Chapter 14: Partial Derivatives (25 February version)

Section 14.1: Functions of Several Variables

We shall study functions of more than one variable.

Definition: A function of n variables
Let D be a set of $n-$tuples $(x_1, x_2, ..., x_n)$ of real numbers.
(i) A **real valued function** f on D is a rule that assigns a unique real number

$$w = f(x_1, x_2, ..., x_n)$$

to each n-tuple $(x_1, x_2, ..., x_n)$ in D.
(ii) The set D is called the **domain** of the function f.
(iii) The set of w-values taken on by f is called the **range of the function** f.

Remarks
(a) The symbol w is called the **dependent variable**, while f is said to be a function of n **independent variables** $x_1, x_{2,} ..., x_n$. We also call $x_1, x_{2,} ..., x_n$ the **input variables**, and w the **output variable**.
(b) When we have two variables we often use x, y rather than x_1, x_2. Likewise, when we have three variables, we often use x, y, z rather than x_1, x_2, x_3.

Example
Let

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

This is the distance from (x, y, z) to $\mathbf{0}$.

Domains and Ranges
(a) The function $f(x) = \sqrt{x}$ is not defined as a real number for $x < 0$. So its domain is $[0, \infty)$. Similar considerations apply for functions of more than one variable. Thus

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

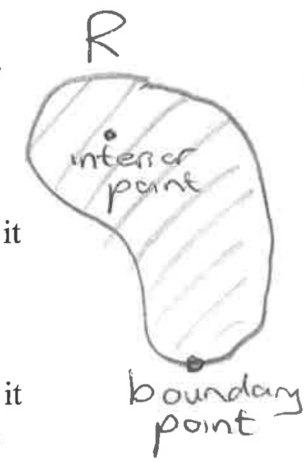
requires $y \geq x^2$ for $f(x, y)$ to be defined. So we can just think of the domain as the set of all (x, y) with $y \geq x^2$. Note that $f(x, y)$ can take on any non-negative value. Thus the range is $[0, \infty)$.
(b) The function

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

is defined as long as $(x, y, z) \neq (0, 0, 0)$. So this describes the domain. Also f takes only positive values and can take any positive value. So the range is $(0, \infty)$.

Functions of Two Variables

Consider the interval $[0, 1]$. The boundary points (or endpoints) are 0 and 1. All other points are interior points of the interval. We can define similar concepts for sets in the plane, also called the xy-plane:



Definitions

(i) A set in the xy-plane is called a **region**.
(ii) A point (x_0, y_0) in a region R in the xy-plane is an **interior point** of R if it is the center of a disk of positive radius that lies entirely in R.
(iii) A point (x_0, y_0) in a set R in the xy-plane is a **boundary point** of R if every disk centered at (x_0, y_0) contains points outside R as well as points inside R.
(iv) The set of all interior points is called the **interior** of R. A set is **open** if it consists entirely of interior points.
(v) The set of all boundary points is called the **boundary** of R. A set is **closed** if it contains all its boundary points.

Examples

(I) Consider

$$R = \{(x, y) \,|x^2 + y^2 < 1\}.$$

This set is open as every point is an interior point. It is called the open unit disk.

(II) Consider

$$R = \{(x, y) \,|x^2 + y^2 \le 1\}.$$

The boundary of this set is $\{(x, y) \,|x^2 + y^2 = 1\}$. The set R is closed as it contains its boundary.
(III) Some sets are neither open nor closed.

Definition

A set in the plane is **bounded** if it is contained in some disk of finite radius. Otherwise, it is **unbounded**.

Graphs, Level Curves, and Contours of Functions of Two Variables

Definitions

Let c be a number and $f(x, y)$ be a function.
(i) The set of points in the plane where $f(x, y) = c$ is called a **level curve** of the function.
(ii) The set of all points $(x, y, f(x, y))$ for (x, y) in the domain of f is called the **graph** of f. We also talk of the **surface** $z = f(x, y)$.

Example

Let

$$f(x, y) = 100 - x^2 - y^2.$$

Plot the level curves

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

and

$$z = f(x, y) = 75.$$

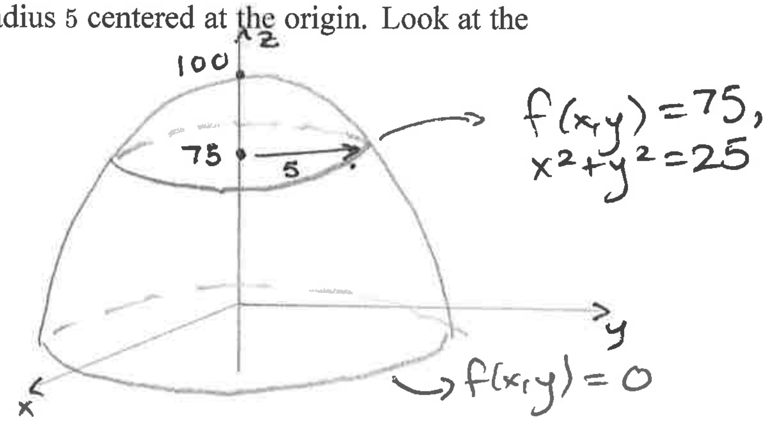
Solution
We see that

$$f(x, y) = 0 \Leftrightarrow x^2 + y^2 = 100.$$

This is the equation of a circle radius 10 centered at the origin. Similarly

$$f(x, y) = 75 \Leftrightarrow x^2 + y^2 = 25.$$

This is the equation of a circle radius 5 centered at the origin. Look at the picture:



Functions of Three Variables

Definition
Let c be a number and $f(x, y, z)$ be a function. The set of points in the plane where $f(x, y, z) = c$ is called a **level surface** of the function.

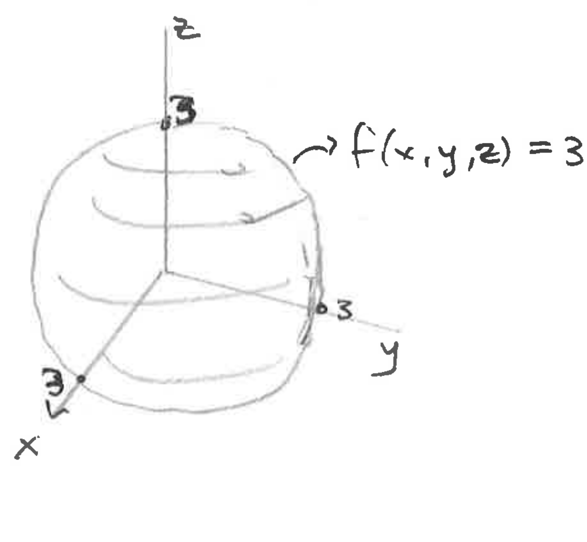
Example
Describe the level surfaces of the function

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

Solution
We see that $f(x, y, z)$ defines the distance from (x, y, z) to the origin. Given $c > 0$, we see that the level surface $f(x, y, z) = c$ is the set of (x, y, z) with

$$x^2 + y^2 + z^2 = c^2,$$

namely the sphere centered on $(0, 0, 0)$ of radius c. When $c = 0$, the level surface is the single point $(0, 0, 0)$.



We can also define the interior and boundary points of a set in space (three dimensions):

Definitions
(i) A point (x_0, y_0, z_0) in a region R in space is an **interior point** of R if it is the center of a solid ball that lies entirely in R.
(ii) A point (x_0, y_0, z_0) in a region R in space is a **boundary point** of R if every solid ball centered at (x_0, y_0, z_0) contains points outside R as well as points inside R.
(iii) The set of all interior points is called the **interior** of R. A set is **open** if it consists entirely of interior points.
(iv) The set of all boundary points is called the **boundary** of R. A set is **closed** if it contains all its boundary points.

Examples
The ball $\{(x, y, z) : x^2 + y^2 + z^2 < 4\}$ is open. The ball $\{(x, y, z) : x^2 + y^2 + z^2 \leq 4\}$ is closed. Its boundary is the surface $\{(x, y, z) : x^2 + y^2 + z^2 = 4\}$.

Section 14.2: Limits and Continuity in Higher Dimensions

You are familiar with limits of functions of one variable, such as

$$\lim_{x \to 0} e^x = 1.$$

Now we explore this concept for functions of two variables:

Definition
We say that a function $f(x, y)$ approaches the **limit** L as (x, y) approaches (x_0, y_0), and write

$$\lim_{(x,y)\to(x_0,y_0)} f(x, y) = L$$

if for every $\varepsilon > 0$, there exists $\delta > 0$ such that whenever (x, y) lies in the domain of f and

$$0 < \sqrt{(x - x_0)^2 + (y - y_0)^2} < \delta,$$

then

$$|f(x, y) - L| < \varepsilon.$$

Remarks
(i) Thus as (x, y) gets closer and closer to (x_0, y_0), so $f(x, y)$ gets closer and closer to L.
(ii) Note that

$$\lim_{(x,y)\to(x_0,y_0)} x = x_0;$$
$$\lim_{(x,y)\to(x_0,y_0)} y = y_0.$$

Here are some basic limit rules. The proofs are similar to those for limits of functions of one variable.

Theorem 1 - Properties of Limits of Functions of Two Variables
Assume that L, M, k are real numbers and

$$\lim_{(x,y)\to(x_0,y_0)} f(x, y) = L \text{ and } \lim_{(x,y)\to(x_0,y_0)} g(x, y) = M.$$

(1) **Sum Rule**

$$\lim_{(x,y)\to(x_0,y_0)} (f(x, y) + g(x, y)) = L + M.$$

(2) **Difference Rule**

$$\lim_{(x,y)\to(x_0,y_0)} (f(x, y) - g(x, y)) = L - M.$$

(3) **Constant Multiple Rule**

$$\lim_{(x,y)\to(x_0,y_0)} kf(x, y) = kL.$$

(4) **Product Rule**

$$\lim_{(x,y)\to(x_0,y_0)} (f(x, y)\, g(x, y)) = LM.$$

(5) **Quotient Rule**

$$\lim_{(x,y)\to(x_0,y_0)} \frac{f(x,y)}{g(x,y)} = \frac{L}{M} \text{ if } M \neq 0.$$

(6) **Power Rule**

$$\lim_{(x,y)\to(x_0,y_0)} f(x,y)^n = L^n \text{ if } n \text{ is a positive integer.}$$

(7) **Root Rule**

$$\lim_{(x,y)\to(x_0,y_0)} f(x,y)^{1/n} = L^{1/n} \text{ if } n \text{ is a positive integer; if } n \text{ is even, we assume } L > 0.$$

By using these rules, we see that we can often evaluate the limit at (x_0, y_0) just by evaluating the function at (x_0, y_0).

Examples
(1)

$$\lim_{(x,y)\to(0,1)} \frac{x + 6xy - y^2}{\sin x - y^3} = \frac{0 + 6(0)(1) - 1^2}{\sin 0 - 1^3} = \frac{-1}{-1} = 1.$$

(2)

$$\begin{aligned}
\lim_{(x,y)\to(0,0)} \frac{x^2 - xy}{\sqrt{x} - \sqrt{y}} &= \lim_{(x,y)\to(0,0)} \frac{x(x-y)}{\sqrt{x} - \sqrt{y}} \\
&= \lim_{(x,y)\to(0,0)} \frac{x\left(\sqrt{x} - \sqrt{y}\right)\left(\sqrt{x} + \sqrt{y}\right)}{\sqrt{x} - \sqrt{y}} \\
&= \lim_{(x,y)\to(0,0)} x\left(\sqrt{x} + \sqrt{y}\right) = 0(0+0) = 0.
\end{aligned}$$

(3) Find

$$\lim_{(x,y)\to(0,0)} \frac{10xy^2}{x^2 + y^2}$$

if it exists.
Solution
We get some intuition by seeing that $0 \leq \frac{y^2}{x^2+y^2} \leq 1$, and see that therefore

$$\left| \frac{10xy^2}{x^2 + y^2} \right| = 10|x| \frac{y^2}{x^2 + y^2} \leq 10|x| \tag{1}$$

and we know $10|x|$ has limit 0. So we expect the function has limit 0 at $(0,0)$. We can make this rigorous: we have to show that for any $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$0 < \sqrt{(x-0)^2 + (y-0)^2} < \delta \Rightarrow \left| \frac{xy^2}{x^2 + y^2} - 0 \right| < \varepsilon.$$

Choose $\delta = \varepsilon/10$. Then

$$0 < \sqrt{(x-0)^2 + (y-0)^2} < \frac{\varepsilon}{10},$$

so that

$$0 < |x| = \sqrt{x^2} < \frac{\varepsilon}{10}$$
$$\Rightarrow \left|\frac{10xy^2}{x^2+y^2} - 0\right| = 10|x|\frac{y^2}{x^2+y^2} \le 10|x| < 10\frac{\varepsilon}{10} = \varepsilon.$$

So indeed,

$$\lim_{(x,y)\to(0,0)} \frac{10xy^2}{x^2+y^2} = 0.$$

Example
Does

$$\lim_{(x,y)\to(0,0)} \frac{y}{x} \text{ exist?}$$

Solution
The domain of $f(x,y) = \frac{y}{x}$ has as its domain $\{(x,y) : x \neq 0\}$. Let us first get some intuition. Firstly, if $x \neq 0$, $f(x,0) = 0$ so if the limit exists it has to be 0. But if $x \neq 0$, $f(x,x) = \frac{x}{x} = 1$, so the limit cannot exist (The limiting values depends on how you approach 0).

Continuity
We can define this as we did for functions of one variable:

Definition
A function $f(x,y)$ is **continuous at** (x_0, y_0) if f is defined and has a limit at (x_0, y_0) and

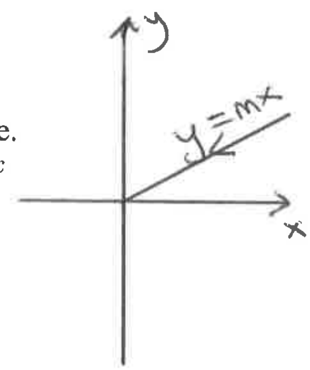
$$\lim_{(x,y)\to(x_0,y_0)} f(x,y) = f(x_0, y_0).$$

A function f is **continuous** if it is continuous at every point of its domain.

