6) = 64 \text{ m}^3\text{s}^{-1}.$$

7.4.7 Chain rule for two independent variables on $f(x, y, z)$

Now suppose $w = f(x, y, z)$ is a function of three variables x, y and z, and $x = x(s, t)$, $y = y(s, t)$, $z = z(s, t)$ are functions of two variables s and t. Then w is a function of s and t:

$w(s,\ t) = f(x(s, t),\ y(s, t),\ z(s, t))$ and we have

$$\frac{\partial w}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial s}$$
$$\frac{\partial w}{\partial t} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial t}$$

7.5 Directional Derivatives

7.5.1 Extension of Partial Derivatives

We have seen earlier that, given a function $f(x, y)$, the partial derivatives give the rates of change of f

with respect to x and y, i.e. along the directions of x- and y-axes.

A natural question to ask is, what is the rate of change of f along an arbitrary direction? This gives rise to the notion of directional derivatives.

Let f be a function of x and y.

The **directional derivative** of f at (a, b) in the direction of a unit vector $\mathbf{u} = u_1\,\mathbf{i} + u_2\,\mathbf{j}$ is

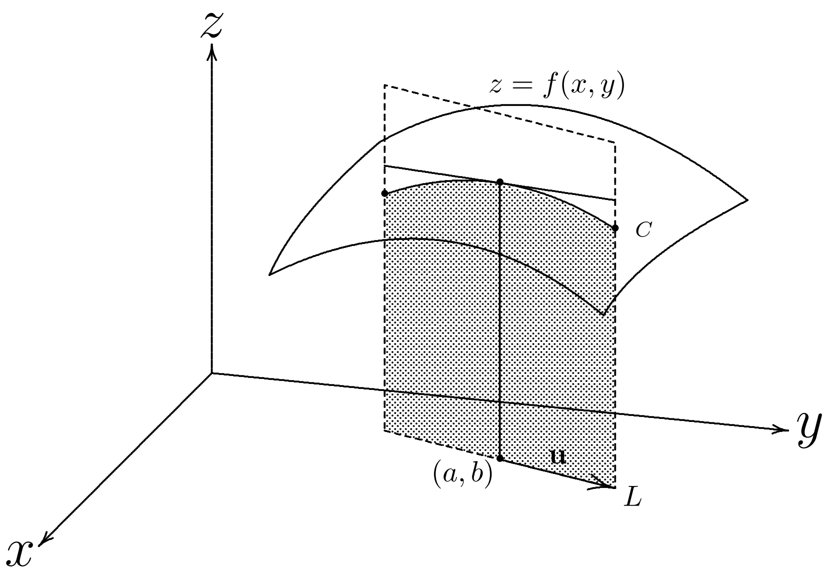
$$D_{\mathbf{u}}f(a,b) = \lim_{h\to 0} \frac{f(a+hu_1, b+hu_2) - f(a,b)}{h}$$

if this limit exists.

Note that $D_{\mathbf{i}}f(a,b) = f_x(a,b)$ and $D_{\mathbf{j}}f(a,b) = f_y(a,b)$, where $\mathbf{i}$ and $\mathbf{j}$ are the standard unit vectors of the xy-plane.

7.5.2 Geometrical meaning

Let L be the line in the xy-plane passing through the point (a, b) and parallel to $\mathbf{u}$. Then L traces out a curve C on the surface represented by $z = f(x, y)$ as shown in the diagram.



Then $D_{\mathbf{u}}f(a, b)$ gives the gradient of the tangent line to the curve C at the point (a, b).

7.5.3 A formula

Since $D_{\mathbf{u}}f(a,b)$ is also the rate of change of $f(x,y)$ at (a,b) in the direction of $\mathbf{u}$, and the coordinates x and y refer to points on the line L:

$$x = a + u_1 t, \quad y = b + u_2 t, \quad z = 0,$$

it follows that the chain rule can be used, as follows:

$$\frac{df}{dt} = \frac{\partial f}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial f}{\partial y} \cdot \frac{dy}{dt}.$$

Thus,

$$D_{\mathbf{u}}f(a,b) = f_x(a,b) \cdot u_1 + f_y(a,b) \cdot u_2.$$

7.5.4 Example

Let $f(x,y) = x^2 - 3xy^2 + 2y^3$. Find $D_{\mathbf{u}}f(2,1)$, where $\mathbf{u} = \frac{\sqrt{3}}{2}\mathbf{i} + \frac{1}{2}\mathbf{j}$.

Solution: First, $f_x = 2x - 3y^2$, $f_y = -6xy + 6y^2$.

Thus $f_x(2,1) = 1$ and $f_y(2,1) = -6$.

