

MAE111 Engineering Mathematics II
(2003/2004 Sem. 1)

SECTION 7: PARTIAL DIFFERENTIATION

D. Burgess
School of Mathematical Sciences
Queen Mary, University of London,
Mile End Road, London E1 4NS, U.K.
D.Burgess@qmul.ac.uk

If you find a mistake, then please tell the author!

Copyright © 2003 by David Burgess.

This material is not in the public domain. It is only for the use of students studying MAE111. Copies, paper or electronic are only permitted for the personal use of such students. No electronic version, in any form, should be made publically available except with authorization.

December 1, 2003

Section 7

Partial Differentiation

7.1 Functions of Two Variables

We are familiar with the idea of a function of one variable:

$$z = g(x)$$

and the idea of finding how this function changes by looking at its rate of change as x changes, which is just the derivative function

$$\frac{dg}{dx}$$

Now, consider a function of *two* variables x and y :

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

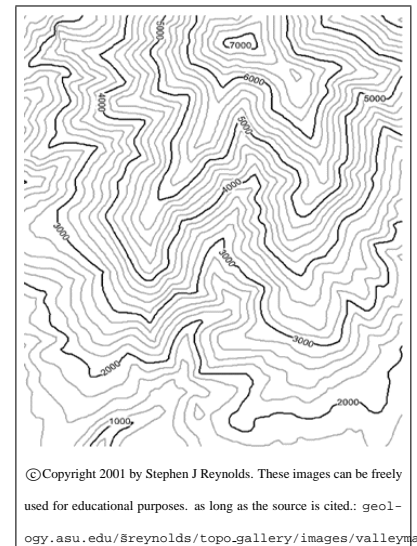
For every pair of values for x and y , the function will produce a value for z . Actually, the function might only be defined in a region of x and y . Such functions are very common in physics and engineering. For example: water speed as a function of depth and position across a river, surface wind speed as a function of position on the globe; current density over the cross section of a conductor, etc. The most obvious example is the following: consider the height at a position on a map, where the map position is given by two coordinates. In this case, the height defines a surface above the x - y plane. This surface can be thought of as the topography of a landscape: the hills and valleys. One way to represent this surface is as a three dimensional view. Alternatively, a *contour map* reveals how the surface goes up and down. A contour is a line of constant height, in other words, moving along a contour does not change the value of the function.

7.2 Partial Derivatives

Now, starting at some position (x, y) one moves in the x direction. How does the function f change? We can use the same idea as differentiation of a function of one variable, because we are not changing y .

The **partial derivative** of the function f wrt x is defined as the rate of change of the function as x changes, while y is kept constant.

$$\left. \frac{\partial z}{\partial x} \right|_y = \lim_{\delta x \rightarrow 0} \frac{f(x + \delta x, y) - f(x, y)}{\delta x}$$



Note the special ∂ symbol indicating partial differentiation. It must be used when partial differentiation is being used. Using the “d” symbol for ordinary differentiation means something completely different! Note also the

$$\left. \vphantom{\frac{\partial z}{\partial x}} \right|_y$$

notation, which indicates that y is to be kept constant while the partial derivative wrt x is taken. Usually, when it is obvious what is being kept constant, this part of the notation is left out.

The same argument can be made for the variation of the function as the y value is changed, but the x is kept constant. This then defines the partial derivative wrt to y :

$$\left. \frac{\partial z}{\partial y} \right|_x = \lim_{\delta y \rightarrow 0} \frac{f(x, y + \delta y) - f(x, y)}{\delta y}$$

7.2.1 Higher Order Partial Derivatives

The operation of partial differentiation can be repeated, eg taking the partial derivative wrt x of the partial derivative wrt x , and so on. There are four second order partial derivatives:

$$\frac{\partial^2 z}{\partial x^2} \quad \frac{\partial^2 z}{\partial y^2} \quad \frac{\partial^2 z}{\partial x \partial y} \quad \frac{\partial^2 z}{\partial y \partial x}$$

Note that the last two of these differ in the order of differentiating wrt x and y .