Example
Let

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

Show that f is continuous at every point except $(0,0)$.
Solution
At any point $(x,y) \neq (0,0)$, the denominator x^2+y^2 in $f(x,y)$ is non-0. The limit of f as we approach (x,y), can then be computed using our rules for sums, products and quotients of limits, and this will equal the function value. Now suppose we consider the limit at $(0,0)$. Let m be a number, and $y = mx$ with $x \neq 0$. We see that



$$f(x,y) = f(x,mx) = \frac{2mx^2}{x^2+m^2x^2} = \frac{2m}{1+m^2}.$$

This gives different values for different values of m. For example, we get 0

for $m = 0$ and $\frac{1}{2}$ for $m = 1$. So the limit

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

cannot exist.

Continuity of Compositions of Functions
Let $f(x,y)$ be continuous at (x_0, y_0) and $g(s)$ be a single variable function continuous at $s = f(x_0, y_0)$. Let $h(x,y) = g(f(x,y)) = g \circ f(x,y)$ denote the composite function. Then h is continuous at (x_0, y_0).

Functions of More than Two Variables
We can define limits and continuity of functions of three or more variables in exactly the same way. The same rules and properties hold.

Maxima and Minima of Continuous Functions on Closed, Bounded Sets
If f is a continuous function of one variable on $[a, b]$, we know that it assumes its maximum at some point in $[a, b]$, and also assumes its minimum at some point in $[a, b]$. Suppose now R is a closed and bounded set in two or more dimensions. Suppose that f is a continuous function on R. Then it also assumes its maximum at some point in R and assumes its minimum at some point in R.

Section 14.3: Partial Derivatives

You know that if $g(t)$ is a function of one variable t, then its derivative at a point t_0 is

$$g'(t_0) = \lim_{h \to 0} \frac{g(t_0 + h) - g(t_0)}{h},$$

if the limit exists. When defining derivatives for functions $f(x, y)$ of two variables, one complication is that we can approach (x_0, y_0) along infinitely many directions - for example along the $x-$axis, and the $y-$axis, and on any ray through (x_0, y_0). Thus it makes sense to focus on specific directions, leading to the notion of partial derivatives:

Definition
The **partial derivative of f with respect to** x at (x_0, y_0) is

$$\frac{\partial f}{\partial x}_{|(x_0, y_0)} = \lim_{h \to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h}$$

provided the limit exists.
Remarks
(a) We could also fix $y = y_0$ and define the partial derivative as

$$\frac{d}{dx} f(x, y_0)_{|x = x_0}.$$

(b) There are many notations for the partial derivative. If we set $z = f(x, y)$, we use all of

$$\frac{\partial f}{\partial x}(x_0, y_0) \text{ and } f_x(x_0, y_0) \text{ and } z_x(x_0, y_0).$$

When we think of the partial derivative as a function, we omit (x_0, y_0).
(c) $\frac{\partial f}{\partial x}_{|(x_0, y_0)}$ is the slope of the curve $z = f(x, y_0)$ in the plane $y = y_0$ at (x_0, y_0).
(d) We can use all the usual rules for sums, products, quotients, when calculating partial derivatives.

Definition
The **partial derivative of f with respect to** y at (x_0, y_0) is

$$\frac{\partial f}{\partial y}_{|(x_0, y_0)} = \lim_{h \to 0} \frac{f(x_0, y_0 + h) - f(x_0, y_0)}{h}$$

provided the limit exists.
Remark
Again, we use f_y etc.

Example
Let

$$f(x, y) = \frac{2y}{y + \sin x + \log y}.$$

Find f_x and f_y.

Solution

$$f_x = \frac{\partial}{\partial x}\left(\frac{2y}{y+\sin x+\log y}\right).$$

Note that we regard y as a constant when taking the partial derivative with respect to x :

$$\begin{aligned} f_x &= 2y\frac{\partial}{\partial x}\left(\frac{1}{y+\sin x+\log y}\right) \\ &= -2y\frac{\frac{\partial}{\partial x}(y+\sin x+\log y)}{(y+\sin x+\log y)^2} \\ &= -2y\frac{0+\cos x+0}{(y+\sin x+\log y)^2} = -\frac{2y\cos x}{(y+\sin x+\log y)^2}. \end{aligned}$$

Next,

$$\begin{aligned} f_y &= \frac{\partial}{\partial y}\left(\frac{2y}{y+\sin x+\log y}\right) \\ &= \frac{\left[\frac{\partial}{\partial y}(2y)\right][y+\sin x+\log y]-[2y]\left[\frac{\partial}{\partial y}(y+\sin x+\log y)\right]}{(y+\sin x+\log y)^2} \\ &= \frac{[2][y+\sin x+\log y]-[2y]\left[1+0+\frac{1}{y}\right]}{(y+\sin x+\log y)^2} \\ &= \frac{2y+2\sin x+2\log y-2y-2}{(y+\sin x+\log y)^2} = \frac{2\sin x+2\log y-2}{(y+\sin x+\log y)^2}. \end{aligned}$$

We can also use implicit differentiation for partial derivatives:

Example

Consider the function z of x, y given by the equation

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

Find $\frac{\partial z}{\partial x}$.

Solution

Differentiate both sides of the equation with respect to x :

$$\begin{aligned} \frac{\partial}{\partial x}(yz-\ln z) &= \frac{\partial}{\partial x}(x+y) \\ \Rightarrow y\frac{\partial z}{\partial x}-\frac{1}{z}\frac{\partial z}{\partial x} &= 1+0 \\ \Rightarrow \frac{\partial z}{\partial x}\left(y-\frac{1}{z}\right) &= 1 \\ \Rightarrow \frac{\partial z}{\partial x} = \frac{1}{y-\frac{1}{z}} &= \frac{z}{zy-1}. \end{aligned}$$

Functions of More than Two Variables

We can also define partial derivatives of functions of 3 or 4 or more variables, in the same way.

Example
Let
$$f(x,y,z) = x\sin(y+3z).$$
Find $\frac{\partial f}{\partial z}$.
Solution
$$\begin{aligned}\frac{\partial f}{\partial z} &= \frac{\partial}{\partial z}(x\sin(y+3z)) \\ &= x\frac{\partial}{\partial z}(\sin(y+3z)) \\ &= x\cos(y+3z)\left(\frac{\partial}{\partial z}(y+3z)\right) \\ &= x\cos(y+3z)\,3 = 3x\cos(y+3z).\end{aligned}$$

Partial Derivatives and Continuity
A function $f(x,y)$ can have both $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ existing at a point even though f is not continuous at the point. Here is an example:

Example
Let
$$f(x,y) = \begin{cases} 0, & xy \neq 0 \\ 1, & xy = 0 \end{cases}.$$
(a) Find the limit of $f(x,y)$ as (x,y) approaches $(0,0)$ along $y = x$ and hence that f is not continuous at the origin.
(b) Show that $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ exist at $(0,0)$.
Solution
(a) Here $xy = x^2 \neq 0$, so $f(x,y) = 0$, and thus the limit of $f(x,y)$ along $y = x$ will just be 0. On the other hand $f(x,0) = 1$ for any x, so the limit of $f(x,y)$ along $y = 0$ will just be 1. So
$$\lim_{(x,y)\to(0,0)} f(x,y)$$
does not exist.
(b) Now
$$\frac{\partial f}{\partial x}_{|(0,0)} = \lim_{h\to 0}\frac{f(0+h,0)-f(0,0)}{h} = \lim_{h\to 0}\frac{1-1}{h} = 0.$$
Also
$$\frac{\partial f}{\partial y}_{|(0,0)} = \lim_{h\to 0}\frac{f(0,0+h)-f(0,0)}{h} = \lim_{h\to 0}\frac{1-1}{h} = 0.$$
So both partial derivatives exist. We shall later define differentiability of f at $(0,0)$ to exclude this pathology.

Second-Order Partial Derivatives

We can partially differentiate more than once. Thus the second partial derivative of f w.r.t x is

$$\frac{\partial^2 f}{\partial x^2} \text{ or } f_{xx}$$

and we can also define

$$\frac{\partial^2 f}{\partial y^2} \text{ or } f_{yy};$$

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) \text{ or } f_{yx} = (f_y)_x \text{; and } \frac{\partial^2 f}{\partial y \partial x} \text{ or } f_{xy}.$$

Example

Let

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

Find

$$\frac{\partial^2 f}{\partial x^2}; \frac{\partial^2 f}{\partial y^2}; \frac{\partial^2 f}{\partial x \partial y}; \frac{\partial^2 f}{\partial y \partial x}.$$

Solution

$$\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}(x \cos y + y e^x) = \cos y + y e^x.$$

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

Now find partial derivatives again:

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial}{\partial x}(\cos y + y e^x) = y e^x.$$

$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial}{\partial y}(-x \sin y + e^x) = -x \cos y.$$

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial}{\partial x}(-x \sin y + e^x) = -\sin y + e^x.$$

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial}{\partial y}(\cos y + y e^x) = -\sin y + e^x.$$

Notice that

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

Theorem 2: The Mixed Derivative Theorem

If $f(x, y)$ and its partial derivative f_x, f_y, f_{xy}, and f_{yx} are defined in an open region containing (a, b) and all are continuous at (a, b), then

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

Partial Derivatives of Higher Order

We can differentiate further and consider

$$f_{xxx} \text{ or } f_{yyxx} \text{ etc.}$$

Differentiability

If $f(x)$ is a function of one variable, and $f'(x)$ exists, then we know that f is continuous at x. We want to make sure that our notion of differentiability for functions of more than one variable preserves this nice property. We have already seen that a function $f(x, y)$ of two variables can be non-differentiable at a point even when f_x and f_y exist. So we need more.

Definition

(a) A function $z = f(x, y)$ is **differentiable at** (x_0, y_0) if $f_x(x_0, y_0)$ and $f_y(x_0, y_0)$ exist, and the change

$$\Delta z = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0)$$

satisfies an equation of the form

$$\Delta z = f_x(x_0, y_0)\,\Delta x + f_y(x_0, y_0)\,\Delta y + \varepsilon_1 \Delta x + \varepsilon_2 \Delta y$$

in which each of $\varepsilon_1, \varepsilon_2 \to 0$ as both $\Delta x, \Delta y \to 0$.

(b) We call f **differentiable** if it is differentiable at every point of its domain, and then say its graph is a **smooth surface**.

Theorem 3/ Corollary

If the partial derivatives f_x and f_y of a function $f(x, y)$ are continuous in an open region R, then f is differentiable at every point of R.

Theorem 4 Differentiability implies Continuity

If a function $f(x, y)$ is differentiable at a point (x_0, y_0), then it is also continuous at (x_0, y_0).

Section 14.4: The Chain Rule

The chain rule for a composite function $w = f(g(t))$ says that

$$\begin{aligned}\frac{dw}{dt} &= f'(g(t))\, g'(t) \\ &= \frac{dw}{dx}\frac{dx}{dt},\end{aligned}$$

where $x = g(t)$. There is an analogue for functions of several variables. We start with

Functions of Two Variables

Theorem 5: Chain Rule for functions of one independent variable t and two intermediate variables x, y
If $w = f(x, y)$ is differentiable and if $x = x(t)$ and $y = y(t)$ are differentiable functions of t, then the composite function $w = f(x(t), y(t))$ is a differentiable function of t and

$$\frac{dw}{dt} = f_x(x(t), y(t))\, x'(t) + f_y(x(t), y(t))\, y'(t),$$

or

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

Remark
We could also write

$$\frac{dw}{dt} = \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt}.$$

Proof
Fix $P_0 = (x_0, y_0)$, and assume that $x_0 = x(t_0)$, $y_0 = y(t_0)$. Since w is differentiable at P_0,

$$\begin{aligned}\Delta w &= f_x(x_0, y_0)\,\Delta x + f_y(x_0, y_0)\,\Delta y + \varepsilon_1 \Delta x + \varepsilon_2 \Delta y \\ &= \frac{\partial w}{\partial x}_{|P_0} \Delta x + \frac{\partial w}{\partial y}_{|P_0} \Delta y + \varepsilon_1 \Delta x + \varepsilon_2 \Delta y,\end{aligned}$$

where $\varepsilon_1, \varepsilon_2 \to 0$ as $\Delta x, \Delta y \to 0$. Now divide by Δt,

$$\frac{\Delta w}{\Delta t} = \frac{\partial w}{\partial x}_{|P_0} \frac{\Delta x}{\Delta t} + \frac{\partial w}{\partial y}_{|P_0} \frac{\Delta y}{\Delta t} + \varepsilon_1 \frac{\Delta x}{\Delta t} + \varepsilon \frac{\Delta y}{\Delta t}.$$

Note that as $\Delta t \to 0$,

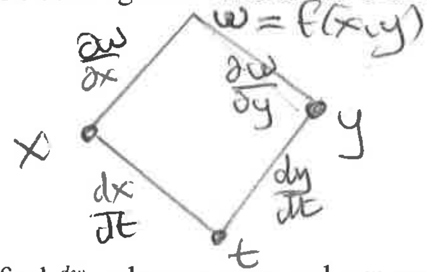
$$\frac{\Delta x}{\Delta t} \to \frac{dx}{dt}_{|t_0};\ \frac{\Delta y}{\Delta t} \to \frac{dy}{dt}_{|t_0}.$$

Then letting $\Delta t \to 0$,

$$\frac{dw}{dt}_{|t_0} = \frac{\partial w}{\partial x}_{|P_0} \frac{dx}{dt}_{|t_0} + \frac{\partial w}{\partial y}_{|P_0} \frac{dy}{dt}_{|t_0}.$$

■

Remembering the chain rule using branch diagrams
One way to remember the chain rule is to use a branch diagram. At the top of the diagram is the dependent variable, e.g. $w = f(x, y)$. Then it branches out to the intermediate variables, and finally to the independent variable. Along each branch, we write the appropriate derivatives, and multiply these. Then finally we add the terms coming from each of the branches.



