2 Applications of partial differentiation

2.1 Higher derivatives

Let $U \subseteq \mathbb{R}^2$ and $f : U \to \mathbb{R}$. The partial derivatives f_x and f_y are functions of x and y, and so we can find their partial derivatives. We write f_{xy} to denote f_y differentiated with respect to x. Similarly f_{xx} denotes f_x differentiated with respect to x, and f_{yx} and f_{yy} denote f_x and f_y, respectively, differentiated with respect to y. The functions f_{xx}, f_{xy}, f_{yx} and f_{yy} are the **second** partial derivatives of f. For the other notation for partial derivatives, we write $\dfrac{\partial}{\partial x}\left(\dfrac{\partial f}{\partial x}\right)$ or just $\dfrac{\partial^2 f}{\partial x^2}$ for f_{xx}, $\dfrac{\partial}{\partial x}\left(\dfrac{\partial f}{\partial y}\right)$ or just $\dfrac{\partial^2 f}{\partial x \partial y}$ for f_{xy}, $\dfrac{\partial}{\partial y}\left(\dfrac{\partial f}{\partial x}\right)$ or just $\dfrac{\partial^2 f}{\partial y \partial x}$ for f_{yx} and $\dfrac{\partial}{\partial y}\left(\dfrac{\partial f}{\partial y}\right)$ or just $\dfrac{\partial^2 f}{\partial y^2}$ for f_{yy}.

We can also differentiate the second partial derivatives to get the third partial derivatives, and so on. For example, f_{xyy}, or $\frac{\partial^3 f}{\partial x \partial y^2}$, is the third partial derivative obtained from differentiating f_{yy} with respect to x.

Example 2.1. Let $f(x, y) = 3x^3y + 2xy^2 - 4x^2y$. Then

$$\begin{aligned} f_x(x,y) &= 9x^2y + 2y^2 - 8xy, \\ f_y(x,y) &= 3x^3 + 4xy - 4x^2, \\ f_{xx}(x,y) &= 18xy - 8y, \\ f_{yx}(x,y) &= 9x^2 + 4y - 8x, \\ f_{xy}(x,y) &= 9x^2 + 4y - 8x, \\ f_{yy}(x,y) &= 4x. \end{aligned}$$

Note that, in Example 2.1, $f_{yx} = f_{xy}$. This is true for all well-behaved functions. In particular, if f_{xy} and f_{yx} are both continuous on U (and U is an open subset of $\mathbb{R}^2$) then $f_{yx} = f_{xy}$ on U.

Example 2.2. Let $z = x \cos 2y$. Then

$$\begin{aligned} \frac{\partial z}{\partial x} &= \cos 2y, \\ \frac{\partial z}{\partial y} &= -2x \sin 2y, \\ \frac{\partial^2 z}{\partial x^2} &= 0, \\ \frac{\partial^2 z}{\partial y \partial x} &= -2 \sin 2y, \\ \frac{\partial^2 z}{\partial x \partial y} &= -2 \sin 2y, \\ \frac{\partial^2 z}{\partial y^2} &= -4x \cos 2y. \end{aligned}$$

Example 2.3. Let $z = \dfrac{xe^{2x}}{y^n}$. Find all the possible values of n given that

$$3x\frac{\partial^2 z}{\partial x^2} - xy^2\frac{\partial^2 z}{\partial y^2} = 12z.$$

For $z = \dfrac{xe^{2x}}{y^n}$ we have

$$\begin{aligned}
\frac{\partial z}{\partial x} &= \frac{e^{2x} + 2xe^{2x}}{y^n} \\
&= \frac{e^{2x}(1+2x)}{y^n}, \\
\frac{\partial z}{\partial y} &= -\frac{nxe^{2x}}{y^{n+1}}, \\
\frac{\partial^2 z}{\partial x^2} &= \frac{2e^{2x}(1+2x) + 2e^{2x}}{y^n} \\
&= \frac{4e^{2x}(x+1)}{y^n}, \\
\frac{\partial^2 z}{\partial y^2} &= \frac{n(n+1)xe^{2x}}{y^{n+2}}.
\end{aligned}$$