For any *continuous* function (which means almost all the functions you will meet in this course), one has the important result that

$$\frac{\partial^2 z}{\partial x \partial y} = \frac{\partial}{\partial x} \left(\frac{\partial z}{\partial y} \right) = \frac{\partial}{\partial y} \left(\frac{\partial z}{\partial x} \right) = \frac{\partial^2 z}{\partial y \partial x}$$

So that, for such functions, there are only three second order partial derivatives to calculate.

7.2.2 Evaluating Partial Derivatives

Partial differentiation is a linear operation, so these two basic rules are obeyed (for two functions f and g of variables x and y):

$$\begin{aligned} \frac{\partial}{\partial x}(f + g) &= \frac{\partial f}{\partial x} + \frac{\partial g}{\partial x} \\ \frac{\partial}{\partial x}(fg) &= f \frac{\partial g}{\partial x} + g \frac{\partial f}{\partial x} \end{aligned}$$

Almost all the rules of **ordinary differentiation** can be applied for evaluating partial derivatives, but remember: *when evaluating a partial derivative wrt to a variable, all other variables are treated as constants.*

Example

Consider

$$z = f(x, y) = 4x^2 + 3xy + 5y^2$$

Differentiating wrt x (keeping y constant) and wrt y (keeping x constant):

$$\begin{aligned} \frac{\partial z}{\partial x} &= 8x + 3y \\ \frac{\partial z}{\partial y} &= 3x + 10y \end{aligned}$$

Differentiating again, in all possible combinations:

$$\begin{aligned}\frac{\partial^2 z}{\partial x^2} &= \frac{\partial}{\partial x}(8x + 3y) = 8 \\ \frac{\partial^2 z}{\partial y^2} &= \frac{\partial}{\partial y}(3x + 10y) = 10 \\ \frac{\partial^2 z}{\partial x \partial y} &= \frac{\partial}{\partial x}(3x + 10y) = 3 \\ \frac{\partial^2 z}{\partial y \partial x} &= \frac{\partial}{\partial y}(8x + 3y) = 3\end{aligned}$$

We have confirmed that the two mixed second order partial derivatives are the same.

Example

$$z = f(x, y) = \sin(xy)$$

First, the first order partial derivatives (and remember function of a function):

$$\begin{aligned}\frac{\partial z}{\partial x} &= y \cos(xy) \\ \frac{\partial z}{\partial y} &= x \cos(xy)\end{aligned}$$

Differentiating again, in all possible combinations (this time remembering the product rule, as well as function of a function):

$$\begin{aligned}\frac{\partial^2 z}{\partial x^2} &= -y^2 \sin(xy) \\ \frac{\partial^2 z}{\partial y^2} &= -x^2 \sin(xy) \\ \frac{\partial^2 z}{\partial x \partial y} &= y(-x \sin(xy)) + 1 \cdot \cos(xy) = \cos(xy) - xy \sin(xy) \\ \frac{\partial^2 z}{\partial y \partial x} &= x(-y \sin(xy)) + 1 \cdot \cos(xy) = \cos(xy) - xy \sin(xy)\end{aligned}$$

We have confirmed (again) that the two mixed second order partial derivatives are the same.

7.3 Small Increments

We now state, and then derive an important result linking partial derivatives to small (and finite) changes.

We can ask how much a function $z(x, y)$, of two variables x and y , changes when the variables are changed by small amounts δx and δy . In other words define the change in the function δz caused by changing the variables from (x, y) by small amounts to $(x + \delta x, y + \delta y)$.

$$\delta z = z(x + \delta x, y + \delta y) - z(x, y)$$

The change in the function is given, approximately, by

$$\delta z \approx \frac{\partial z}{\partial x} \delta x + \frac{\partial z}{\partial y} \delta y$$

This is an approximation, and gets more accurate as δx and δy tend to zero. This important result has many applications, but first we will show where it comes from.