Example 1
Use the chain rule to find $\frac{dw}{dt}$, where $w = xy$ and $x = \cos t$ and $y = \sin t$. Hence find the value when $t = \frac{\pi}{4}$.
Solution
By the chain rule

$$\begin{aligned}
\frac{dw}{dt} &= \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt} \\
&= \frac{\partial (xy)}{\partial x}\frac{d(\cos t)}{dt} + \frac{\partial (xy)}{\partial y}\frac{d(\sin t)}{dt} \\
&= (y)(-\sin t) + (x)(\cos t) \\
&= -\sin^2 t + \cos^2 t \\
&= \cos 2t.
\end{aligned}$$

When $t = \frac{\pi}{4}$, we obtain

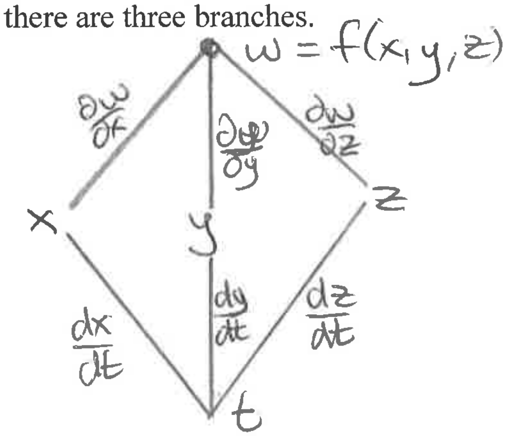
$$\frac{dw}{dt} = \cos\frac{\pi}{2} = 0.$$

Next, we consider a function $f(x, y, z)$ with x, y, z functions of t:

Theorem 6: Chain Rule for functions of one independent variable t and three intermediate variables x, y, z
If $w = f(x, y, z)$ is differentiable and x, y, z are differentiable functions of t, then

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

Branch Diagram
There are now three intermediate variables x, y, z, so there are three branches.

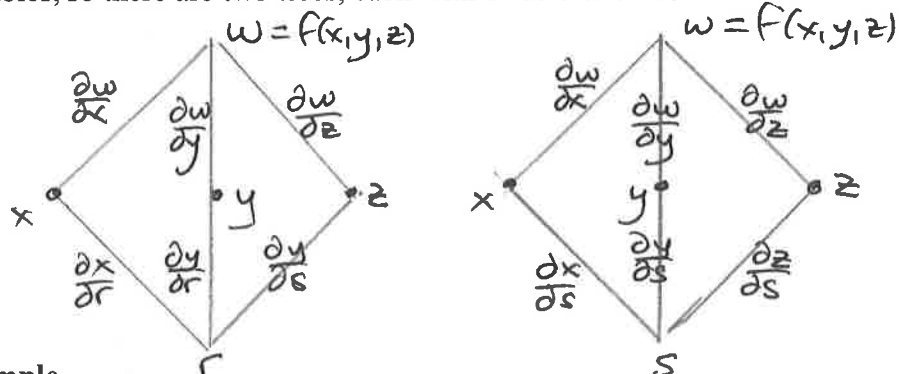


Now consider the case where x, y, z are functions of two, rather than one, variable:

Theorem 7: Chain Rule for functions of two independent variables r, s and three intermediate variables x, y, z
If $w = f(x, y, z)$ is differentiable and x, y, z are differentiable functions of r, s, then

$$\frac{dw}{dr} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial r} + \frac{\partial w}{\partial z}\frac{\partial y}{\partial r}.$$
$$\frac{dw}{ds} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial s}.$$

Branch Diagram
There are now three intermediate variables x, y, z, and two independent variables, so there are two trees, each with three branches.



Example
Let

$$w = f(x, y, x) = x + 2y + z^2;$$

and

$$x = \frac{r}{s};\; y = r^2 + \ln s;\; z = 2r,$$

Find

$$\frac{\partial w}{\partial r} \text{ and } \frac{\partial w}{\partial s}.$$

Solution

Now

$$\begin{aligned}
\frac{dw}{dr} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial r} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial r} \\
&= \frac{\partial\left(x+2y+z^2\right)}{\partial x}\frac{\partial\left(r/s\right)}{\partial r} \\
&\quad + \frac{\partial\left(x+2y+z^2\right)}{\partial y}\frac{\partial\left(r^2+\ln s\right)}{\partial r} \\
&\quad + \frac{\partial\left(x+2y+z^2\right)}{\partial z}\frac{\partial\left(2r\right)}{\partial r} \\
&= (1)\left(\frac{1}{s}\right) + (2)(2r) + (2z)\,2 \\
&= \frac{1}{s} + 4r + 4(2r) = \frac{1}{s} + 12r.
\end{aligned}$$

Next

$$\begin{aligned}
\frac{dw}{ds} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial s} \\
&= \frac{\partial\left(x+2y+z^2\right)}{\partial x}\frac{\partial\left(r/s\right)}{\partial s} \\
&\quad + \frac{\partial\left(x+2y+z^2\right)}{\partial y}\frac{\partial\left(r^2+\ln s\right)}{\partial s} \\
&\quad + \frac{\partial\left(x+2y+z^2\right)}{\partial z}\frac{\partial\left(2r\right)}{\partial s} \\
&= (1)\left(-\frac{r}{s^2}\right) + (2)\left(\frac{1}{s}\right) + (2z)\,0 \\
&= -\frac{r}{s^2} + \frac{2}{s}.
\end{aligned}$$

Functions of one or two intermediate variables and two independent variables

(I) If $w = f(x, y)$ and $x = g(r, s)$ and $y = h(r, s)$, then the equations above simplify by losing one term, namely that for z :

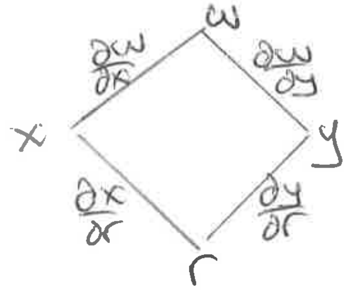
$$\frac{dw}{dr} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial r}.$$

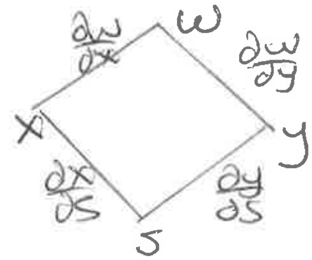
$$\frac{dw}{ds} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial s}.$$

(II) If $w = f(x)$ and $x = g(r, s)$ and $y = h(r, s)$, then the equations above simplify more by losing one more term, namely that for y :

$$\frac{dw}{dr} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial r}.$$

$$\frac{dw}{ds} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial s}.$$





Example

$$w = x^2 + y^2 \text{ and } x = r - s \text{ and } y = r + s.$$

Find $\frac{\partial w}{\partial r}$ and $\frac{\partial w}{\partial s}$.
Solution

$$\begin{aligned}
\frac{dw}{dr} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial r} \\
&= \frac{\partial\left(x^2+y^2\right)}{\partial x}\frac{\partial\left(r-s\right)}{\partial r} + \frac{\partial\left(x^2+y^2\right)}{\partial y}\frac{\partial\left(r+s\right)}{\partial r} \\
&= (2x)(1) + (2y)(1) \\
&= 2(r-s) + 2(r+s) = 4r.
\end{aligned}$$

$$\begin{aligned}
\frac{dw}{ds} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial s} \\
&= \frac{\partial\left(x^2+y^2\right)}{\partial x}\frac{\partial\left(r-s\right)}{\partial s} + \frac{\partial\left(x^2+y^2\right)}{\partial y}\frac{\partial\left(r+s\right)}{\partial s} \\
&= (2x)(-1) + (2y)(1) \\
&= 2(r-s)(-1) + 2(r+s) = 4s.
\end{aligned}$$

Implementation Differentiation Revisited
We can use the two-variable rule to do implicit differentiation. Suppose that
(1) $F(x,y)$ is differentiable;
(2) The equation $w = F(x,y) = 0$ defines y implicitly as a function of x, say $y = h(x)$. Then $w = F(x, h(x))$ is a function of two intermediate variables and one independent variable. By the chain rule, and as $w = 0$,

$$\begin{aligned}
0 &= \frac{dw}{dx} = \frac{\partial F}{\partial x}\frac{dx}{dx} + \frac{\partial F}{\partial y}\frac{dy}{dx} \\
&= (F_x)(1) + (F_y)\left(\frac{dy}{dx}\right).
\end{aligned}$$

Assuming that $F_y \neq 0$, we can solve for $\frac{dy}{dx}$, obtaining

$$\frac{dy}{dx} = -\frac{F_x}{F_y}.$$

Let's state this formally:

Theorem 8: Formula for Implicit Differentiation
Suppose that $F(x,y)$ is differentiable, and that the equation $F(x,y) = 0$

defines y as a differentiable function of x. Then at any point where $F_y \neq 0$,

$$\frac{dy}{dx} = -\frac{F_x}{F_y}.$$

Example

Let

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

and use the equation $F(x, y) = 0$ to define y as a function of x. Find $\frac{dy}{dx}$.

Solution

Then

$$\begin{aligned} \frac{dy}{dx} &= -\frac{F_x}{F_y} \\ &= -\frac{-2x - \cos(xy)\, y}{2y - \cos(xy)\, x} = \frac{2x + \cos(xy)\, y}{2y - \cos(xy)\, x}. \end{aligned}$$

We can also extend this to functions of three variables. Suppose we have a differentiable function $F(x, y, z)$ of three variables, and the equation

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

defines z implicitly as a function $z = f(x, y)$. Proceeding as above, we obtain the equations

$$\frac{\partial z}{\partial x} = -\frac{F_x}{F_z} \text{ and } \frac{\partial z}{\partial y} = -\frac{F_y}{F_z}.$$

Example

Find $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ at $(0, 0, 0)$ if z is defined as a function of x, y by

$$x^3 + z^2 + ye^{xz} + z\cos y = 0.$$

Solution

Let

$$F(x, y, z) = x^3 + z^2 + ye^{xz} + z\cos y.$$

Then

$$\begin{aligned} F_x &= 3x^2 + yze^{xz}; \\ F_y &= e^{xz} - z\sin y; \\ F_z &= 2z + yxe^{xz} + \cos y. \end{aligned}$$

Then

$$\begin{aligned} \frac{\partial z}{\partial x} &= -\frac{F_x}{F_z} = -\frac{3x^2 + yze^{xz}}{2z + yxe^{xz} + \cos y} \\ \text{and } \frac{\partial z}{\partial y} &= -\frac{F_y}{F_z} = -\frac{e^{xz} - z\sin y}{2z + yxe^{xz} + \cos y}. \end{aligned}$$

At $(x, y, z) = (0, 0, 0)$, we have

$$\frac{\partial z}{\partial x} = -\frac{0}{0 + 0 + \cos 0} = 0; \text{ and } \frac{\partial z}{\partial y} = -\frac{1 - 0}{0 + 0 + \cos 0} = -1.$$

Functions of more than three variables

One can form a chain rule for functions of more than three variables.

Section 14.5: Directional Derivatives and Gradient Vectors

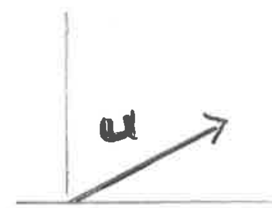
If $z = f(x, y)$ is a differentiable function of (x, y), then we know that $f_x = \frac{\partial f}{\partial x}$ is the rate of change of f in the horizontal or x direction. Similarly, $f_y = \frac{\partial f}{\partial y}$ is the rate of change of f in the vertical or y direction. We can more generally define the derivative in any direction defined by a unit vector:

Definition
Let $\mathbf{u} = (u_1, u_2) = u_1\mathbf{i} + u_2\mathbf{j}$ be a unit vector, that is $|\mathbf{u}| = 1$. The **derivative of a function f of two variables at $P(x_0, y_0)$ in the direction of the unit vector $\mathbf{u}$ is**

$$\left(\frac{df}{ds}\right)_{\mathbf{u},P_0} = \lim_{s\to 0} \frac{f(x_0 + su_1, y_0 + su_2) - f(x_0, y_0)}{s}$$

provided the limit exists. We also denote it by

$$(D_{\mathbf{u}}f)_{P_0}.$$



Example
Find the derivative of $f(x, y) = x^2 + xy$ at $P_0(1, 2)$ in the direction of the unit vector at $\mathbf{u} = \frac{1}{\sqrt{2}}\mathbf{i} + \frac{1}{\sqrt{2}}\mathbf{j}$.
Solution

$$\begin{aligned}
\left(\frac{df}{ds}\right)_{\mathbf{u},P_0} &= \lim_{s\to 0} \frac{f\left(1 + s\frac{1}{\sqrt{2}}, 2 + s\frac{1}{\sqrt{2}}\right) - f(1,2)}{s} \\
&= \lim_{s\to 0} \frac{\left(1 + s\frac{1}{\sqrt{2}}\right)^2 + \left(1 + s\frac{1}{\sqrt{2}}\right)\left(2 + s\frac{1}{\sqrt{2}}\right) - (1+2)}{s} \\
&= \lim_{s\to 0} \frac{1 + 2s\frac{1}{\sqrt{2}} + \frac{s^2}{2} + 2 + 3s\frac{1}{\sqrt{2}} + \frac{s^2}{2} - (1+2)}{s} \\
&= \lim_{s\to 0} \frac{5s\frac{1}{\sqrt{2}} + s^2}{s} = \frac{5}{\sqrt{2}}.
\end{aligned}$$

Interpretation of the Directional Derivative
Suppose for example that $T(x, y)$ is the temperature at a point (x, y) in the region. Then $T_{\mathbf{u}}(x_0, y_0)$ is the rate of change of temperature in the direction $\mathbf{u}$ at the point (x_0, y_0).

There is a way to calculate the directional derivative in any direction u using the partial derivatives in the x, y directions. First we need:

Definition: The Gradient Vector

The **gradient vector (gradient)** of $f(x, y)$ at a point $P(x_0, y_0)$ is the vector

$$\nabla f = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial x}\mathbf{j} = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right),$$

where the partial derivatives are evaluated at $P_0 = (x_0, y_0)$.

Remark

We call this grad f or del f. We also write $(\nabla f)_{P_0}$.