Thus

$$3x\frac{\partial^2 z}{\partial x^2} - xy^2\frac{\partial^2 z}{\partial y^2} = \frac{12xe^{2x}(x+1)}{y^n} - \frac{n(n+1)x^2e^{2x}}{y^n} = \frac{xe^{2x}}{y^n}\left(12(x+1) - n(n+1)x\right),$$

and so

$$\begin{aligned}
3x\frac{\partial^2 z}{\partial x^2} - xy^2\frac{\partial^2 z}{\partial y^2} = 12z \quad &\Leftrightarrow \quad \frac{xe^{2x}}{y^n}\left(12(x+1) - n(n+1)x\right) = \frac{12xe^{2x}}{y^n} \\
&\Leftrightarrow \quad 12(x+1) - n(n+1)x = 12 \\
& \quad \text{(since the result is true for all } x \text{ and } y\text{)} \\
&\Leftrightarrow \quad 12x + 12 - n(n+1)x = 12 \\
&\Leftrightarrow \quad 12 - n(n+1) = 0 \\
& \quad \text{(since the result is true for all } x\text{)} \\
&\Leftrightarrow \quad n^2 + n - 12 = 0 \\
&\Leftrightarrow \quad (n+4)(n-3) = 0 \\
&\Leftrightarrow \quad n = -4 \text{ or } n = 3.
\end{aligned}$$

Exercises 2.1.

1. Let $f(x,y) = 3x^3y^2 - 5xy^3 + 3x^2y^4 - y^5$. Find $\frac{\partial^2 f}{\partial x \partial y}$.

2. Determine all of the values of n such that $z = 2xy + x^n y^{2n}$ satisfies

$$2x^2\frac{\partial^2 z}{\partial x^2} - y^2\frac{\partial^2 z}{\partial y^2} + 18z = 36xy.$$

2.2 Stationary points

Let $U \subseteq \mathbb{R}^2$, and let $f : U \to \mathbb{R}$. Then $(a,b) \in U$ is called a **stationary point** of f if the tangent plane at (a,b) is horizontal. That is, the tangent plane exists and is parallel to the (x,y)-plane.

Recall from Chapter 1 that the equation of the tangent plane at (a,b) is

$$z = f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b).$$

Now, if (a,b) is a stationary point then this tangent plane is horizontal and hence $f_x(a,b) = 0$ and $f_y(a,b) = 0$. Conversely, if $f_x(a,b) = 0$ and $f_y(a,b) = 0$ then the tangent plane is horizontal and thus (a,b) is a stationary point. We therefore have the following theorem.

Theorem 2.4. *Suppose the $f(x,y)$ has a tangent plane at (a,b). Then (a,b) is a stationary point if and only if $f_x(a,b) = f_y(a,b) = 0$.*

The point concerning the existence of the tangent plane ensures that the function is "well behaved". The functions we are mainly dealing with below are all well behaved (polynomial functions, trigonometric functions, exponentials etc).

Example 2.5. Find all stationary points of

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

The partial derivatives of f are

$$\begin{aligned} f_x(x,y) &= 2xy + 3y^2 - 3y = y(2x+3y-3), \\ f_y(x,y) &= x^2 + 6xy - 3x = x(x+6y-3). \end{aligned}$$

Putting $f_x(x,y) = f_y(x,y) = 0$ gives

$$\begin{aligned} y(2x+3y-3) &= 0, & (1) \\ x(x+6y-3) &= 0. & (2) \end{aligned}$$

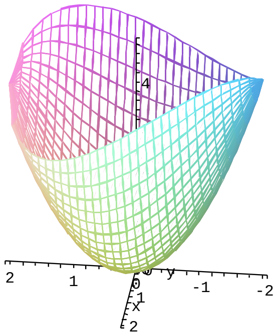
From equation (1) either $y = 0$ or $2x + 3y = 3$. If $y = 0$ then equation (2) gives $x(x-3) = 0$, and so $x = 0$ or 3. If $2x + 3y = 3$ then $6y = 6 - 4x$. Thus, in this case, equation (2) gives $x(3-3x) = 0$, and so $x = 0$ or 1; for $x = 0$ we get $y = 1$, and for $x = 1$ we get $y = \frac{1}{3}$. Thus the stationary points of f are $(0,0)$, $(3,0)$, $(0,1)$ and $(1, \frac{1}{3})$.