Consider the definition

$$\frac{\partial z}{\partial x} = \lim_{\delta x \rightarrow 0} \frac{f(x + \delta x, y) - f(x, y)}{\delta x}$$

When δx is finite we must be able to write this as:

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

where α is an “error” term that gets smaller as δx tends to zero:

$$\lim_{\delta x \rightarrow 0} \alpha = 0$$

And similarly:

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

where

$$\lim_{\delta x \rightarrow 0} \beta = 0$$

We can write

$$\begin{aligned} \delta z &= z(x + \delta x, y + \delta y) - z(x, y) \\ \delta z &= z(x + \delta x, y + \delta y) - z(x + \delta x, y) + z(x + \delta x, y) - z(x, y) \\ \delta z &= \frac{\partial z}{\partial x} \delta x + \frac{\partial z}{\partial y} \delta y + \alpha + \beta \end{aligned}$$

Note, that we have assumed that the partial derivative wrt y at x and at $x + \delta x$ are the same, since δx is small. Since α and β both get smaller as δx and δy tend to zero, therefore we can write:

$$\delta z \approx \frac{\partial z}{\partial x} \delta x + \frac{\partial z}{\partial y} \delta y$$

Actually, there is a better way of proving this result, using Taylor series:

The aim is to find $z(x + \delta x, y + \delta y)$, so we first expand, using a Taylor’s series, about x by an amount δx , but at fixed $(y + \delta y)$:

$$z(x + \delta x, y + \delta y) = z(x, y + \delta y) + \frac{\partial z}{\partial x} \Big|_{y+\delta y} \delta x + O(\delta x^2)$$

Note that we use the notation $O(\delta x^2)$ to indicate the rest of the terms, which include at least δx^2 . Note also, that the terms on the rhs (including partial derivatives) are evaluated at $(x, y + \delta y)$. So, we next expand about y , at fixed x :

$$z(x, y + \delta y) = z(x, y) + \frac{\partial z}{\partial y} \delta y + O(\delta y^2)$$

This is then substituted in the first expansion:

$$z(x + \delta x, y + \delta y) = z(x, y) + \frac{\partial z}{\partial y} \delta y + \frac{\partial}{\partial x} \left[z(x, y) + \frac{\partial z}{\partial y} \delta y + O(\delta y^2) \right] \delta x + O(\delta x^2) + O(\delta y^2)$$

Keeping only the terms depending linearly on δx and δy , we obtain the required approximation:

$$\delta z \approx \frac{\partial z}{\partial x} \delta x + \frac{\partial z}{\partial y} \delta y$$

Example

Consider the height near some mountain is given by the following function of x and y coordinates:

$$H = (1 - x^2)e^{-y^2}$$

Starting from the position $(\sqrt{2}/2, \sqrt{2}/2)$, find the *approximate* drop in height if one moves in the following ways: (i) $\delta x = 0.1, \delta y = 0$; (ii) $\delta x = 0, \delta y = 0.1$; (iii) $\delta x = 0.1, \delta y = -0.2$.

First calculate:

$$\frac{\partial H}{\partial x} = -2xe^{-y^2} \quad \frac{\partial H}{\partial y} = -2ye^{-y^2}(1 - x^2)$$

The change in height is then given approximately by

$$\delta H = \frac{\partial H}{\partial x} \delta x + \frac{\partial H}{\partial y} \delta y$$

So, evaluating the partial derivatives at the point $(\sqrt{2}/2, \sqrt{2}/2)$ one finds:

$$\frac{\partial H}{\partial x} = -2 \frac{\sqrt{2}}{2} e^{-1/2} = -\sqrt{2} e^{-1/2}$$

$$\frac{\partial H}{\partial y} = \left(1 - \frac{1}{2}\right) \left(-2 \frac{\sqrt{2}}{2}\right) e^{-1/2} = \frac{-\sqrt{2}}{2} e^{-1/2}$$

We find the approximate change in height from that point:

$$\delta H \approx -\sqrt{2} e^{-1/2} \left(\delta x + \frac{1}{2} \delta y\right) = -0.858 \left(\delta x + \frac{1}{2} \delta y\right)$$

We can simply substitute for different value of δx and δy to find out how the height changes.

$$\delta x = 0.1, \delta y = 0 \quad \Rightarrow \delta H \approx -0.0858$$

$$\delta x = 0, \delta y = 0.1 \quad \Rightarrow \delta H \approx -0.0429$$

$$\delta x = 0.1, \delta y = -0.2 \quad \Rightarrow \delta H \approx 0$$

7.3.1 Error Estimation

Suppose that we have the following situation: two quantities x and y , are measured, but their measurement is in error by amounts δx and δy . Then the error in the measurement of a third quantity z which depends on x and y , can be found from the rule for small increments. The value of δz gives an estimate on the error in the measurement of z .