Therefore,

$$D_{\mathbf{u}}f(2,1) = (1)(\frac{\sqrt{3}}{2}) + (-6)(\frac{1}{2}) = \frac{\sqrt{3}-6}{2}.$$

Gradient Vector

In view of the expression of the directional derivative in terms of partial derivatives, it is convenient and useful to introduce the notion of a *gradient vector*.

The **gradient** of $f(x, y)$ is the vector (function)

$$\nabla f \;=\; f_x\,\mathbf{i} \;+\; f_y\,\mathbf{j}.$$

For a given unit vector $\mathbf{u} \;=\; u_1\mathbf{i} + u_2\mathbf{j}$, we obtain

$$\begin{aligned}
&\nabla f(a, b) \cdot \mathbf{u} \\
&= \; (f_x(a, b)\,\mathbf{i} \;+\; f_y(a, b)\,\mathbf{j}) \cdot (u_1\mathbf{i} + u_2\mathbf{j}) \\
&= \; f_x(a, b) \cdot u_1 \;+\; f_y(a, b) \cdot u_2 \\
&= \; D_{\mathbf{u}}f(a, b).
\end{aligned}$$

Thus,

$$D_{\mathbf{u}}f(a, b) \;=\; \nabla f(a, b) \cdot \mathbf{u}.$$

Noting that $\nabla f(a,b)$ and $\mathbf{u}$ are vectors, let θ be the angle $(0 \le \theta \le \pi)$ between them. Then

$$D_{\mathbf{u}} f(a,b) \;=\; \nabla f(a,b) \cdot \mathbf{u} \;=\; ||\nabla f(a,b)||\, \cos\theta.$$

Since $-1 \le \cos\theta \le 1$, we obtain some useful properties of the above formula when $\nabla f(a,b) \ne \mathbf{0}$:

Facts.

(1) The function f increases most rapidly in the direction $\nabla f(a,b)$.

(2) The function f decreases most rapidly in the direction $-\nabla f(a,b)$.

Example. Let $f(x,y)=\sqrt{9-x^2-y^2}$. Find the largest possible value of $D_{\mathbf{u}}f(2,1)$.

Solution. The surface $z = f(x,y)$ is the upper hemisphere of the sphere of radius 3 and centred at $(0,0,0)$. First compute

$$f_x = \frac{-x}{\sqrt{9-x^2-y^2}}, \quad f_y = \frac{-y}{\sqrt{9-x^2-y^2}}.$$

The largest possible value of $D_{\mathbf{u}}f(2,1)$ is obtained when $\mathbf{u}$ is in the direction of

$$\begin{aligned}\nabla f(2,1) &= f_x(2,1)\mathbf{i} + f_y(2,1)\mathbf{j}\\ &= -\mathbf{i} - \frac{1}{2}\mathbf{j}.\end{aligned}$$

Now,

$$||\nabla f(2,1)|| = \sqrt{(-1)^2+\left(-\frac{1}{2}\right)^2} = \frac{\sqrt{5}}{2}.$$

Let

$$\mathbf{u} = \frac{\nabla f(2,1)}{||\nabla f(2,1)||} = -\frac{2}{\sqrt{5}}\mathbf{i} - \frac{1}{\sqrt{5}}\mathbf{j}.$$

Thus, the largest possible value of $D_{\mathbf{u}}f(2,1)$ is

$$\begin{aligned}
&\nabla f(2,1)\cdot \mathbf{u}\\
&= (-1)\cdot\left(-\frac{2}{\sqrt{5}}\right) + \left(-\frac{1}{2}\right)\cdot\left(-\frac{1}{\sqrt{5}}\right)\\
&= \frac{\sqrt{5}}{2}.
\end{aligned}$$

7.5.5 Physical meaning

As we mentioned at the beginning of this section, the directional derivative $D_{\mathbf{u}}f(a,b)$ measures the change in the value df of a function f when we move a small distance dt from the point (a,b) in the direction of the vector $\mathbf{u}$:

$$df = D_{\mathbf{u}} f(a, b) \cdot dt.$$

7.5.6 Example

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

Estimate how much the value of f will change if a point Q moves 0.1 unit from $(2, 1)$ towards $(3, 0)$.

Solution: Q moves in the direction $(3\ \mathbf{i} + 0\ \mathbf{j}) - (2\ \mathbf{i} +\ \mathbf{j}) = \mathbf{i} - \mathbf{j}$.

The unit vector $\mathbf{u}$ along this direction is $\dfrac{1}{\sqrt{2}}\mathbf{i} - \dfrac{1}{\sqrt{2}}\mathbf{j}$.

Now $f_x = 2xy^3$, $f_y = 3x^2y^2$.