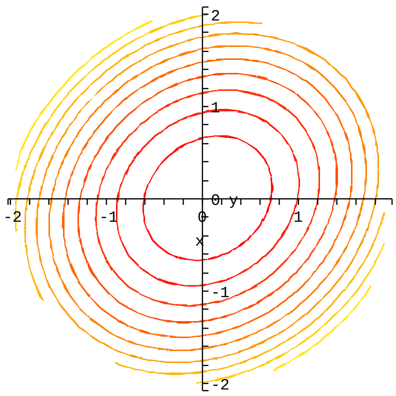
Types of stationary points

There are three types of stationary points: local minima, local maxima and saddle points. A function of two variables $f(x,y)$ has a **local maximum** at (a,b) if

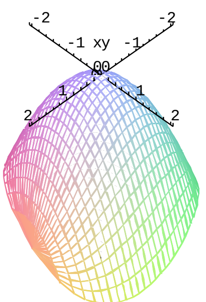
$f(x,y) \le f(a,b)$ for all (x,y) in some small enough disk around (a,b). The number $f(a,b)$ is the local maximum value. We will similarly call a point (a,b) a **local minimum** if $f(x,y) \ge f(a,b)$ for all (x,y) in some small enough disk around (a,b).[1] A **saddle point** is a stationary point that is neither a local maximum nor a local minimum. The following diagrams show a typical graph and contour-plot for each of these three types.

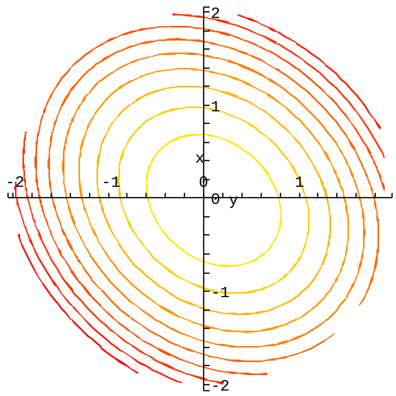
A local minimum





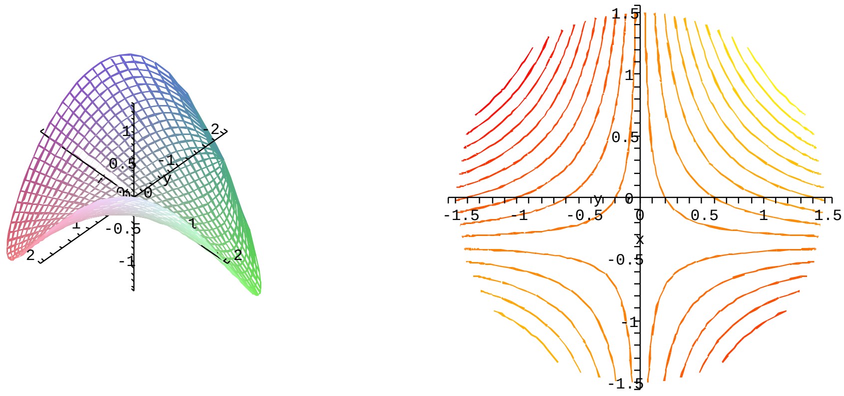
A local maximum





[1]Notice that the definitions of a local maximum and a local minimum do not require that (a,b) is a stationary point, so we can talk about local maxima and local minima (a,b) when the tangent plane does not exist at (a,b). However, in this section we will only consider local maxima and minima that are stationary points.