Example

A factory makes oil drums, which are cylindrical, of radius r , and height h . The volume of one oil drum is

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

Suppose that in measuring the radius there is an error δr , and similarly an error δh in the height measurement. Using the formula for small increments, the (maximum) error in the measurement of the volume will be

$$\delta V = \frac{\partial V}{\partial r} \delta r + \frac{\partial V}{\partial h} \delta h = 2\pi r h \delta r + 2\pi r^2 \delta h$$

So, if $h = 1\text{m}$, and $r = 0.25\text{m}$, so that $V = 0.196\text{m}^3$, and the error in manufacture in h and r is $+0.005\text{m}$ (ie 5mm), then

$$\delta V = 2\pi(0.25)(0.005) + 2\pi(0.0625)(0.005) = 2\pi(0.00156) = 0.0098\text{m}^3$$

So a positive error on both r and h could result in giving a free extra 0.01 m^3 (ie 10 litres) of oil. Of course, if the errors on r and h were negative, then the oil drum would contain less than intended.

7.4 Total Derivative

Suppose we have a function $z = f(x, y)$ of two variables. Furthermore, both x and y depend on a third variable t :

$$x = x(t), \quad y = y(t)$$

Since z depends on x and y , it must also be the case that we can consider z as a function of t

$$z = z(t)$$

The question is: What is the rate of change of z with respect to changing t ?

The rule of small increments applies:

$$\delta z = \frac{\partial f}{\partial x} \delta x + \frac{\partial f}{\partial y} \delta y$$

For this small change in the function there must be a corresponding small change in the variable t , namely δt . Taking the small increments equation, dividing through by δt and letting $\delta t \rightarrow 0$ we find the following result:

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

This is called the **total derivative** of $z(t)$ wrt t . When t represents time, this gives the rate of change of z in terms of the rates of change of x and y . Note that we use the notation for ordinary differentiation for functions of one variable.

Example

Consider

$$z = xy^2$$

So,

$$\frac{\partial z}{\partial x} = y^2, \quad \frac{\partial z}{\partial y} = 2xy$$

The variables x and y vary according to

$$x = t^2 + 1, \quad y = t + 2$$

Substituting into the formula for the total derivative of z wrt t :

$$\begin{aligned} \frac{dz}{dt} &= y^2 \frac{d}{dt}(t^2 + 1) + 2xy \frac{d}{dt}(t + 2) \\ \frac{dz}{dt} &= (t + 2)^2(2t) + 2(t^2 + 1)(t + 2)(1) \end{aligned}$$

So,

$$\frac{dz}{dt} = 2(t + 2)(2t^2 + 2t + 1)$$

Note, that the same result could be obtained by substituting for x and y in terms of t :

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

and then using ordinary differentiation (using product rule)

$$\frac{dz}{dt} = (t + 2)^2(2t) + (t^2 + 1)2(t + 2) = 2(t + 2)(2t^2 + 2t + 1)$$

7.5 Implicit Differentiation

This section is not given in lectures.

A special case of the total derivative is when y is itself a function of x . This has the effect of replacing t by x . It follows that z is a function only of x and the equation for the total derivative becomes

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

From which it obviously follows that

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

Example

Suppose

$$u = f(x, y) = \tan^{-1} \frac{x}{y}$$

and

$$y = \sin x$$

Then by the above formula for implicit differentiation:

$$\frac{du}{dx} = \frac{y}{x^2 + y^2} - \frac{x}{x^2 + y^2} \cos x = \frac{\sin x - x \cos x}{x^2 + \sin^2 x}$$

Remember that

$$\frac{d}{dx} \tan^{-1} x = \frac{1}{x^2 + 1}, \quad \Rightarrow \frac{d}{dx} \tan^{-1} \frac{x}{a} = \frac{a^2}{x^2 + a^2} \frac{d}{dx} \left(\frac{x}{a} \right) = \frac{a}{x^2 + a^2}$$

and the rest follows from application of the chain rule.