Theorem 9: Directional Derivatives and Dot Products

If $f(x, y)$ is differentiable in an open region containing $P_0 = (x_0, y_0)$, then

$$\left(\frac{df}{ds}\right)_{\mathbf{u}, P_0} = (\nabla f)_{P_0} \cdot \mathbf{u}.$$

Equivalently

$$D_{\mathbf{u}} f = \nabla f \cdot \mathbf{u}.$$

Proof

We can use the chain rule. Write $x = x_0 + su_1$ and $y = y_0 + su_2$, where s is small but not zero. By the chain rule, with partials taken at P_0,

$$\begin{aligned}\left(\frac{df}{ds}\right)_{\mathbf{u}, P_0} &= \frac{\partial f}{\partial x}\frac{dx}{ds} + \frac{\partial f}{\partial y}\frac{dy}{ds} \\ &= \frac{\partial f}{\partial x}u_1 + \frac{\partial f}{\partial y}u_2 \\ &= \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) \cdot (u_1, u_2) \\ &= (\nabla f)_{P_0} \cdot \mathbf{u}.\end{aligned}$$

■

Example

Let

$$f(x, y) = xe^y + \cos(xy).$$

Find the derivative of f at the point $(x_0, y_0) = (2, 0)$ in the direction of $\mathbf{v} = 3i - 4j$.

Solution

We first need to produce a unit vector in the same direction as v :

$$\mathbf{u} = \frac{\mathbf{v}}{|\mathbf{v}|} = \frac{3\mathbf{i} - 4\mathbf{j}}{\sqrt{3^2 + 4^2}} = \frac{3}{5}\mathbf{i} - \frac{4}{5}\mathbf{j}.$$

Firstly

$$f_x = e^y - \sin(xy)\, y \Rightarrow (f_x)_{(2,0)} = e^0 - \sin 0 = 1;$$
$$f_y = xe^y - \sin(xy)\, x \Rightarrow (f_x)_{(2,0)} = 2e^0 - \sin 0 = 2.$$

The gradient of f at $(2, 0)$ is

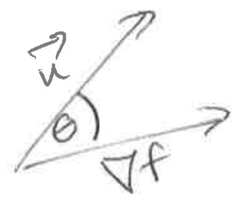
$$(\nabla f)_{(2,0)} = f_x\mathbf{i} + f_y\mathbf{j} = \mathbf{i} + 2\mathbf{j} = (1, 2).$$

Then

$$\begin{aligned} D_{\mathbf{u}}f &= (\nabla f)_{P_0} \cdot \mathbf{u} \\ &= (\mathbf{i} + 2\mathbf{j}) \cdot \left(\frac{3}{5}\mathbf{i} - \frac{4}{5}\mathbf{j}\right) \\ &= \frac{3}{5}\mathbf{i} \cdot \mathbf{i} - \frac{8}{5}\mathbf{j} \cdot \mathbf{j} = -\mathbf{1}. \end{aligned}$$

Remark
If θ is the angle between ∇f and u, then we know from our results on dot products that

$$D_{\mathbf{u}}f = |\nabla f|\,|\mathbf{u}| \cos\theta = |\nabla f| \cos\theta.$$



This gives:

Properties of the Directional Derivative.
(1) The function f increases most rapidly when $\cos\theta = 1$ or $\theta = 0$, and $\mathbf{u}$ is the direction of ∇f. In this case,

$$D_{\mathbf{u}}f = |\nabla f|\,.$$

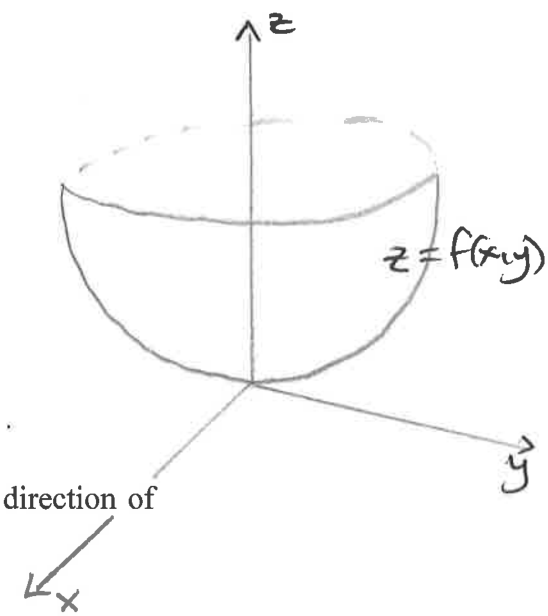
(2) The function f decreases most rapidly when $\cos\theta = -1$ or $\theta = \pi$, and $\mathbf{u}$ is the direction of $-\nabla f$. In this case,

$$D_{\mathbf{u}}f = -|\nabla f|\,.$$

(3) Any direction $\mathbf{u}$ orthogonal to ∇f is a direction of zero change of f, that is

$$D_{\mathbf{u}}f = 0.$$

Example
Find the directions in which $f(x, y) = \frac{x^2}{2} + y^2$
(a) increases most rapidly at $(1, 1)$;
(b) decreases most rapidly at $(1, 1)$;
(c) What are the directions of zero change of f?



Solution
We see that

$$\nabla f = (f_x, f_y) = (x, 2y)\,.$$

In particular at $(1, 1)$,

$$(\nabla f)_{(1,1)} = (1, 2)\,.$$

The direction of ∇f at $(1, 1)$ is

$$\mathbf{u} = \frac{(\nabla f)_{(1,1)}}{\left|(\nabla f)_{(1,1)}\right|} = \frac{(1, 2)}{\sqrt{1+4}} = \left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right).$$

(a) The function increases most rapidly at $(1, 1)$ in the direction of

$u = \left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)$ and the rate of change is

$$\left|(\nabla f)_{(1,1)}\right| = \sqrt{5}.$$

(b) The function decreases most rapidly at $(1,1)$ in the direction of $-u = -\left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)$ and the rate of change is

$$-\left|(\nabla f)_{(1,1)}\right| = -\sqrt{5}.$$

(c) The directions of zero change at $(1,1)$ are the directions orthogonal to $(\nabla f)_{(1,1)} = (1,2)$. One can check that these are

$$\left(-\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right) \text{ and } \left(\frac{2}{\sqrt{5}}, -\frac{1}{\sqrt{5}}\right).$$

Gradients and Tangents to Level Curves

Recall that a level curve of a function $f(x,y)$ is formulated as follows: we choose a constant c, and consider the set of all (x,y) with

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

This is a **level curve** of f. If we assume that

$$x = g(t) \text{ and } y = h(t),$$

so that $r = (g(t), h(t)) = g(t)\,i + h(t)\,k$ is our level curve, then

$$f(g(t), h(t)) = c.$$

We differentiate both sides with respect to t and use the chain rule:

$$\frac{d}{dt} f(g(t), h(t)) = \frac{dc}{dt} = 0,$$

or

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

that is

$$\frac{\partial f}{\partial x}\frac{dg}{dt} + \frac{\partial f}{\partial y}\frac{dh}{dt} = 0.$$

This is the same as

$$\left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) \cdot \left(\frac{dg}{dt}, \frac{dh}{dt}\right) = 0$$

or

$$\nabla f \cdot \frac{d\mathbf{r}}{dt} = \mathbf{0}.$$

Here $\frac{dr}{dt}$ is the tangent to the curve r. Thus ∇f is the normal to the tangent vector, so is a normal to the curve.

Summary

At every point (x_0, y_0) in the domain of a differentiable function $f(x,y)$, the gradient of f is normal to the level curve through (x_0, y_0).

We can use this normal ∇f to find the equation for the tangent line to level curves. The line through a point (x_0, y_0) that is normal to a vector $\mathbf{N} = A\mathbf{i} + B\mathbf{j}$ is

$$A(x - x_0) + B(y - y_0) = 0.$$

Choosing $\mathbf{N} = \nabla f$, so that $A = f_x$ and $B = f_y$, we obtain:

Tangent Line to a Level Curve of $f(x, y)$ at (x_0, y_0) on the curve
This is given by the equation

$$f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) = 0.$$

Example
Find an equation for the tangent to the ellipse

$$\frac{x^2}{4} + y^2 = 2$$

at the point $(-2, 1)$.
Solution
Here $f(x, y) = \frac{x^2}{4} + y^2$ and we are considering the level curve $f(x, y) = 2$, while $(x_0, y_0) = (-2, 1)$. We see that

$$f_x = \frac{1}{2}x \text{ and } f_y = 2y,$$

so that at $(-2, 1)$,

$$f_x = -1 \text{ and } f_y = 2.$$

Then the tangent line to the ellipse is

$$(-1)(x + 2) + 2(y - 1) = 0,$$

or

$$2y - x = 4.$$

Rules for Gradients
Let f, g be differentiable functions of two variables, and k be a constant.
(1) **Sum Rule**

$$\nabla(f + g) = \nabla f + \nabla g.$$

(2) **Difference Rule**

$$\nabla(f - g) = \nabla f - \nabla g.$$

(3) **Pull out a constant**

$$\nabla(kf) = k\nabla f.$$

(4) **Product Rule**

$$\nabla(fg) = f\nabla g + g\nabla f.$$

(5) **Quotient Rule**

$$\nabla\left(\frac{f}{g}\right) = \frac{g\nabla f - f\nabla g}{g^2}.$$

Proof of the Product Rule

Using the usual product rule (which applies to partial derivatives)

$$\begin{aligned}
\nabla(fg) &= \left(\frac{\partial}{\partial x}(fg), \frac{\partial}{\partial y}(fg)\right) \\
&= \left(f\frac{\partial g}{\partial x} + g\frac{\partial f}{\partial x}, f\frac{\partial g}{\partial y} + g\frac{\partial f}{\partial y}\right) \\
&= \left(f\frac{\partial g}{\partial x}, f\frac{\partial g}{\partial y}\right) + \left(g\frac{\partial f}{\partial x}, g\frac{\partial f}{\partial y}\right) \\
&= f\left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}\right) + g\left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) \\
&= f\nabla g + g\nabla f.
\end{aligned}$$

■

Functions of Three Variables
Let $f(x, y, z)$ be a function of three variables. Let $\mathbf{u} = (u_1, u_2, u_3)$ be a unit vector. Then

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

The directional derivative is still

$$\begin{aligned}
D_{\mathbf{u}}f &= \nabla f \cdot \mathbf{u} \\
&= \frac{\partial f}{\partial x}u_1 + \frac{\partial f}{\partial y}u_2 + \frac{\partial f}{\partial z}u_3,
\end{aligned}$$

and if θ is the angle between ∇f and $\mathbf{u}$, then still

$$D_{\mathbf{u}}f = |\nabla f|\cos\theta.$$

Example
Let

$$f(x, y, z) = x^3 - xy^2 - z.$$

(a) Find the derivative of f at $P_0(1, 1, 0)$ in the direction $\mathbf{v} = 2i - 3j + 6k$.
(b) In what directions does f change most rapidly at P_0, and what is the rate of this change?
Solution
(a)

$$f_x = 3x^2 - y^2;\ f_y = -2xy;\ f_z = -1.$$

At $(1, 1, 0)$, these take the values

$$f_x = 3 - 1 = 2;\ f_y = -2;\ f_z = -1.$$

So

$$\nabla f = (2, -2, -1).$$

Next, we find the unit vector $\mathbf{u}$ in the direction $\mathbf{v}$:

$$|\mathbf{v}| = \sqrt{4 + 9 + 36} = 7,$$

so

$$\mathbf{u} = \frac{\mathbf{v}}{|\mathbf{v}|} = \frac{2}{7}\mathbf{i} - \frac{3}{7}\mathbf{j} + \frac{6}{7}\mathbf{k} = \left(\frac{2}{7}, -\frac{3}{7}, \frac{6}{7}\right).$$

Then

$$\begin{aligned} D_{\mathbf{u}}f &= \nabla f \cdot \mathbf{u} \\ &= (2, -2, -1) \cdot \left(\frac{2}{7}, -\frac{3}{7}, \frac{6}{7}\right) \\ &= \frac{2(2) + (-2)(-3) + (-1)6}{7} = \frac{4}{7}. \end{aligned}$$

(b) The function increases most rapidly in the direction of $\nabla f = (2, -2, -1)$ and decreases most rapidly in the direction of $-\nabla f = -(2, -2, -1) = (-2, 2, 1)$. The rate of change in the direction of ∇f is

$$|\nabla f| = \sqrt{4 + 4 + 1} = 3.$$

The rate of change in the direction of $-\nabla f$ is -3.

The Chain Rule for Paths

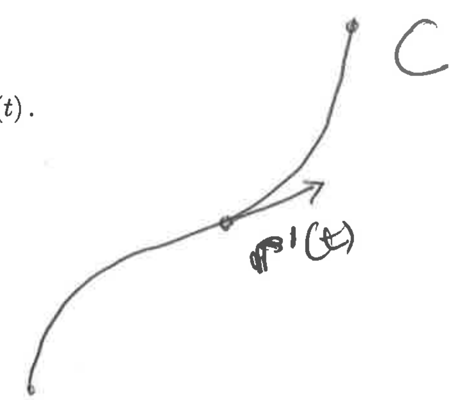
Let $\mathbf{r}(t) = x(t)\,\mathbf{i} + y(t)\,\mathbf{j} + z(t)\,\mathbf{k}$ be a smooth path C and $w = f(\mathbf{r}(t))$ be a scalar function along C, then the chain rule gives

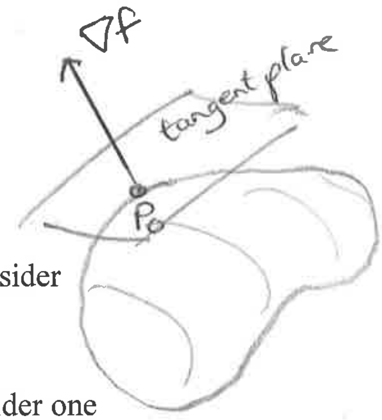
$$\begin{aligned} \frac{dw}{dt} &= \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt} + \frac{\partial w}{\partial z}\frac{dz}{dt} \\ &= f_x\frac{dx}{dt} + f_y\frac{dy}{dt} + f_z\frac{dz}{dt} \\ &= (f_x, f_y, f_z) \cdot \left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt}\right) \\ &= \nabla f \cdot \mathbf{r}'(t). \end{aligned}$$

That is,

$$\frac{d}{dt}(f(\mathbf{r}(t))) = \nabla f \cdot \mathbf{r}'(t).$$

This is the **derivative of f along the path** C.