A saddle point



Classifying stationary points

For the remainder of this section we assume that $f_{xy} = f_{yx}$; as mentioned earlier, this is the case for all well-behaved functions and, in particular, it will be true for all examples below. To determine the nature of a stationary point (a, b) of f we first find the **Hessian matrix**,

$$\begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix}.$$

We then evaluate the determinant Δ of the Hessian matrix at the point (a, b). That is, we find the value, at (a, b), of

$$\Delta = \det \begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix} = f_{xx}f_{yy} - (f_{xy})^2.$$

- If $\Delta > 0$ and $f_{xx} > 0$ then (a, b) is a local minimum.
- If $\Delta > 0$ and $f_{xx} < 0$ then (a, b) is a local maximum.
- If $\Delta < 0$ then (a, b) is a saddle point.

If $\Delta = 0$ then this test provides no information about the nature of the stationary point (a, b), and other methods are required to classify the stationary point in this case.[2]

Example 2.6. Find all stationary points of

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

[2] The conditions on Δ and f_{xx} can be expressed more conceptually using the theory of symmetric matrices: $\Delta > 0$ and $f_{xx} > 0$ if and only if the Hessian matrix is positive definite, $\Delta > 0$ and $f_{xx} < 0$ if and only if the Hessian matrix is negative definite, $\Delta < 0$ if and only if the Hessian matrix is indefinite, and $\Delta = 0$ if and only if the Hessian matrix is singular.

and determine their nature.

From Example 2.5 the stationary points of f are $(0,0)$, $(3,0)$, $(0,1)$ and $(1,\frac{1}{3})$, and the partial derivatives are $f_x(x,y) = 2xy + 3y^2 - 3y = y(2x+3y-3)$ and $f_y(x,y) = x^2 + 6xy - 3x = x(x+6y-3)$. Now

$$\begin{aligned} f_{xx} &= 2y, \\ f_{xy} &= 2x + 6y - 3 \quad (= f_{yx}), \\ f_{yy} &= 6x. \end{aligned}$$

Thus

$$\Delta = \det \begin{pmatrix} 2y & 2x+6y-3 \\ 2x+6y-3 & 6x \end{pmatrix} = 12xy - (2x+6y-3)^2.$$

When $x = y = 0$, then $\Delta = -9 < 0$, and so $(0,0)$ is a saddle point.
When $x = 3$ and $y = 0$, then $\Delta = -9 < 0$, and so $(3,0)$ is a saddle point.
When $x = 0$ and $y = 1$, then $\Delta = -9 < 0$, and so $(0,1)$ is a saddle point.
When $x = 1$ and $y = \frac{1}{3}$, then $\Delta = 3 > 0$, $f_{xx} = \frac{2}{3} > 0$, and so $(1,\frac{1}{3})$ is a local minimum.

Example 2.7. Find and classify the stationary points of

$$f(x,y) = xye^{x+y}.$$

First we find the partial derivatives of f.

$$\begin{aligned} f_x(x,y) &= ye^{x+y} + xye^{x+y} = y(x+1)e^{x+y}, \\ f_y(x,y) &= x(y+1)e^{x+y} \qquad \text{(since } f \text{ is symmetrical in } x \text{ and } y\text{)}. \end{aligned}$$

Putting $f_x(x,y) = f_y(x,y) = 0$, and using the fact that $e^{x+y} \neq 0$, gives

$$\begin{aligned} y(x+1) &= 0, & (3) \\ x(y+1) &= 0. & (4) \end{aligned}$$

From equation (3) either $y = 0$ or $x = -1$. If $y = 0$ then equation (4) gives $x = 0$. If $x = -1$ then equation (4) gives $y = -1$. Thus the stationary points of f are $(0,0)$, and $(-1,-1)$. Now

$$\begin{aligned} f_{xx} &= ye^{x+1} + y(x+1)e^{x+y} = y(x+2)e^{x+y}, \\ f_{xy} &= (y+1)e^{x+y} + x(y+1)e^{x+y} = (y+1)(x+1)e^{x+y}, \\ f_{yy} &= x(y+2)e^{x+y}. \end{aligned}$$