Thus $f_x(2, 1) = 4$ and $f_y(2, 1) = 12$.

Therefore,

$$D_{\mathbf{u}}f(2,1) = (4)(\frac{1}{\sqrt{2}}) + (12)(-\frac{1}{\sqrt{2}}) = -\frac{8}{\sqrt{2}}.$$

So

$$df = D_{\mathbf{u}}f(2,1) \cdot dt = (-\frac{8}{\sqrt{2}})(0.1) \approx -0.57.$$

So the value of f decreases by approximately 0.57 unit.

7.5.7 Functions of Three Variables

We can also define directional derivatives for functions of three variables.

Let f be a function of x, y and z. The directional derivative of f at (a, b, c) in the direction of a unit vector $\mathbf{u} = u_1\,\mathbf{i} + u_2\,\mathbf{j} + u_3\,\mathbf{k}$ in the xyz space is

$$D_{\mathbf{u}}f(a,b,c)$$

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

if this limit exists.

Similarly, we have the formula

$$D_{\mathbf{u}}f(a,b,c) = f_x(a,b,c)u_1 + f_y(a,b,c)u_2 + f_z(a,b,c)u_3$$

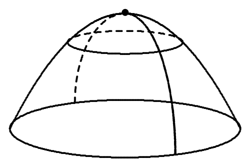
7.6 Maximum and Minimum Values

7.6.1 Local maximum and minimum

(1) $f(x,y)$ has a **local maximum** at (a,b) if

$$f(x,y) \le f(a,b) \quad \text{for all points} \quad (x,y) \ \text{near} \ (a,b).$$

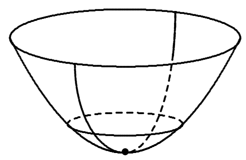
The number $f(a,b)$ is called a **local maximum value**.



(2) $f(x, y)$ has a **local minimum** at (a, b) if

$$f(x, y) \geq f(a, b) \text{ for all points } (x, y) \text{ near } (a, b).$$

The number $f(a, b)$ is called a **local minimum value**.



7.6.2 Critical Points

A function f may have a local maximum or minimum at (a, b) if

(i) $f_x(a, b) = 0$ and $f_y(a, b) = 0$; or

(ii) $f_x(a, b)$ or $f_y(a, b)$ does not exist.

A point of f that satisfies (i) or (ii) above is called a *critical point.*

7.6.3 Example

Let $f(x,y) = x^2 + y^2 + 4x - 8y + 24$. Find the local maxima and local minima of f, if any.

Solution: The partial derivatives exist for any point.

So we solve $f_x(x,y) = 0$ and $f_y(x,y) = 0$.

$$f_x = 2x + 4 = 0 \Rightarrow x = -2$$

$$f_y = 2y - 8 = 0 \Rightarrow y = 4.$$

This gives a solution $(x,y) = (-2,4)$.

Is this point a local maximum, a local minimum, or none?

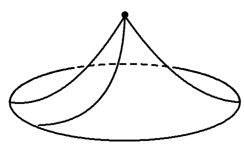
By completing squares, we have

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

Therefore, we conclude that $(-2, 4)$ is a local minimum of f with minimum value 4.

7.6.4 Example

The following diagram shows the graph of a function $f(x, y)$ which has a local maximum at a point (a, b) but $f_x(a, b)$ and $f_y(a, b)$ does not exist.



7.6.5 Saddle points

Let (a, b) be a point of f with $f_x(a, b) = 0$ and $f_y(a, b) = 0$. We say (a, b) is a **saddle point** of

f if there are some directions along which f has a local maximum at (a, b) and some directions along which f has a local minimum at (a, b).

7.6.6 Example

Find the local maximum or local minimum of $f(x, y) = 2y^2 - 3x^2$, if any.

Solution: As before, we solve

$$f_x = -6x = 0 \quad \text{and} \quad f_y = 4y = 0.$$

The only solution is $(x, y) = (0, 0)$. So this is the only critical point of f.

However, this point $(0, 0)$ is neither a local maximum nor a local minimum of f.

To see this, consider the function f along the x-axis which has equation $y = 0$. Substituting this equation into $f(x, y)$, we have

$$f(x, 0) = -3x^2 \leq 0.$$

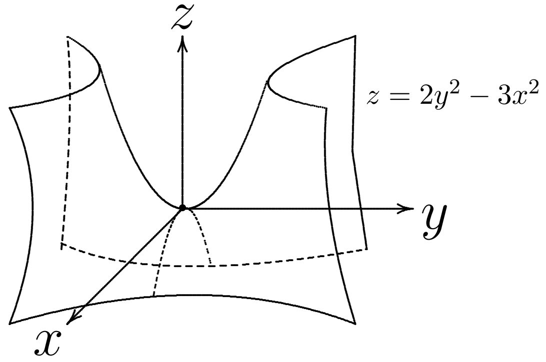
So along x-axis, f has a local maximum at $(0, 0)$.