Section 14.6: Tangent Planes and Differentials

So far we have considered about tangent lines, now we shall consider tangent planes to a surface.

Tangent Planes and Normal Lines

Let $f(x,y,z)$ be a differentiable function of three variables, and consider one of its level surfaces $f(x,y,z)=c$. Let $\mathbf{r}(t)=x(t)\,\mathbf{i}+y(t)\,\mathbf{j}+z(t)\,\mathbf{k}$ be a smooth curve on this level surface $f(x,y,z)=c$. We saw above that

$$\frac{d}{dt}(f(\mathbf{r}(t)))=\nabla f\cdot\mathbf{r}'(t),$$

and since $f(\mathbf{r}(t))=c$, we have

$$\nabla f\cdot\mathbf{r}'(t)=0.$$

That is, the gradient ∇f is orthogonal to the curve's velocity vector $\mathbf{r}'$.

Now let us fix a point P_0 and consider all velocity vectors at P_0. They are all orthogonal to $\nabla f_{|P_0}$. Thus all tangent lines through P_0 have the same normal $(\nabla f)_{P_0}$, so all lie in the plane through P_0 that has normal $\nabla f_{|P_0}$. We now define this plane:

Definition

(a) The **tangent plane** at the point $P_0(x_0,y_0,z_0)$ on the level surface $f(x,y,z)=c$ of a differentiable function f is the plane through P_0 normal to $\nabla f_{|P_0}$.

(b) The **normal line** of the surface at P_0 is the line through P_0 parallel to $\nabla f_{|P_0}$.

Here are the formulae for these:

Tangent plane to $f(x,y,z)=c$ at $P_0(x_0,y_0,z_0)$

This is

$$f_x(P_0)(x-x_0)+f_y(P_0)(y-y_0)+f_z(P_0)(z-z_0)=0.$$

Normal line to $f(x,y,z)=c$ at $P_0(x_0,y_0,z_0)$

This is defined for all $t\in(-\infty,\infty)$ by

$$\begin{aligned}(x,y,z)&=P_0+t\nabla f_{|P_0}\\&=(x_0,y_0,z_0)+t\left(f_x(P_0),f_y(P_0),f_z(P_0)\right).\end{aligned}$$

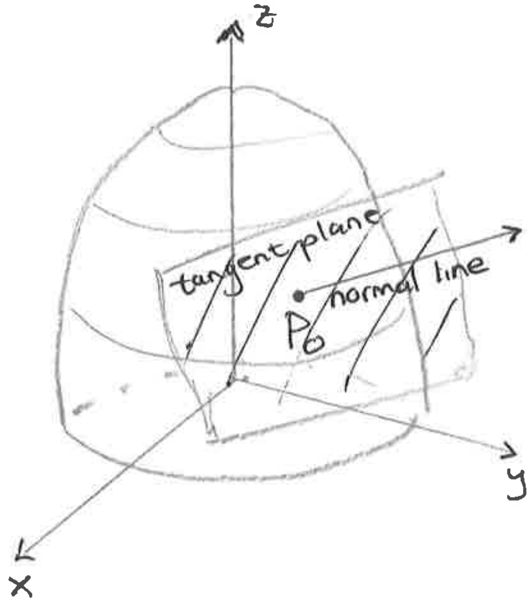
Example

Find the tangent plane and the normal line of the level surface

$$f(x,y,z)=x^2+y^2+z-9=0$$

(a circular paraboloid) at the point $P_0(1,2,4)$.

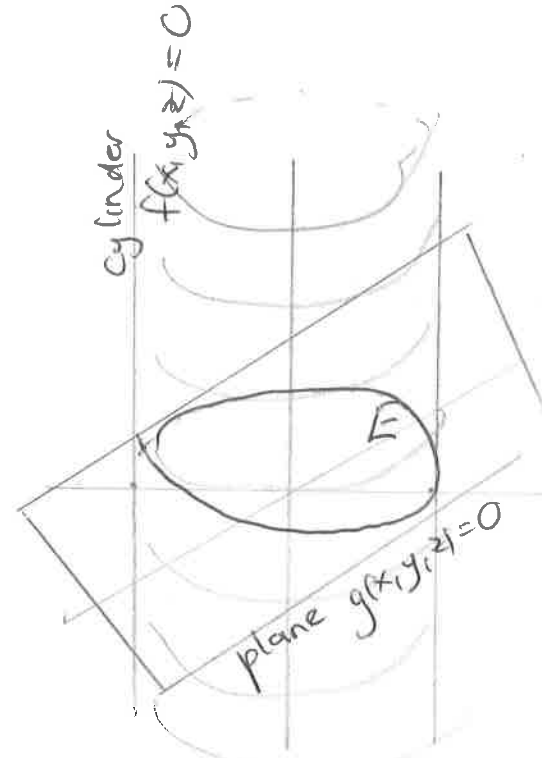
Solution



We see that
$$\nabla f = (2x, 2y, 1)$$
so at P_0,
$$\nabla f_{|P_0} = (2, 4, 1) = 2\mathbf{i} + 4\mathbf{j} + \mathbf{k}.$$
The tangent plan is therefore
$$2(x-1) + 4(y-2) + 1(z-4) = 0,$$
or
$$2x + 4y + z = 14.$$
The normal line to the surface at P_0 is
$$(x, y, z) = (1, 2, 4) + t(2, 4, 1),$$
or equivalently
$$x = 1 + 2t;\ y = 2 + 4t;\ z = 4 + t.$$
What if we don't start with a level surface, but want the tangent plane to a surface at a specific point? The above applies with minor changes. Let $z = f(x, y)$ be a smooth surface, and fix a point $P_0(x_0, y_0, z_0)$. where $z_0 = f(x_0, y_0)$. Let
$$F(x, y, z) = f(x, y) - z.$$
We can think of $z = f(x, y)$ as defining a level surface $F(x, y, z) = 0$, and then apply the theory above to F, a function of three variables:
$$\begin{aligned} F_x &= f_x; \\ F_y &= f_y; \\ F_z &= -1. \end{aligned}$$
The formula for the tangent plane is
$$F_x(P_0)(x - x_0) + F_y(P_0)(y - y_0) + F_z(P_0)(z - z_0) = 0,$$
that is
$$f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) + (-1)(z - z_0) = 0.$$
Let's summarize:

Tangent Plane to a Surface $z = f(x, y)$ at $(x_0, y_0, f(x_0, y_0))$
The plane tangent to the surface $z = f(x, y)$ of a differentiable function f at the point $P_0(x_0, y_0, z_0) = (x_0, y_0, f(x_0, y_0))$ is given by
$$f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) - (z - z_0) = 0.$$

Example
Find the tangent plane to the surface
$$z = x\cos y - ye^x$$
at $(0, 0, 0)$.
Solution
Here
$$f(x, y) = x\cos y - ye^x$$



so

$$f_x = \cos y - ye^x \text{ and } f_y = -x\sin y - e^x.$$

At $(x, y) = (0, 0)$,

$$f_x = \cos 0 - 0 = 1 \text{ and } f_y = -0 - 1 = -1.$$

The tangent plane is therefore

$$1\,(x - 0) + (-1)\,(y - 0) - (z - 0) = 0,$$

or

$$x - y - z = 0.$$

Example
The surfaces

$$f\,(x, y, z) = x^2 + y^2 - 2 = 0 \text{ (a cylinder)}$$

and

$$g\,(x, y, z) = x + z - 4 = 0 \text{ (a plane)}$$

intersect in an ellipse, which we call E. Find parametric equations for the tangent line to E at the point $P_0\,(1, 1, 3)$.
Solution
The ellipse E lies on both level curves $f\,(x, y, z) = 0$ and $g\,(x, y, z) = 0$. Recall that the tangent line to E lies on the tangent planes of both level surfaces, so is orthogonal to ∇f and similarly to ∇g. Then it will be parallel to $\mathbf{v} = \nabla f \times \nabla g$. So let us first compute the latter:

$$\begin{aligned} \nabla f &= (2x, 2y, 0)\,; \\ \nabla g &= (1, 0, 1). \end{aligned}$$

Then at the point $(1, 1, 3)$,

$$\begin{aligned} \nabla f &= (2, 2, 0) = 2\mathbf{i}+2\mathbf{j}; \\ \nabla g &= (1, 0, 1) = \mathbf{i} + \mathbf{k}. \end{aligned}$$

So

$$\begin{aligned} \mathbf{v} &= (2\mathbf{i}+2\mathbf{j}) \times (\mathbf{i} + \mathbf{k}) \\ &= \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix} \\ &= \mathbf{i}\,(2 - 0) - \mathbf{j}\,(2 - 0) + \mathbf{k}\,(0 - 2) \\ &= 2\mathbf{i}-2\mathbf{j}-2\mathbf{k} = (2, -2, -2)\,. \end{aligned}$$

The tangent line will pass through $(1, 1, 3)$ and have direction $2\mathbf{i} - 2\mathbf{j} - 2\mathbf{k}$, so is given by

$$(1, 1, 3) + t\,(2, -2, -2)\,.$$

In parametric form this gives

$$\begin{aligned} x &= 1+2t; \\ y &= 1-2t; \\ z &= 3-2t. \end{aligned}$$

Estimating Change in a Specific Direction

If we want to estimate how much a function $f(x)$ of one variable changes when we move a small distance ds from a point P_0, we can use

$$df = f'(P_0)\, ds.$$

Suppose now that f is a function of two or more variables, and we want to estimate how much f changes in the direction of a unit vector $\mathbf{u}$. We use the formula

$$df = \left(\nabla f_{|P_0} \cdot \mathbf{u}\right) ds$$

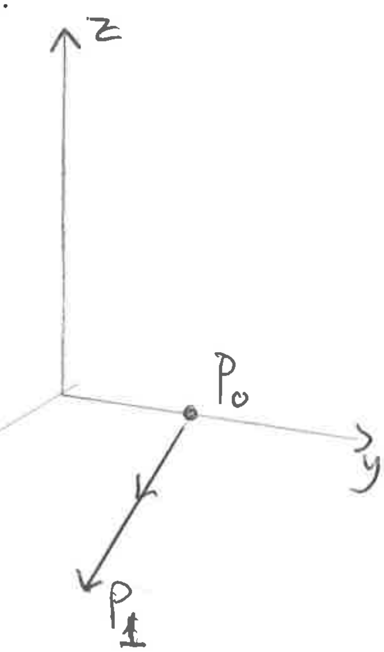
Example

Let

$$f(x,y,z) = y \sin x + 2yz.$$

Estimate how much f will change if $P(x,y,z)$ moves 0.1 unit from the point $P_0(0,1,0)$ in the direction of the line segment from P_0 to $P_1(2,2,-2)$.

Solution



$$\nabla f = (y\cos x, \sin x + 2z, 2y)$$

so at $P_0(0,1,0)$,

$$\nabla f = (1\cos 0, \sin 0 + 0, 2) = (1,0,2).$$

The vector from P_0 to P_1 is given by

$$\overrightarrow{P_0P_1} = (2,2,-2) - (0,1,0) = (2,1,-2),$$

and

$$\left|\overrightarrow{P_0P_1}\right| = \sqrt{4+1+4} = 3,$$

and the unit vector in the direction of $\overrightarrow{P_0P_1}$ is

$$\mathbf{u} = \frac{\overrightarrow{P_0P_1}}{\left|\overrightarrow{P_0P_1}\right|} = \frac{1}{3}(2,1,-2) = \left(\frac{2}{3}, \frac{1}{3}, -\frac{2}{3}\right) = \frac{2}{3}\mathbf{i} + \frac{1}{3}\mathbf{j} - \frac{2}{3}\mathbf{k}.$$

Then

$$\nabla f_{|P_0} \cdot \mathbf{u} = (1,0,2) \cdot \left(\frac{2}{3}, \frac{1}{3}, -\frac{2}{3}\right) = \frac{2}{3} + 0 - \frac{4}{3} = -\frac{2}{3}.$$

so

$$df = \left(\nabla f_{|P_0} \cdot \mathbf{u}\right) ds = \left(-\frac{2}{3}\right)(0.1) = -0.0666\ldots \text{ units.}$$

How to Linearize a Function of two Variables

If we move a small distance from a given point, how can we estimate the function values $f(x, y)$ at the new point? To do this, we use linearization. Suppose (x_0, y_0) is given, and we move to a new point (x, y). Let

$$\Delta x = x - x_0 \text{ and } \Delta y = y - y_0,$$

and

$$\Delta z = f(x, y) - f(x_0, y_0).$$

From the definition of differentiability,

$$\Delta z = f_x(x_0, y_0)\Delta x + f_y(x_0, y_0)\Delta y + \varepsilon_1 \Delta x + \varepsilon_2 \Delta y$$

in which each of $\varepsilon_1, \varepsilon_2 \to 0$ as both $\Delta x, \Delta y \to 0$. If Δx and Δy are small, then $\varepsilon_1 \Delta x + \varepsilon_2 \Delta y$ will be even smaller. Thus approximately,

$$\begin{aligned} f(x, y) - f(x_0, y_0) &\approx f_x(x_0, y_0)\Delta x + f_y(x_0, y_0)\Delta y \\ &= f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0), \end{aligned}$$

so defining

$$L(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0),$$

we have

$$f(x, y) \approx L(x, y).$$

Hence:

Definition
The **linearization of a function** $f(x, y)$ at a point (x_0, y_0) where f is differentiable, is the function

$$L(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0).$$

The **standard linear approximation** is

$$f(x, y) \approx L(x, y).$$

Example
Find the linearization of

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

at the point $(3, 2)$.
Solution
We need the values of the function and its partial derivatives:

$$f(3, 2) = 9 - 6 + 2 + 3 = 8;$$

$$f_x = 2x - y \text{ and } f_y = -x + y$$

so

$$f_x(3, 2) = 6 - 2 = 4 \text{ and } f_x(3, 2) = -3 + 2 = -1.$$

Then

$$\begin{aligned} L(x, y) &= 8 + 4(x - 3) + (-1)(y - 2) \\ &= 4x - y - 2. \end{aligned}$$

How good an approximation is the linearization?