Thus

$$\begin{aligned} \Delta &= \det \begin{pmatrix} y(x+2)e^{x+y} & (y+1)(x+1)e^{x+y} \\ (y+1)(x+1)e^{x+y} & x(y+2)e^{x+y} \end{pmatrix} \\ &= e^{2(x+y)} \left(xy(x+2)(y+2) - (y+1)^2(x+1)^2 \right). \end{aligned}$$

When $x = y = 0$, then $\Delta = -1 < 0$, and so $(0,0)$ is a saddle point.

When $x = y = -1$, then $\Delta = e^{-4} > 0$, $f_{xx} = -e^{-2} < 0$, and so $(-1,-1)$ is a local maximum.

Examples when $\Delta = 0$

We now consider the problem of classifying a stationary point (a, b) when $\Delta = 0$ at (a, b). There is no single method for determining the nature of (a, b) in this case, although it is often useful to find $f(x, y) - f(a, b)$ in order to determine the behaviour of f near (a, b); an alternative method is to consider how f varies along curves (normally straight lines) passing through (a, b).

Example 2.8. Find and classify the stationary points of

$$f(x, y) = x^2 + y^4 + 1.$$

First we find the partial derivatives of f.

$$\begin{aligned} f_x(x, y) &= 2x, \\ f_y(x, y) &= 4y^3. \end{aligned}$$

Putting $f_x(x, y) = f_y(x, y) = 0$, gives $x = 0$, $y = 0$. Thus the only stationary points of f is $(0, 0)$. Now

$$\begin{aligned} f_{xx} &= 2, \\ f_{xy} &= 0, \\ f_{yy} &= 12y^2. \end{aligned}$$

Thus

$$\begin{aligned} \Delta &= \det \begin{pmatrix} 2 & 0 \\ 0 & 12y^2 \end{pmatrix} \\ &= 24y^2. \end{aligned}$$

Thus at $(0, 0)$ we have $\Delta = 0$, and so the Hessian gives us no information about the nature of this stationary point. However, for $(x, y) \neq (0, 0)$, note that

$$\begin{aligned} f(x, y) - f(0, 0) &= x^2 + y^4 + 1 - (0 + 0 + 1) \\ &= x^2 + y^4 > 0. \end{aligned}$$

Thus $f(x, y) > f(0, 0)$ for all $(x, y) \neq (0, 0)$. Hence $(0, 0)$ is a local minimum.

Example 2.9. Find and classify the stationary points of

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

First we find the partial derivatives of f.

$$\begin{aligned} \frac{\partial f}{\partial x} &= y^2 - 2xy^2 + 4x^3, \\ \frac{\partial f}{\partial y} &= 2xy - 2x^2y. \end{aligned}$$

Putting $\dfrac{\partial f}{\partial x} = \dfrac{\partial f}{\partial y} = 0$ gives

$$\begin{aligned} y^2 - 2xy^2 + 4x^3 &= 0, & (5)\\ 2xy(1-x) &= 0. & (6)\end{aligned}$$

From equation (6) $x = 0$, $y = 0$ or $x = 1$. If either $x = 0$ or $y = 0$ then equation (5) gives $y = 0$ and $x = 0$, respectively. If $x = 1$ then equation (5) gives $y^2 = 4$, and so $y = 2$ or $y = -2$. Thus the stationary points of f are $(0, 0)$, $(1, 2)$ and $(1, -2)$. Now

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2} &= -2y^2 + 12x^2,\\ \frac{\partial^2 f}{\partial x \partial y} &= 2y - 4xy,\\ \frac{\partial^2 f}{\partial y^2} &= 2x - 2x^2. \end{aligned}$$

At the stationary point $(0, 0)$, each of the second derivatives is zero. Thus $\Delta = 0$ at $(0, 0)$, and so the Hessian gives no information about the nature of this stationary point. However, for $(x, y) \neq (0, 0)$, note that

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

In particular, for $x = y$ we get

$$f(x, x) - f(0, 0) = x^3 - x^4 + x^4 = x^3.$$

Since the sign of x^3 is the same as the sign of x, it follows that $f(x, x) - f(0, 0) > 0$, when $x > 0$, and $f(x, x) - f(0, 0) < 0$, when $x < 0$. Thus $(0, 0)$ is a saddle point.