The equation

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

implicitly defines y as a function of x . Since z is identically zero, its derivative must be zero also, so the total derivative will vanish, and we are left with

$$\frac{dy}{dx} = - \frac{\partial f}{\partial x} \left(\frac{\partial f}{\partial y} \right)^{-1}$$

7.6 Change of Variables – The Chain Rule

For a function of one variable $g(x)$, if x is a function of another variable u , then we can find the derivative of g wrt u (ie, the how g changes as u changes) by using the chain rule:

$$\frac{dg}{du} = \frac{dg}{dx} \frac{dx}{du}$$

We can obtain a similar rule for functions of two variables. Consider a function z of two variables x and y , so that $z = z(x, y)$. The variables x and y are themselves functions of two other variables u and v :

$$x = x(u, v), \quad y = y(u, v)$$

Now, small changes in u and v of δu and δv will produce corresponding small changes in x and y of δx and δy . Using the rule for small increments we have

$$\begin{aligned} \delta x &= \frac{\partial x}{\partial u} \delta u + \frac{\partial x}{\partial v} \delta v \\ \delta y &= \frac{\partial y}{\partial u} \delta u + \frac{\partial y}{\partial v} \delta v \end{aligned}$$

But, from z considered as a function of x and y , we have

$$\delta z = \frac{\partial z}{\partial x} \delta x + \frac{\partial z}{\partial y} \delta y$$

Substituting and rearranging:

$$\begin{aligned} \delta z &= \frac{\partial z}{\partial x} \left(\frac{\partial x}{\partial u} \delta u + \frac{\partial x}{\partial v} \delta v \right) + \frac{\partial z}{\partial y} \left(\frac{\partial y}{\partial u} \delta u + \frac{\partial y}{\partial v} \delta v \right) \\ \delta z &= \left(\frac{\partial z}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial u} \right) \delta u + \left(\frac{\partial z}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial v} \right) \delta v \end{aligned}$$

Now, we can also consider z as a function of u and v , in which case we have the following rule of small increments:

$$\delta z = \frac{\partial z}{\partial u} \delta u + \frac{\partial z}{\partial v} \delta v$$

and by comparing these equations, we find the following identities:

$$\begin{aligned} \frac{\partial z}{\partial u} &= \frac{\partial z}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial u} \\ \frac{\partial z}{\partial v} &= \frac{\partial z}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial v} \end{aligned}$$

These identities are the equivalent for the chain rule for functions of two variables. They show how the rate of change wrt u can be expressed in terms of the rates of change of x and y wrt u .

Note that the same results can be obtained by considering

$$\delta z = \frac{\partial z}{\partial x} \delta x + \frac{\partial z}{\partial y} \delta y$$

and dividing by δu (or δv) and letting δu tend to zero.

Example

Consider

$$z = \sin(x + y)$$

with

$$x = u^2 + v^2, \quad y = 2uv$$

Evaluating partial derivatives:

$$\frac{\partial x}{\partial u} = 2u, \quad \frac{\partial y}{\partial u} = 2v, \quad \frac{\partial x}{\partial v} = 2v, \quad \frac{\partial y}{\partial v} = 2u$$

$$\frac{\partial z}{\partial x} = \cos(x + y), \quad \frac{\partial z}{\partial y} = \cos(x + y)$$

Therefore:

$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial u} = 2 \cos(x + y)(u + v)$$

$$\frac{\partial z}{\partial v} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial v} = 2 \cos(x + y)(v + u)$$

Example

$$z = u$$

incomplete ...

Example

Given that $u = f(x, y)$ and that x and y are related to variables r and θ by

$$x = r \cos \theta, \quad y = r \sin \theta$$

prove that

$$r \frac{\partial u}{\partial r} = x \frac{\partial f}{\partial x} + y \frac{\partial f}{\partial y}$$

$$\frac{\partial u}{\partial \theta} = x \frac{\partial f}{\partial y} - y \frac{\partial f}{\partial x}$$

This result follows from

$$\frac{\partial u}{\partial r} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial r} = \frac{\partial f}{\partial x} \cos \theta + \frac{\partial f}{\partial y} \sin \theta$$

And multiplying by r , we obtain the required result.