The Error in the Standard Linear Approximation
Suppose that f has continuous first and second derivatives in an open set containing a rectangle R centered at (x_0, y_0). Let M be an upper bound for $|f_{xx}|$, $|f_{yy}|$, and $|f_{xy}|$ in R. Then the error $E(x,y) = f(x,y) - L(x,y)$ in replacing $f(x,y)$ by its linearization $L(x,y)$ satisfies

$$|E(x,y)| \leq \frac{M}{2}\left(|x - x_0| + |y - y_0|\right)^2.$$

Differentials
If $f(x,y)$ is a function of two variables, we can think of dx and dy as **differentials.**

Definition
If we move from (x_0, y_0) to a point $(x_0 + dx, y_0 + dy)$ nearby, the resulting change

$$df = f_x(x_0, y_0)\,dx + f_y(x_0, y_0)\,dy$$

in the linearization of f is called the **total differential** of f.

Example
Suppose that a cylindrical can is supposed to have a radius r of 1 inch and a height h of 5 inches. Suppose, however, that the radius and height are off by amounts of $dr = 0.03$ and $dh = -0.1$. Estimate the resulting absolute change in the volume of the can.
Solution
The volume of a cylinder of radius r and height h is

$$V = \pi r^2 h.$$

To estimate the change, we use the total differential

$$\Delta V \approx dV = V_r(r_0, h_0)\,dr + V_h(r_0, h_0)\,dh.$$

In our case $(r_0, h_0) = (1, 5)$ and $dr = 0.03$, while $dh = -0.1$ We see that

$$V_r = 2\pi r h \text{ and } V_h = \pi r^2.$$

At $(r_0, h_0) = (1, 5)$,

$$V_r = 10\pi \text{and } V_h = \pi.$$

Then

$$\begin{aligned} \Delta V &\approx dV = (10\pi)(0.03) + (\pi)(-0.1) \\ &= 0.2\pi. \end{aligned}$$

Example
A company produces storage tanks that are 25 feet high and have a radius of 5 feet. How sensitive are the volumes to small variations in height and

radius?

Solution

Again, $V = \pi r^2 h$, and

$$\begin{aligned} \Delta V &\approx dV = V_r(r_0, h_0)\,dr + V_h(r_0, h_0)\,dh \\ &= (2\pi r_0 h_0)\,dr + \left(\pi r_0^2\right) dh. \end{aligned}$$

In this case, $(r_0, h_0) = (5, 25)$, so

$$\Delta V \approx dV = (250\pi)\,dr + (25\pi)\,dh.$$

Functions of More than Two Variables

One can define a linearization $L(x, y, z)$ of a function $f(x, y, z)$ of three variables:

$$L(x, y, z) = f(P_0) + f_x(P_0)(x - x_0) + f_y(P_0)(y - y_0) + f_z(P_0)(z - z_0).$$

The total differential is now

$$df = f_x(P_0)\,\Delta x + f_y(P_0)\,\Delta y + f_z(P_0)\,\Delta z.$$

Section 14.7: Extreme Values and Saddle Points

You have seen that finding maxima and minima of functions of one variable can be useful in practical problems. Now we do this for functions of more than one variable, First, let us define what we mean by a maximum or minimum:

Definitions
Let $f(x,y)$ be a function defined on a region R containing the point (a,b).
(1) $f(a,b)$ is a **local maximum value** of f if $f(a,b) \geq f(x,y)$ for all domain points (x,y) that lie in an open disk centered on (a,b).
(2) $f(a,b)$ is a **local minimum value** of f if $f(a,b) \leq f(x,y)$ for all domain points (x,y) that lie in an open disk centered on (a,b).
(3) A local minimum or maximum is called a **local extremum** or **relative extremum**.

Remark
A local maximum is like a mountain peak, while a local minimum is like the bottom of a valley. At such points (at least for smooth functions), the tangent plane is horizontal, so that $\nabla f = 0$. This leads to:

Theorem 10: First Derivative Test for Local Extreme Values
If $f(x,y)$ has a local maximum or local minimum at an interior point (a,b) of its domain, and if the first partial derivatives exists there, then

$$f_x(a,b) = 0 \text{ and } f_y(a,b) = 0.$$

Proof
Suppose f has a local extremum at (a,b). Then the function of one variable

$$g(x) = f(x,b)$$

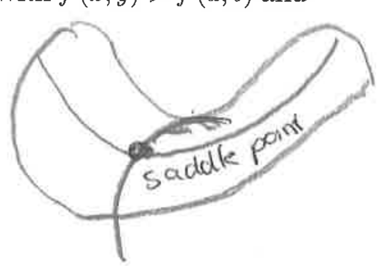
is a local extremum of a differentiable function of one variable, so $g'(a) = 0$, that is $f_x(a,b) = 0$. Similarly for $f_y(a,b)$. ■

We give a special name to points where $f_x = 0$ and $f_y = 0$:

Definition
An interior point (a,b) of the domain of a function $f(x,y)$ is called a **critical point** of f if at (a,b), either
(a) Both $f_x = 0$ and $f_y = 0$
(b) Or one or both of f_x and f_y do not exist.

One example of a critical point that is not a local maximum or minimum, is a saddle point, where the graph of the function looks like a saddle:

Definition
A differentiable function $f(x,y)$ has a **saddle point** at a critical point (a,b) if in every open disk center (a,b), there are points (x,y) with $f(x,y) > f(a,b)$ and



other points with $f(x,y) < f(a,b)$. The point $(a,b,f(a,b))$ is called a **saddle point on the surface** $z = f(x,y)$.

Example
Find the local extreme values of $f(x,y) = x^2 + y^2 - 4y + 9$.
Solution
Note that $f(x,y)$ is defined for all (x,y) in the plane. Now let us compute the critical points: we see that

$$f_x = 2x \text{ and } f_y = 2y - 4$$

so we solve

$$\begin{aligned} f_x &= 2x = 0 \Rightarrow x = 0; \\ f_y &= 2y - 4 = 0 \Rightarrow y = 2. \end{aligned}$$

So there is only one critical point, $(0,2)$ and we see that $f(0,2) = 0+4-8+9 = 5$. By completing the square, we see that

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

and hence for all (x,y),

$$f(x,y) \geq 5 = f(0,2).$$

Thus $(0,2)$ gives a local minimum (and in fact a global minimum).

Example
Find the local extreme values of $f(x,y) = x^2 - y^2$.
Solution
Again, the domain of f is all (x,y) in the plane. We see that

$$f_x = 2x \text{ and } f_y = -2y$$

so $f_x = 0$ and $f_y = 0$ occurs only at $(0,0)$. This is the only critical point. In fact it is a saddle point, for if $x \neq 0$

$$f(x,0) = x^2 > 0,$$

while if $y \neq 0$

$$f(0,y) = -y^2 < 0,$$

so in every open disk center $(0,0)$ that are points (x,y) where $f(x,y) > 0 = f(0,0)$ and also points (x,y) where $f(x,y) < 0 = f(0,0)$.

So we see from this example that $f_x = 0$ and $f_y = 0$ is not enough to have a local minimum or maximum. Here is what we need:

Theorem 11: Sufficient conditions for local extrema
Suppose that $f(x,y)$ and its second partial derivatives are continuous throughout a disk centered on (a,b) and that $f_x(a,b) = 0$ and $f_y(a,b) = 0$.
(I) If at (a,b), $f_{xx} < 0$ and $f_{xx}f_{yy} - f_{xy}^2 > 0$, then f has a local maximum at (a,b).
(II) If at (a,b), $f_{xx} > 0$ and $f_{xx}f_{yy} - f_{xy}^2 > 0$, then f has a local minimum at (a,b).

(III) If at (a,b), $f_{xx}f_{yy} - f_{xy}^2 < 0$, then f has a saddle point at (a,b).
(The test is inconclusive if at (a,b), $f_{xx}f_{yy} - f_{xy}^2 = 0$).

Remarks
(a) We shall prove/motivate Theorem 11 in Section 4.9.
(b) The expression

$$f_{xx}f_{yy} - f_{xy}^2 = \det \begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix}$$

is called the discriminant or Hessian of f.

Example
Find the local extreme values of the function

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

Solution
First, the domain of f is all (x,y). So we compute

$$f_x = y - 2x - 2 \text{ and } f_y = x - 2y - 2.$$

Thus we have simultaneous equations

$$\begin{aligned} -2x + y &= 2; \\ x - 2y &= 2. \end{aligned}$$

Solving these, we find the unique solution $(x,y) = (-2,-2)$. So this is the only critical point. Now let us check the second derivatives.

$$f_{xx} = -2 \text{ and } f_{yy} = -2 \text{ and } f_{xy} = 1,$$

so $f_{xx} < 0$ and the discriminant is

$$f_{xx}f_{yy} - f_{xy}^2 = (-2)(-2) - 1 = 3 > 0.$$

Thus at $(-2,-2)$ there is a local maximum of f. (The value of this local maximum is 8).

Example
Find the local extreme values of

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

Solution
First, the domain of f is all (x,y). So we compute

$$f_x = -6x + 6y \text{ and } f_y = 6y - 6y^2 + 6x.$$

So we solve

$$f_x = -6x + 6y = 0 \Rightarrow y = x$$

and

$$\begin{aligned} f_y &= 6y - 6y^2 + 6x = 0 \\ &\Rightarrow 6x - 6x^2 + 6x = 0 \\ &\Rightarrow 2x - x^2 = 0 \\ &\Rightarrow (2 - x)\, x = 0 \\ &\Rightarrow x = 0 \text{ or } x = 2. \end{aligned}$$

So the critical points ar $(0,0)$ and $(2,2)$. Next, we calculate the second derivatives:

$$f_{xx} = -6 \text{ and } f_{yy} = 6 - 12y \text{ and } f_{xy} = 6.$$

The discriminant is

$$\begin{aligned} f_{xx} f_{yy} - f_{xy}^2 &= (-6)(6 - 12y) - 6^2 \\ &= -36 + 72y - 36 = 72\,(y - 1)\,. \end{aligned}$$

Let us do the points separately:
First, $(0,0)$
Here $f_{xx} = -6 < 0$ and $f_{xx} f_{yy} - f_{xy}^2 = -72 < 0$, so $(0,0)$ is a saddle point.
Second $(2,2)$
Here $f_{xx} = -6 < 0$ and $f_{xx} f_{yy} - f_{xy}^2 = 72 > 0$, so $(2,2)$ is a local maximum (of value $f\,(2,2) = 8$).

Example
Find the critical points and extrema of

$$f\,(x,y) = 10xye^{-(x^2+y^2)}.$$

Solution
Again, the function f is defined for all x, y, so there are no boundary points to the domain. We see that

$$\begin{aligned} f_x &= 10ye^{-(x^2+y^2)} + 10xye^{-(x^2+y^2)}\,(-2x) \\ &= 10ye^{-(x^2+y^2)}\left(1 - 2x^2\right); \end{aligned}$$

Also,

$$\begin{aligned} f_y &= 10xe^{-(x^2+y^2)} + 10xye^{-(x^2+y^2)}\,(-2y) \\ &= 10xe^{-(x^2+y^2)}\left(1 - 2y^2\right). \end{aligned}$$

So

$$f_x = 0 \Rightarrow y = 0 \text{ or } x = \pm\frac{1}{\sqrt{2}};$$
$$f_y = 0 \Rightarrow x = 0 \text{ or } y = \pm\frac{1}{\sqrt{2}}.$$

Since we need both $f_x = 0$ and $f_y = 0$, we see that when $y = 0$, then also $x = 0$. Thus these are the critical points, five in all:

$$(0,0)\,, \left(\pm\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}}\right).$$

Next, calculate the second derivatives:

$$\begin{aligned} f_{xx} &= 10ye^{-(x^2+y^2)}\left(1-2x^2\right)(-2x)+10ye^{-(x^2+y^2)}(-4x) \\ &= -20xye^{-(x^2+y^2)}\left(1-2x^2+2\right) \\ &= -20xye^{-(x^2+y^2)}\left(3-2x^2\right). \end{aligned}$$

Similarly,

$$f_{yy} = -20xye^{-(x^2+y^2)}\left(3-2y^2\right).$$

Also

$$\begin{aligned} f_{xy} &= 10e^{-(x^2+y^2)}\left(1-2x^2\right)+10ye^{-(x^2+y^2)}\left(1-2x^2\right)(-2y) \\ &= 10e^{-(x^2+y^2)}\left(1-2x^2\right)\left(1-2y^2\right). \end{aligned}$$

Then the discriminant is

$$\begin{aligned} & f_{xx}f_{yy}-f_{xy}^2 \\ = & \left(20xye^{-(x^2+y^2)}\right)^2\left(3-2x^2\right)\left(3-2y^2\right) \\ & -\left(10e^{-(x^2+y^2)}\left(1-2x^2\right)\left(1-2y^2\right)\right)^2. \end{aligned}$$

Let's analyze it all in a table:

Critical point	f_{xx}	f_{yy}	f_{xy}	Discriminant	Analysis
$(0,0)$	0	0	10	-100	Saddle Point
$\left(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right)$	$-20e^{-1}$	$-20e^{-1}$	0	$400e^{-2}$	Local Maximum
$\left(-\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right)$	$20e^{-1}$	$20e^{-1}$	0	$400e^{-2}$	Local Minimum
$\left(\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}}\right)$	$20e^{-1}$	$20e^{-1}$	0	$400e^{-2}$	Local Minimum
$\left(-\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}}\right)$	$-20e^{-1}$	$-20e^{-1}$	0	$400e^{-2}$	Local Maximum

Remarks

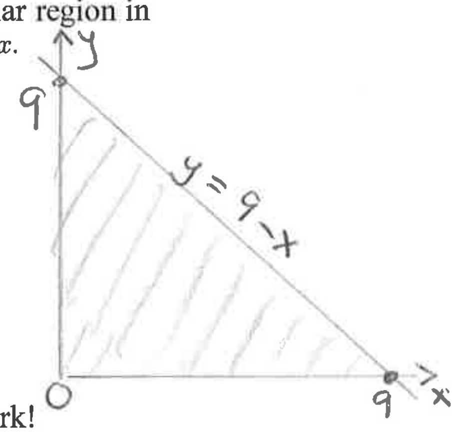
All the above is for finding local maxima/minima/saddle points. What if we want to find a global maximum or minimum of a function f in a region R? That is, we want to find the points where f is the largest or the smallest in all of R? Let us suppose that R is a closed bounded region. Then we know the global maximum and minimum are attained. If these are inside the interior of R, then they are local extrema also, and so will be critical points. If they are on the boundary, then it will be different. This suggests a possibly messy procedure:

(I) Find the critical points of f in the interior of R. Evaluate f at these points.