The nature of the other stationary points of f can be determined using the Hessian. This is left as an exercise.

Example 2.10. Find the stationary points of $z = (x^2 + y^2)e^{-x^2-y^2}$, and determine their nature.

The partial derivatives of z are

$$\begin{aligned} \frac{\partial z}{\partial x} &= 2xe^{-x^2-y^2} + (x^2 + y^2)(-2x)e^{-x^2-y^2} = 2xe^{-x^2-y^2}(1 - x^2 - y^2),\\ \frac{\partial z}{\partial y} &= 2ye^{-x^2-y^2}(1 - x^2 - y^2) \qquad \text{(since z is symmetrical in x and y).} \end{aligned}$$

Putting $\dfrac{\partial z}{\partial x} = \dfrac{\partial z}{\partial y} = 0$, and using the fact that $e^{-x^2-y^2} \neq 0$ gives

$$\begin{aligned} x(1 - x^2 - y^2) &= 0, & (7)\\ y(1 - x^2 - y^2) &= 0. & (8)\end{aligned}$$

From equation (7) either $x = 0$ or $x^2 + y^2 = 1$. If $x = 0$ then equation (8) gives $y(1 - y^2) = 0$, and so $y = 0, -1$ or 1. Thus $(0, 0)$, $(0, 1)$ and $(0, -1)$ are stationary points of z. If $x^2 + y^2 = 1$ then equation (8) always holds. Thus every point on the circle $x^2 + y^2 = 1$ is also a stationary point of z.

Note that the points $(0, 1)$ and $(0, -1)$ lie on the circle $x^2 + y^2 = 1$. Thus the stationary points of z are $(0, 0)$ and every point on the circle $x^2 + y^2 = 1$.

Finding the second derivatives of z we get:

$$\begin{aligned}
\frac{\partial^2 z}{\partial x^2} &= (2e^{-x^2-y^2} + 2x(-2x)e^{-x^2-y^2})(1 - x^2 - y^2) + 2xe^{-x^2-y^2}(-2x) \\
&= 2e^{-x^2-y^2}\left((1 - 2x^2)(1 - x^2 - y^2) - 2x^2\right) \\
&= 2e^{-x^2-y^2}\left(1 - x^2 - y^2 - 2x^2 + 2x^4 + 2x^2y^2 - 2x^2\right) \\
&= 2e^{-x^2-y^2}\left(1 - 5x^2 - y^2 + 2x^4 + 2x^2y^2\right) \\
&= 2e^{-x^2-y^2}\left(1 - 4x^2 - (x^2 + y^2) + 2x^2(x^2 + y^2)\right), \\
\frac{\partial^2 z}{\partial x \partial y} &= (-2y)2xe^{-x^2-y^2}(1 - x^2 - y^2) + 2xe^{-x^2-y^2}(-2y) \\
&= -4xye^{-x^2-y^2}(2 - x^2 - y^2), \\
\frac{\partial^2 z}{\partial y^2} &= 2e^{-x^2-y^2}\left(1 - 4y^2 - (x^2 + y^2) + 2y^2(x^2 + y^2)\right) \\
&\qquad \text{(since } z \text{ is symmetrical in } x \text{ and } y\text{).}
\end{aligned}$$