Similarly

$$\frac{\partial u}{\partial \theta} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial \theta} = \frac{\partial f}{\partial x} (-r \sin \theta) + \frac{\partial f}{\partial y} (r \cos \theta) = x \frac{\partial f}{\partial y} - y \frac{\partial f}{\partial x}$$

7.7 Stationary Points

For functions of a single variable, we know that the extrema (eg minima or maxima) of a curve can be found by finding points when the derivative vanishes. A similar rule applies to functions of two variables.

We wish to find so called “stationary points” where the rate of change of the function is zero. But this zero rate of change must be true for departures in whatever direction. In other words, we require, for a function $z(x, y)$:

$$\delta z = 0 = \frac{\partial z}{\partial x} \delta x + \frac{\partial z}{\partial y} \delta y$$

for *all* values of δx and δy . This is only true if both

$$\begin{aligned}\frac{\partial z}{\partial x} &= 0 \\ \frac{\partial z}{\partial y} &= 0\end{aligned}$$

These two equations, simultaneously, determine the stationary points of a function.

The question whether a stationary point is a maximum, or a minimum, or some other kind of point, is more difficult to test. A minimum is a point, where any departure produces an increase in the function, whatever the direction of departure; and similarly for a maximum. These conditions can be shown to be equivalent to the following: (They can be shown by using the Taylor series expansion for a function of two variables.)

For a MAXIMUM:

$$\frac{\partial^2 z}{\partial x^2} < 0 \quad \text{AND} \quad \left(\frac{\partial^2 z}{\partial x^2}\right)\left(\frac{\partial^2 z}{\partial y^2}\right) - \left(\frac{\partial^2 z}{\partial x \partial y}\right)^2 > 0$$

For a MINIMUM:

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

If

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

then the point is called a SADDLE POINT, since the function decreases in some directions, but increases in others.

Example

Find the stationary points, and their nature, of the function

$$z = x^4 + 2x^2y - 2x^2 + 2y^2$$

The stationary points have

$$\begin{aligned}\frac{\partial z}{\partial x} &= 0 = 4x^3 + 4xy - 4x \\ \frac{\partial z}{\partial y} &= 0 = 2x^2 + 4y \quad \Rightarrow 4y = -2x^2\end{aligned}$$

Substituting from second into first:

$$4x^3 - 2x^3 - 4x = 2x(x^2 - 2) \quad x = 0, +\sqrt{2}, -\sqrt{2}$$

So the stationary points are:

$$(0, 0), \quad (\sqrt{2}, -1), \quad (-\sqrt{2}, -1)$$

Next, evaluating the second order derivatives:

$$\begin{aligned}\frac{\partial^2 z}{\partial x^2} &= 12x^2 + 4y - 4 \\ \frac{\partial^2 z}{\partial y^2} &= 4 \\ \frac{\partial^2 z}{\partial x \partial y} &= 4x\end{aligned}$$

And comparing with the conditions for the different types of stationary points we find:

Point	$\frac{\partial^2 z}{\partial x^2}$	$\left(\frac{\partial^2 z}{\partial x^2}\right)\left(\frac{\partial^2 z}{\partial y^2}\right) - \left(\frac{\partial^2 z}{\partial x \partial y}\right)^2$	Type
$(0, 0)$	-4	-16	Saddle point
$(\sqrt{2}, -1)$	16	32	Minimum
$(-\sqrt{2}, -1)$	16	32	Minimum

7.8 Method of Least Squares

Often one collects data, and then wants to find a curve that “best” fits that data. The usual method is to choose a type of curve (say a straight line) and then to adjust the parameters of that curve (eg slope) to give a “best” fit. The most useful definition of “best” is to minimize the distance of the points from the curve. This is called the Method of Least Squares, and is an application of partial differentiation.

Consider a set of n data points:

$$(x_i, y_i), \quad i = 1, \dots, n$$

To demonstrate the method we will show how to obtain a best fit to a straight line:

$$y = a + bx$$

The aim is to find suitable values for a and b .