(II) List the boundary points of f where f has local maxima or minima (when f is regarded as a function on the boundary only).

(III) Look through the lists of values from (I) and (II) and see where f is largest and smallest.

Example
Let
$$f(x,y) = 2 + 2x + 4y - x^2 - y^2.$$
Find the absolute maximum and minimum of f on the triangular region in the first quadrant bounded by the lines $x = 0, y = 0$, and $y = 9 - x$.



Solution
First we examine interior points:
(I) **Interior Points**
Now
$$f_x = 2 - 2x \text{ and } f_y = 4 - 2y$$
so $f_x = 0$ and $f_y = 0$ imply
$$(x, y) = (1, 2).$$
This is the only critical point in the interior. We see that
$$f(1,2) = 2 + 2 + 8 - 1 - 4 = 7.$$
(II) **Boundary Points**
We look at each of the three bounding lines. This is a lot of work!
(a) For $y = 0$, $0 \leq x \leq 9$, and
$$f(x, 0) = 2 + 2x - x^2.$$
We see that
$$\frac{d}{dx}(f(x,0)) = 2 - 2x = 0 \text{ when } x = 1,$$
and 1 is inside $(0, 9)$, and
$$f(1,0) = 2 + 2 - 1 = 3.$$
Also we have to consider the endpoints $(0,0)$ and $(9,0)$:
$$\begin{aligned} f(0,0) &= 2; \\ f(9,0) &= -61. \end{aligned}$$
(b) For $x = 0$, $0 \leq y \leq 9$, and
$$f(0, y) = 2 + 4y - y^2.$$
We see that
$$\frac{d}{dy}(f(0,y)) = 4 - 2y = 0 \text{ when } y = 2,$$
and 2 is inside $(0, 9)$ and
$$f(0,2) = 2 + 8 - 4 = 6.$$
We already know $f(0,0) = 2$; the other endpoint is
$$f(0,9) = 2 + 36 - 81 = -43.$$
(c) For $y = 9 - x$, we have
$$\begin{aligned} f(x, 9 - x) &= 2 + 2x + 4(9 - x) - x^2 - (9 - x)^2 \\ &= 2 + 2x + 36 - 4x - x^2 - 81 + 18x - x^2 \\ &= -2x^2 + 16x - 43. \end{aligned}$$

Note too that $0 \leq x \leq 9$. We see that

$$\frac{d}{dx}\left(f\left(x, 9-x\right)\right) = -4x + 16 = 0 \Rightarrow x = 4.$$

This is inside $(0, 9)$, and

$$f\left(4, 5\right) = -32 + 64 - 43 = -11.$$

We already have done the endpoints $(0, 9)$ and $(9, 0)$.

Finally, let us draw up a table:

Point	Function value f
$(1, 2)$	7
$(1, 0)$	3
$(0, 0)$	2
$(9, 0)$	-61
$(0, 2)$	6
$(0, 9)$	-43
$(4, 5)$	-11

Thus we see the global minimum of -61 occurs at $(9, 0)$, and the global maximum of 7 occurs at (1.2).

Section 14.8: Lagrange Multipliers

Constrained Maxima and Minima

Sometimes we want to find the minimum of a function subject to some extra condition or constraint.

Example
Find the points on the hyperbolic cylinder $x^2 - z^2 - 1 = 0$ that are closest to the origin.
Solution
We measure distance from the origin using the square of the distance function:

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

Thus we want to minimize the function $f(x, y, z)$ subject to $x^2 - z^2 = 1 = 0$. One way to do this is to solve for z^2 from the constraint, and then get rid of z^2 in f:

$$z^2 = x^2 - 1$$

and so

$$\begin{aligned} f(x, y, z) &= x^2 + y^2 + z^2 \\ &= 2x^2 - 1 + y^2 = h(x, y), \text{ say.} \end{aligned}$$

We now look for the points that minimize this, by finding critical points:

$$h_x = 4x \text{ and } h_y = 2y.$$

Solve for $h_x = 0$ and $h_y = 0$:

$$4x = 0 \Rightarrow x = 0 \text{ and } 2y = 0 \Rightarrow y = 0.$$

So there is only one critical point $(0, 0)$ of h. But there is no z with $(0, 0, z)$ on the cylinder $x^2 - z^2 = 1$, that is $0 - z^2 = 1$.

Why the problem? The first derivative test found a critical point in the domain of f. However, we want points on the cylinder, where h has a minimum value. Remarkably, if we instead solve for x and repeat this process above, it works.

The Method of Lagrange Multipliers
There is another way that works, and it involves **Lagrange multipliers**. This method says that the local extreme values of a function $f(x, y, z)$ whose variables are subject to a constraint $g(x, y, z) = 0$ are to be found on the surface $g = 0$ among the points where

$$\nabla f = \lambda \nabla g,$$

for some scalar λ. We call λ a **Lagrange multiplier**. The theoretical basis is given in:

Theorem 12: The Orthogonal Gradient Theorem
Suppose that $f(x,y,z)$ is differentiable in a region whose interior contains a smooth curve C, given by

$$\mathbf{r}(t) = x(t)\,\mathbf{i} + y(t)\,\mathbf{j} + z(t)\,\mathbf{k}.$$

If P_0 is a point on C where f has a local maximum or minimum relative to its values on C, then ∇f is orthogonal to C at P_0.

Proof
We show that ∇f is orthogonal to the tangent vector, that is $\nabla f \cdot \mathbf{r}'(t) = 0$. To see this, note that on C, the values of f are given by $f(\mathbf{r}(t)) = f(x(t), y(t), g(t))$. As f has a local min or max at $P_0 = \mathbf{r}(t_0)$, say, so as a function of one variable, we must have

$$\frac{d}{dt} f(\mathbf{r}(t)) = 0 \text{ at } t = t_0.$$

But using the chain rule,

$$\frac{d}{dt} f(\mathbf{r}(t)) = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt} = \nabla f \cdot \mathbf{r}'(t).$$

So at t_0,

$$\nabla f_{|P_0} \cdot \mathbf{r}'(t_0) = 0.$$

■

For functions of two variables, we have

Corollary
At the points on a smooth curve $\mathbf{r}(t) = x(t)\,\mathbf{i} + y(t)\,\mathbf{j}$, where a differentiable function $f(x,y)$ has a local maximum and minimum relative to the curve, we have $\nabla f \cdot \mathbf{r}' = 0$.

Now we motivate the use of Lagrange multipliers. Suppose that $f(x,y,z)$ and $g(x,y,z)$ are differentiable, and that P_0 is a point on the surface $g(x,y,z) = 0$ where f has a local max or min relative to its other values on the surface. We assume also that $\nabla g \neq 0$ at points on the surface $g(x,y,z) = 0$. Then f takes on a local maximum or minimum at P_0 relative to its values on every differentiable curve through P_0. Therefore, by Theorem 12, ∇f is orthogonal to the tangent vector of every such differentiable curve through P_0. So is ∇g, because it is orthogonal to the surface $g = 0$, as we showed in Section 14.5. Therefore, ∇f and ∇g are either parallel or at 180° one another, so that for some scalar λ,

$$\nabla f = \lambda \nabla g.$$

Let us state this formally:

The Method of Lagrange Multipliers
Suppose that $f(x,y,z)$ and $g(x,y,z)$ are differentiable and $\nabla g \neq 0$ whenever $g(x,y,z) = 0$. To find local maxima and minima of f subject to the constraint

$g(x, y, z) = 0$, we find the values of x, y, z, λ that satisfy

$$\nabla f = \lambda \nabla g \text{ and } g(x, y, z) = 0.$$

Remarks

(a) For functions of two variables, the result is similar, we just omit the variable z.

(b) Note that this method does not always give a min or max - an extreme value may not exist.

Example

Find the greatest and smallest values that the function $f(x, y) = xy$ takes on the ellipse

$$\frac{x^2}{8} + \frac{y^2}{2} = 1.$$

Solution

We must find extreme values of $f(x, y) = xy$ subject to

$$g(x, y) = \frac{x^2}{8} + \frac{y^2}{2} - 1 = 0.$$

So we solve the equation

$$\nabla f = \lambda \nabla g,$$

that is

$$(y, x) = \lambda \left(\frac{x}{4}, y\right)$$

or if you prefer,

$$y\mathbf{i} + x\mathbf{j} = \lambda \frac{x}{4}\mathbf{i} + \lambda y\mathbf{j}$$

so

$$y = \lambda \frac{x}{4} \text{ and } x = \lambda y.$$

Then

$$y = \lambda \frac{\lambda y}{4} = \frac{\lambda^2}{4} y,$$

so $y = 0$ or $\lambda = \pm 2$. We now examine these cases:

Case 1: $y = 0$

Then also $x = 0$, and our point is $(0, 0)$. However, $(0, 0)$ is not on the ellipse.

Case 2: $y \neq 0$ and $\lambda = \pm 2$

Then $x = \pm 2y$. Substituting into our constraint equation $g(x, y) = 0$ gives

$$\begin{aligned} \frac{(\pm 2y)^2}{8} + \frac{y^2}{2} - 1 &= 0 \\ &\Rightarrow \frac{1}{2}y^2 + \frac{1}{2}y^2 = 1 \\ &\Rightarrow y = \pm 1 \Rightarrow x = \pm 2 \end{aligned}$$

The function $f(x, y) = xy$ therefore takes on its extreme values at the four points $(\pm 2, 1)$ and $(\pm 2, -1)$. The extreme values are the maximum $xy = 2$ and minimum $xy = -2$.

Example
Find the maximum and minimum of the function $f(x, y) = 3x + 4y$ on the circle $x^2 + y^2 = 1$.
Solution
So we want to find extrema of $f(x, y) = 3x + 4y$ subject to the constraint

$$g(x, y) = x^2 + y^2 - 1 = 0.$$

So as usual, we solve

$$\nabla f = \lambda \nabla g,$$

or

$$(3, 4) = \lambda (2x, 2y)$$

(or

$$3\mathbf{i} + 4\mathbf{j} = 2\lambda x\mathbf{i} + 2\lambda y\mathbf{j}).$$

Thus

$$3 = 2\lambda x \text{ and } 4 = 2\lambda y,$$

which forces $\lambda \neq 0$. So

$$(x, y) = \left(\frac{3}{2\lambda}, \frac{2}{\lambda}\right).$$

Now we substitute in the constraint $g(x, y) = 0$:

$$\begin{aligned} 0 &= x^2 + y^2 - 1 = \left(\frac{3}{2\lambda}\right)^2 + \left(\frac{2}{\lambda}\right)^2 - 1 \\ &= \frac{1}{\lambda^2}\left(\frac{9}{4} + 4\right) - 1 \\ &= \frac{1}{\lambda^2}\frac{25}{4} - 1 \end{aligned}$$

so

$$\left(\frac{5}{2\lambda}\right)^2 = 1$$

so

$$\lambda = \pm\frac{5}{2}.$$

Then either

$$(x, y) = \left(\frac{3}{2\lambda}, \frac{2}{\lambda}\right) = \pm\left(\frac{3}{5}, \frac{4}{5}\right).$$

Thus $f(x, y) = 3x + 4y$ has extreme values at $\pm\left(\frac{3}{5}, \frac{4}{5}\right)$. When $(x, y) = \left(\frac{3}{5}, \frac{4}{5}\right)$, we see that $f(x, y) = 3\left(\frac{3}{5}\right) + 4\left(\frac{4}{5}\right) = 5$, which is a maximum, and when $(x, y) = -\left(\frac{3}{5}, \frac{4}{5}\right)$, we see that $f(x, y) = 3\left(-\frac{3}{5}\right) + 4\left(-\frac{4}{5}\right) = -5$, which is a minimum.

Lagrange Multipliers with Two Constraints
Suppose we want to maximize or minimize a function $f(x, y, z)$ subject to **two constraints:**

$$g_1(x, y, z) = 0 \text{ and } g_2(x, y, z) = 0.$$

To make the problem meaningful, we need that ∇g_1 and ∇g_2 are not parallel, so that the constraints are independent. Rather than one, we need two Lagrange multipliers:

To find the minima and maxima of f subject to the two constraints above, we locate the points by finding the values of x, y, z, λ, μ that satisfy the three equations

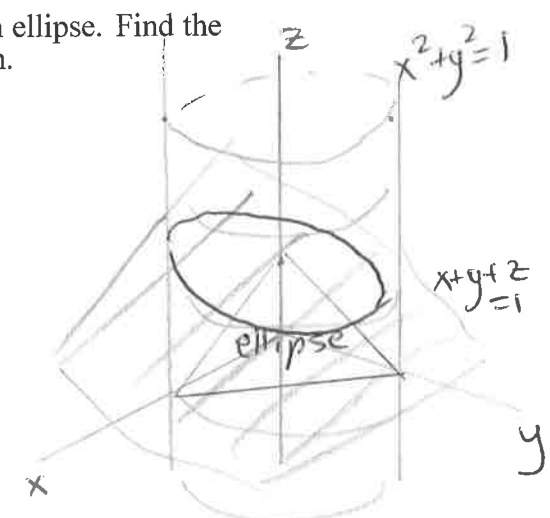
$$\nabla f = \lambda \nabla g_1 + \mu \nabla g_2 \text{ and } g_1(x, y, z) = 0 \text{ and } g_2(x, y, z) = 0.$$

Motivation

Note that the surfaces $g_1 = 0$ and $g_2 = 0$ intersect in a smooth curve C. Along this curve, we seek the points where f has local maximum and minimum values relative to its other values on the curve. These are points where ∇f is normal to C. Now ∇g_1 and ∇g_2 are also normal to C at these points because C lies in the surfaces $g_1 = 0$ and $g_2 = 0$. Thus ∇f lies in the plane determined by ∇g_1 and ∇g_2, so for some scalars λ and μ, $\nabla f = \lambda \nabla g_1 + \mu \nabla g_2$.