If $x^2 + y^2 = 1$ then

$$\begin{aligned}
\frac{\partial^2 z}{\partial x^2} &= 2e^{-x^2-y^2}\left(1 - 4x^2 - (x^2 + y^2) + 2x^2(x^2 + y^2)\right) \\
&= 2e^{-1}\left(1 - 4x^2 - 1 + 2x^2\right) \\
&= -4x^2e^{-1}, \\
\frac{\partial^2 z}{\partial x \partial y} &= -4xye^{-x^2-y^2}(2 - x^2 - y^2) \\
&= -4xye^{-1}, \\
\frac{\partial^2 z}{\partial y^2} &= 2e^{-x^2-y^2}\left(1 - 4y^2 - (x^2 + y^2) + 2y^2(x^2 + y^2)\right) \\
&= 2e^{-1}\left(1 - 4y^2 - 1 + 2y^2\right) \\
&= -4y^2e^{-1}.
\end{aligned}$$

Thus, in this case,

$$\begin{aligned}
\Delta &= \frac{\partial^2 z}{\partial x^2}\frac{\partial^2 z}{\partial y^2} - \left(\frac{\partial^2 z}{\partial x \partial y}\right)^2 \\
&= \left(-4x^2e^{-1}\right)\left(-4y^2e^{-1}\right) - \left(-4xye^{-1}\right)^2 \\
&= 16x^2y^2e^{-2} - 16x^2y^2e^{-2} = 0.
\end{aligned}$$

Thus the Hessian gives us no information about the nature of the stationary points on the circle $x^2 + y^2 = 1$. To determine the nature of these stationary points, we consider the behaviour of z on the lines $y = ax$, where $a \in \mathbb{R}$. On such a line

$$\begin{aligned} z(x,y) &= z(x,ax) \\ &= (x^2 + a^2x^2)e^{-x^2-a^2x^2} \\ &= (a^2+1)x^2e^{-(a^2+1)x^2}. \end{aligned}$$

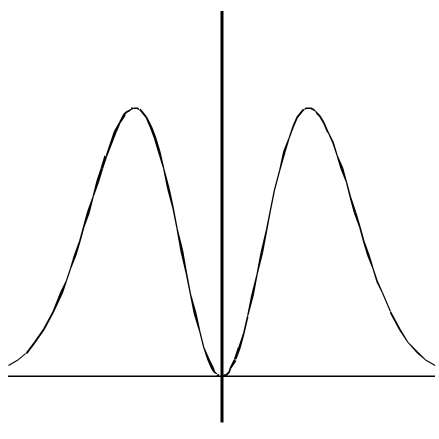
Let $g(x) = (a^2+1)x^2e^{-(a^2+1)x^2}$. We make the following observations about g.

- Since $a^2 + 1 > 0$, $x^2 \geq 0$ and $e^{-(a^2+1)x^2} > 0$, $g(x) \geq 0$ for all $x \in \mathbb{R}$. Furthermore, $g(x) = 0$ if and only if $x = 0$.

- As $x \to \pm\infty$, $g(x) \to 0$.

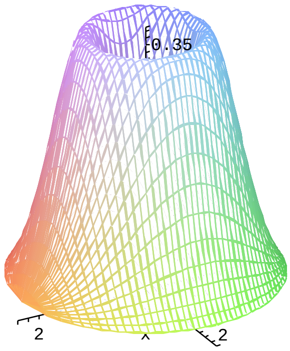
- The stationary points of g (i.e. the values x such that $g'(x) = 0$) are $x = 0$ and $x = \pm\frac{1}{\sqrt{a^2+1}}$.

From these observations we can deduce that $g(x)$ takes its minimum value at $x = 0$. The value of $g(x)$ then increases as we move away from $x = 0$ until g reaches a maximum at $x = \pm\frac{1}{\sqrt{a^2+1}}$; the value then decreases and approaches 0 as x approaches $\pm\infty$. The following diagram shows g for a typical value of a.



From this we can see that every point on the circle $x^2 + y^2 = 1$ is a local maximum of z. We also get that $(0,0)$ is a local minimum.

The following diagram illustrates the graph of f that clearly shows the local maxima on the circle $x^2 + y^2 = 1$.