The (vertical) distance to the line from any data point at x_i is

$$y_i - y_{\text{line}} = y_i - (a + bx_i)$$

Since this could be negative or positive, we square it, add all such values for all data points, to find the total square distance:

$$S = \sum_{i=1}^n (y_i - a - bx_i)^2 = S(a, b)$$

The best fit is when the function $S(a, b)$ is a minimum.

$$\begin{aligned} \frac{\partial S}{\partial a} &= 0 = \sum_{i=1}^n 2(y_i - a - bx_i)(-1) \\ \frac{\partial S}{\partial b} &= 0 = \sum_{i=1}^n 2(y_i - a - bx_i)(-x_i) \end{aligned}$$

These two equations lead to:

$$\begin{aligned} na + \left(\sum x_i\right)b &= \sum y_i \\ a\left(\sum x_i\right) + b\left(\sum x_i^2\right) &= \sum (x_i y_i) \end{aligned}$$

These are two simultaneous equations that can be solved for a and b the parameters of the “best” fit straight line.

Example

Here is a table of data (with $n = 5$), and the required sum totals.

	x	y	x^2	xy
	-2.4	-5.0	5.8	12.0
	-0.8	-1.5	0.64	1.2
	0.3	2.5	0.09	0.8
	1.9	6.4	3.6	12.2
	3.2	11.0	10.2	35.2
\sum	2.2	13.4	20.3	61.4

The equations to solve are:

$$\begin{aligned}5a + 2.2b - 13.4 &= 0 \\2.2a + 20.3b - 61.4 &= 0\end{aligned}$$

Dividing first by 5, second by 2.2, and subtracting we find:

$$b \left(\frac{20.3}{2.2} - \frac{2.2}{5} \right) = \frac{61.4}{2.2} - \frac{13.4}{5}$$

And

$$a = \frac{1}{5}(13.4 - 2.2b)$$

Giving

$$b = 2.87, \quad a = 1.42$$

Thus the best fit straight line is

$$y = 1.42 + 2.87x$$

7.9 Constrained Maxima and Minima

Often, with functions of several variables, the problem is not just to find the maximum or minimum, but to find the maximum or minimum which also satisfies a further constraint. There is a general technique to deal with such problems (Lagrange multipliers) but here we will deal with simpler problems. In such simple cases the following rule should be used: *use the constraint (or constraints) to eliminate one (or more) of the variables. And then look for maxima/minima.*

Example

Consider a sports field consisting of a rectangular area a by b (with $a > b$) and semicircular regions at each end (radius $b/2$). The rectangular area is to be used for football, and the perimeter is to be used for a running track. The track is to be 400 m long.

How can a and b be chosen to maximize the area of the football pitch, while maintaining the perimeter at the required length?

Area

$$A = ab$$

Perimeter:

$$P = a + \frac{\pi b}{2} + a + \frac{\pi b}{2} = 2a + \pi b = 400$$

Substituting to eliminate a :

$$a = 200 - \frac{\pi b}{2} \Rightarrow A = \left(200 - \frac{\pi b}{2}\right)b = 200b - \frac{\pi b^2}{2}$$

Now looking for maximum of A :

$$\frac{\partial A}{\partial b} = 200 - \pi b = 0 \Rightarrow b = \frac{200}{\pi} \text{ m}$$

And

$$a = 200 - \frac{\pi}{2} \frac{200}{\pi} = 100 \text{ m}$$

So that with these values for a and b the maximum area is

$$A = \frac{20000}{\pi} \text{ m}^2$$

Example

A cylindrical tin is to be made out of 600π cm² of sheet metal. What choice of radius r and height h will give the tin a maximum volume?

Volume:

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

Surface area (including ends)

$$S = 2\pi r h + 2\pi r^2 \quad \Rightarrow \quad h = \frac{S}{2\pi r} - r$$

Substituting for h into equation for V :

$$V = \pi r^2 \left(\frac{S}{2\pi r} - r \right) = \frac{Sr}{2} - \pi r^3$$

Looking for maximum volume (remember S is fixed):

$$\frac{\partial V}{\partial r} = \frac{S}{2} - 3\pi r^2 = 0 \quad \Rightarrow \quad r^2 = \frac{S}{6\pi}$$

Substituting $S = 600\pi$, one finds

$$r = 10 \text{ cm} \quad \Rightarrow \quad h = \frac{600\pi}{2\pi r} - r = \frac{600}{2 \cdot 10} - 10 = 20 \text{ cm}$$

These values maximize the volume of the tin for the given surface area.

$$V = 2000\pi \text{ cm}^3$$