Example

The plane $x + y + z = 1$ cuts the cylinder $x^2 + y^2 = 1$ in an ellipse. Find the points on this ellipse closest to and further from the origin.



Solution

We want to maximize or minimize

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

subject to

$$g_1(x, y, z) = x^2 + y^2 - 1 = 0$$

and

$$g_2(x, y, z) = x + y + z - 1 = 0.$$

The equation

$$\nabla f = \lambda \nabla g_1 + \mu \nabla g_2$$

becomes

$$(2x, 2y, 2z) = \lambda (2x, 2y, 0) + \mu (1, 1, 1)$$

so

$$\begin{aligned} 2x &= \lambda 2x + \mu; \\ 2y &= \lambda 2y + \mu; \\ 2z &= \mu. \end{aligned}$$

Thus

$$\begin{aligned} 2x(1 - \lambda) &= \mu = 2z; \\ 2y(1 - \lambda) &= \mu = 2z. \end{aligned}$$

So

$$\begin{aligned} x(1 - \lambda) &= z; \\ y(1 - \lambda) &= z. \end{aligned}$$

Then either $\lambda = 1$ and $z = 0$ or $\lambda \neq 1$ and $x = y = z/(1 - \lambda)$.

Case I: $z = 0$
Then our constraints $g_1 = 0 = g_2$ give

$$x^2 + y^2 = 1 \text{ and } x + y = 1$$

Solving for $y = 1 - x$, we have

$$\begin{aligned} x^2 + (1-x)^2 &= 1 \\ &\Rightarrow 2x^2 - 2x + 1 = 1 \\ &\Rightarrow 2x(x-1) = 0 \\ &\Rightarrow x = 0 \text{ or } x = 1. \end{aligned}$$

Thus we obtain the points $(x, y, z) = (0, 1, 0)$ or $(1, 0, 0)$. These are both at a distance 1 to the origin. Next, let's look at the other points:
Case II: $z \neq 0$
Then $x = y$ and our first constraint gives

$$2x^2 = 1 \Rightarrow x = \pm \frac{1}{\sqrt{2}}.$$

Our second constraint gives

$$2x + z = 1 \Rightarrow z = 1 - 2x = 1 \mp \sqrt{2}.$$

Thus our points are $(x, y, z) = P_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 1 - \sqrt{2}\right)$ and $P_2 = \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 1 + \sqrt{2}\right)$. These do give local maxima. We see that the point P_2 is further than P_1 from the origin:

$$\begin{aligned} &\left(-\frac{1}{\sqrt{2}}\right)^2 + \left(-\frac{1}{\sqrt{2}}\right)^2 + \left(1 + \sqrt{2}\right)^2 \\ = &\ \frac{1}{2} + \frac{1}{2} + 1 + 2\sqrt{2} + 2 = 4 + 2\sqrt{2}, \end{aligned}$$

while

$$\begin{aligned} &\left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{\sqrt{2}}\right)^2 + \left(1 - \sqrt{2}\right)^2 \\ = &\ \frac{1}{2} + \frac{1}{2} + 1 - 2\sqrt{2} + 2 = 4 - 2\sqrt{2} > 1. \end{aligned}$$

Summary:
The closest points on the ellipse to the origin are $(1, 0, 0)$ and $(0, 1, 0)$ which have a distance 1 from the origin. The furthest point is $\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 1 + \sqrt{2}\right)$ which is at a distance $4 + 2\sqrt{2}$ from the origin.

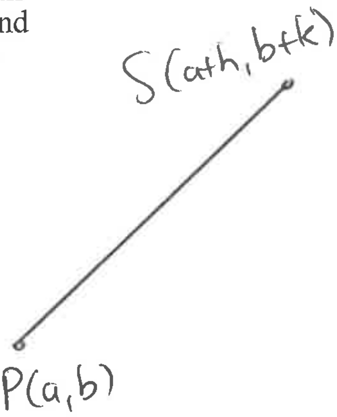
Section 14.9: Taylor's Formula for Two Variables

We begin with a proof of Theorem 11 from Section 4.7, namely the two variable version of the second derivative test:

Derivation of the Second Derivative Test (Theorem 11 of Section 4.7)

Let $f(x,y)$ have continuous second partial derivatives in an open region R containing a point $P(a,b)$ where $f_x = 0$ and $f_y = 0$. Let h,k be small increments. Let S be the point $(a+h, b+k)$ and assume that both S and the line segment joining S and P lie in R. The line segment has the parametrization

$$x = a + th \text{ and } y = b + tk \text{ for } t \in [0,1].$$



Let

$$F(t) = f(a+th, b+tk), t \in [0,1].$$

By the chain rule

$$F'(t) = f_x \frac{dx}{dt} + f_y \frac{dy}{dt} = f_x h + f_y k.$$

Then

$$\begin{aligned} F''(t) &= \frac{\partial}{\partial x}(f_x h + f_y k)\frac{dx}{dt} + \frac{\partial}{\partial y}(f_x h + f_y k)\frac{dy}{dt} \\ &= (f_{xx} h + f_{yx} k) h + (f_{xy} h + f_{yy} k) k \\ &= h^2 f_{xx} + 2hk f_{xy} + k^2 f_{yy}. \end{aligned}$$

(Recall that $f_{yx} = f_{xy}$). We apply the usual one variable Taylor's formula for the function F with $n = 2$ and $a = 0$: for some c between 0 and 1,

$$\begin{aligned} F(1) &= F(0) + F'(0)(1-0) + F''(c)\frac{(1-0)^2}{2} \\ &= F(0) + F'(0) + \frac{1}{2}F''(c). \end{aligned}$$

In terms of f this says

$$\begin{aligned} f(a+h, b+k) &= f(a,b) + hf_x(a,b) + kf_y(a,b) \\ &\quad + \frac{1}{2}\left(h^2 f_{xx} + 2hk f_{xy} + k^2 f_{yy}\right)_{|(a+ch, b+ck)}. \end{aligned}$$

We assumed that $f_x(a,b) = 0 = f_y(a,b)$, so that

$$f(a+h, b+k) - f(a,b) = \frac{1}{2}\left(h^2 f_{xx} + 2hk f_{xy} + k^2 f_{yy}\right)_{|(a+ch, b+ck)}.$$

Observe that if f has a local maximum at (a,b) then for small h,k, we have

$$Q(c) = \left(h^2 f_{xx} + 2hk f_{xy} + k^2 f_{yy}\right)_{|(a+ch, b+ck)} < 0$$

If $Q(0) \neq 0$, then for small enough c, $Q(c)$ will have the same sign as $Q(0)$, by continuity. So let us examine

$$\begin{aligned} f_{xx}Q(0) &= \left(h^2 f_{xx}^2 + 2hk f_{xy} f_{xx} + k^2 f_{yy} f_{xx}\right)_{|(a,b)} \\ &= (h f_{xx} + k f_{xy})^2 + k^2 \left(f_{yy} f_{xx} - f_{xy}^2\right). \end{aligned}$$

From this, we see that
(1) If $f_{xx} < 0$ and $f_{yy}f_{xx} - f_{xy}^2 > 0$ at (a,b), then $Q(0) < 0$ and f has a **local maximum** at (a,b);
(2) If $f_{xx} > 0$ and $f_{yy}f_{xx} - f_{xy}^2 > 0$ at (a,b), then $Q(0) > 0$ and f has a **local minimum** at (a,b).
(3) If $f_{yy}f_{xx} - f_{xy}^2 < 0$ at (a,b), then we can find arbitrarily small values of (k,h) for which $Q(0) < 0$ and others for which $Q(0) > 0$, so f has a **saddle point** at (a,b).
(4) If $f_{yy}f_{xx} - f_{xy}^2 = 0$ at (a,b), need other tests.

Next, we turn to

Taylor's Formula for Functions of Two Variables
Above we saw that

$$F'(t) = f_x h + f_y k = \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right) f(x,y).$$

and

$$F''(t) = \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^2 f(x,y).$$

We can do this for higher derivartives:

$$F^{(n)}(t) = \frac{d^n}{dt^n}F(t) = \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^n f(x,y).$$

We can interpret it as follows: we expand the operator $\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^n$ using the binomial theorem and then apply the partial derivative operators to f. If F has enough derivatives, we can expand F by the usual one variable Taylor formula for F :

$$F(t) = F(0) + F'(0)t + \frac{F''(0)}{2!}t^2 + \ldots + \frac{F^{(n)}(0)}{n!}t^n + \text{remainder.}$$

Applying the relations above leads to Taylor's formula for functions of two variables:

Taylor's Formula for $f(x,y)$ at (a,b) :
Suppose that $f(x,y)$ and its partial derivatives through order $n+1$ are continuous in an open rectangular region R centered at (a,b). Then throughout

R,

$$\begin{aligned} f(a+h,b+k) &= f(a,b) + (f_x h + f_y k)_{|(a,b)} + \frac{1}{2!}\left(h^2 f_{xx} + 2hk f_{xy} + k^2 f_{yy}\right)_{|(a,b)} \\ &+ \frac{1}{3!}\left(h^3 f_{xxx} + 3h^2 k f_{xxy} + 3hk^2 f_{xyy} + k^3 f_{yyy}\right)_{|(a,b)} + \ldots + \frac{1}{n!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^n f_{|(a,b)} \\ &+ \frac{1}{(n+1)!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^{n+1} f_{|(a+ch,b+ck)} \end{aligned}$$

Here c is some point between 0 and 1.

When $(a,b) = (0,0)$, and we take $(h,k) = (x,y)$, this simplifies to:

Taylor's Formula for $f(x,y)$ at $(0,0)$

$$\begin{aligned} f(x,y) &= f(0,0) + (f_x x + f_y y) + \frac{1}{2!}\left(x^2 f_{xx} + 2xy f_{xy} + y^2 f_{yy}\right) \\ &+ \frac{1}{3!}\left(x^3 f_{xxx} + 3x^2 y f_{xxy} + 3xy^2 f_{xyy} + y^3 f_{yyy}\right) \\ &+ \ldots + \frac{1}{n!}\left(x^n \frac{\partial^n f}{\partial x^n} + n x^{n-1} y \frac{\partial^n f}{\partial x^{n-1}\partial y} + \ldots + y^n \frac{\partial^n f}{\partial y^n}\right) \\ &+ \frac{1}{(n+1)!}\left(x^{n+1}\frac{\partial^{n+1} f}{\partial x^{n+1}} + (n+1)\, x^n y \frac{\partial^{n+1} f}{\partial x^n \partial y} + \ldots + y^{n+1}\frac{\partial^{n+1} f}{\partial y^{n+1}}\right)_{|(cx,cy)}. \end{aligned}$$

(All derivatives except the last term are evaluated at $(0,0)$.)

Example
Find a quadratic approximation to $f(x,y) = \sin x \sin y$ near $(0,0)$. (That is find the terms up to those involving x^2 and y^2 in Taylor's formula). Estimate the error if $|x| \leq 0.1$ and $|y| \leq 0.1$.
Solution
We use Taylor's formula at $(0,0)$ with $n = 2$:

$$\begin{aligned} f(x,y) &= f(0,0) + (f_x x + f_y y) + \frac{1}{2!}\left(x^2 f_{xx} + 2xy f_{xy} + y^2 f_{yy}\right) \\ &+ \frac{1}{3!}\left(x^3 f_{xxx} + 3x^2 y f_{xxy} + 3xy^2 f_{xyy} + y^3 f_{yyy}\right)_{|(cx,cy)}. \end{aligned}$$

Here $f(0,0) = 0$ and

$$\begin{aligned} f_x &= \cos x \sin y \text{ so at } (0,0),\ f_x = 0; \\ f_y &= \sin x \cos y \text{ so at } (0,0),\ f_y = 0; \\ f_{xx} &= -\sin x \sin y \text{ so at } (0,0),\ f_{xx} = 0; \\ f_{yy} &= -\sin x \sin y \text{ so at } (0,0),\ f_{yy} = 0; \\ f_{xy} &= \cos x \cos y \text{ so at } (0,0),\ f_{xy} = 1; \end{aligned}$$

Thus

$$\begin{aligned} f(x,y) &= 0+(0)+\frac{1}{2!}(0+2xy+0) \\ &\quad +\frac{1}{3!}\left(x^3 f_{xxx}+3x^2 y f_{xxy}+3xy^2 f_{xyy}+y^3 f_{yyy}\right)_{|(cx,cy)} \\ &= xy+\frac{1}{3!}\left(x^3 f_{xxx}+3x^2 y f_{xxy}+3xy^2 f_{xyy}+y^3 f_{yyy}\right)_{|(cx,cy)} \end{aligned}$$

The quadratic approximation is just

$$f(x,y) \approx xy.$$

Since all partial derivatives are products of sines and cosines, their absolute values are bounded by 1, so

$$\begin{aligned} |Error| &= \left|\frac{1}{3!}\left(x^3 f_{xxx}+3x^2 y f_{xxy}+3xy^2 f_{xyy}+y^3 f_{yyy}\right)_{|(cx,cy)}\right| \\ &\le \frac{1}{3!}\left(|x|^3+3|x|^2|y|+3|x||y|^2+|y|^3\right) \\ &\le \frac{1}{6}(1+3+3+1)(0.1)^3=\frac{4}{3}(0.001)=0.0013333... \le 0.00134. \end{aligned}$$

Section 14.10: Partial Derivatives with Constrained Variables

In differentiating functions $f(x, y, z)$ of three variables, we assumed that x, y, z are independent. Sometimes they are not independent. This section discusses how to cope with this added difficulty.