On the other hand, if we consider f along the y-axis which has equation $x = 0$, we have

$$f(0, y) = 2y^2 > 0.$$

So along y-axis, f has a local minimum at $(0, 0)$.

Therefore f has a saddle point at $(0, 0)$.



7.6.7 Second Derivative Test

When the partial derivatives exist, we can determine the type of critical point using the following systematic approach:

Let $f_x(a,b) = 0$ and $f_y(a,b) = 0$.

$$D = f_{xx}(a,b)f_{yy}(a,b) - f_{xy}(a,b)^2.$$

(a) If $D > 0$ and $f_{xx}(a,b) > 0$, then f has a local minimum at (a,b).

(b) If $D > 0$ and $f_{xx}(a,b) < 0$, then f has a local maximum at (a,b).

(c) If $D < 0$, then f has a saddle point at (a,b).

(d) If $D = 0$, then no conclusion can be drawn.

7.6.8 Example

Find the local maximum, local minimum and saddle points (if any) of $f(x, y) = x^3 + y^3 + 3x^2 - 3y^2 - 8$.

Solution: First, we solve $f_x = 3x^2 + 6x = 0$ and $f_y = 3y^2 - 6y = 0$.

On solving, we obtain $x = 0$ or -2 and $y = 0$ or 2.

Therefore, the critical points are $(0, 0), (0, 2), (-2, 0)$ and $(-2, 2)$.

To apply the second derivative test, we compute the second order partial derivatives.

$$f_{xx} = 6x + 6, \quad f_{yy} = 6y - 6, \quad f_{xy} = 0.$$

Thus $D(x, y) = f_{xx}f_{yy} - f_{xy}^2 = 36xy - 36x + 36y - 36$.

At $(0, 0)$, $D(0, 0) = -36 < 0$.

Hence, f has a saddle point at $(0, 0)$.

At $(0, 2)$, $D(0, 2) = 36 > 0$ and $f_{xx}(0, 2) = 6 > 0$.

Hence f has a local minimum at $(0, 2)$.

At $(-2, 0)$, $D(-2, 0) = 36 > 0$ and $f_{xx}(-2, 0) = -6 < 0$. Hence f has a local maximum at $(-2, 0)$.

At $(-2, 2)$, $D(-2, 2) = -36 < 0$. Hence f has a saddle point at $(-2, 2)$.

7.6.9 Lagrange Multipliers

Many optimization models are subject to certain constraints. For example, production levels depend on labour input and capital expenditure. With a given budget (constraint), a manufacturer aims to maximize production. The method of **Lagrange multipliers** is illustrated below.

7.6.10 Example

Find relative extrema of

$$z = f(x, y) = 12x - 16y + 50$$

subject to the constraint $x^2 + y^2 = 25$.

Solution. The constraint is written as

$$g(x,y) = x^2 + y^2 - 25.$$

The following function is constructed:

$$\begin{aligned} F(x,y,\lambda) &= f(x,y) - \lambda g(x,y) \\ &= 12x - 16y + 50 - \lambda(x^2 + y^2 - 25). \end{aligned}$$

Then set $F_x = 0$, $F_y = 0$, $F_\lambda = 0$ to get respectively

$$\begin{aligned} 12 - 2\lambda x &= 0 \\ -16 - 2\lambda y &= 0 \\ -x^2 - y^2 + 25 &= 0 \end{aligned}$$

Writing the first two equations as $x = \dfrac{6}{\lambda}$, $y = \dfrac{-8}{\lambda}$

and substituting into the third equation, we obtain

$$\begin{aligned} -\frac{36}{\lambda^2} - \frac{64}{\lambda^2} + 25 &= 0 \\ \lambda^2 &= \frac{100}{25} = 4 \\ \lambda &= \pm 2 \end{aligned}$$

If $\lambda = 2$, then $x = \frac{6}{2} = 3$, $y = \frac{-8}{2} = -4$ and

$$z = f(3,-4) = 12(3) - 16(-4) + 50 = 150.$$

If $\lambda = -2$, then $x = \frac{6}{-2} = -3$, $y = \frac{-8}{-2} = 4$ and

$$z = f(-3,4) = 12(-3) - 16(4) + 50 = -50.$$

Thus, subject to the constraint $x^2 + y^2 = 25$, the function $z = f(x,y) = 12x - 16y + 50$ attains:

(1) a local maximum of $z = f(3,-4) = 150$; and

(2) a local minimum of $z = f(-3,4) = -50$.