Generalization to functions of several variables

We can generalize the idea of a stationary point to a function of several variables. Let $U \subseteq \mathbb{R}^n$ and let $f : U \to \mathbb{R}$. Then $(a_1, a_2, \ldots, a_n) \in U$ is a stationary point of f if

$$f_{x_1}(a_1, a_2, \ldots, a_n) = f_{x_2}(a_1, a_2, \ldots, a_n) = \cdots = f_{x_n}(a_1, a_2, \ldots, a_n) = 0.$$

It is also possible to classify the stationary points, but this is beyond the scope of the course.

Example 2.11. Find the stationary points of

$$f(x, y, z) = (x + y + z)^2 - 6xyz.$$

The partial derivatives of f are

$$\begin{aligned} f_x(x, y, z) &= 2(x + y + z) - 6yz, \\ f_y(x, y, z) &= 2(x + y + z) - 6xz, \\ f_z(x, y, z) &= 2(x + y + z) - 6xy. \end{aligned}$$

Putting $f_x(x, y, z) = f_y(x, y, z) = f_z(x, y, z) = 0$ gives

$$x + y + z = 3yz, \tag{9}$$
$$x + y + z = 3xz, \tag{10}$$
$$x + y + z = 3xy. \tag{11}$$

Combining equations (9) and (10) gives

$$3yz = 3xz.$$

Thus either $z = 0$ or $x = y$.

If $z = 0$ then equations (9)–(11) reduce to $x + y = 0$ and $x + y = 3xy$. From this we get $x = y = 0$, and so $(0, 0, 0)$ is a stationary point of f.

If $x = y$ then equations (9)–(11) reduce to $2x + z = 3xz$ and $2x + z = 3x^2$. From this we get $xz = x^2$, and so either $x = 0$ or $x = z$. If $x = 0$ we get $x = y = z = 0$, giving the stationary point found earlier. If $x = z$ and $x \neq 0$ then $2x + z = 3xz$ and $2x + z = 3x^2$ reduce to $x = x^2$, and so $x = 1$. This gives $x = y = z = 1$, and so $(1, 1, 1)$ is a stationary point of f.

Thus the stationary points of f are $(0, 0, 0)$ and $(1, 1, 1)$.

Exercises 2.2.

1. Find, and classify, the stationary points of
$$g(x, y) = \frac{1}{2}x^2y - 2xy + \frac{2}{3}y^3.$$
2. Determine the nature of the non-zero stationary points of the function f given in Example 2.9.
3. Find the stationary points of
$$f(x, y, z) = x^3 - 3x + y^3 - 3yz + 2z^2.$$

2.3 Local extrema vs global extrema

Let $U \subseteq \mathbb{R}^2$ and $f : U \to \mathbb{R}$. Recall that a point $(a, b) \in U$ is called a local maximum (resp. local minimum) of f if $f(a, b) \geq f(x, y)$ (resp. $f(a, b) \leq f(x, y)$) for all points (x, y) in some small disk around (a, b). We call (a, b) a local extremum if it is a local maximum or a local minimum.

We say that the point (a, b) is a **global maximum** (resp. **global minimum**) of f if $f(a, b) \geq f(x, y)$ (resp. $f(a, b) \leq f(x, y)$) for all $(x, y) \in U$. A global extremum is a point that is a global maximum or a global minimum. If (a, b) is a global maximum (resp. global minimum), then the value $f(a, b)$ is called the global maximum value (resp. global minimum value) or simply the maximum (resp. minimum) of f.[3]

In the previous section we discussed how to find the local extrema among the stationary points. We now consider the question under what conditions global extrema exist and how we can find them. Before answering this question for functions of two variables, we first look at the case of single variable functions. Figure 1 illustrates three different scenarios for the existence of global extrema.

The following theorem, though somewhat difficult to prove, is intuitively clear.

[3] Sometimes the point (a, b) and the value $f(a, b)$ are both called global maximum (resp. minimum). Obviously the terminology can be confusing, but it should always be clear from the context if we mean the point (a, b) or the value $f(a, b)$.

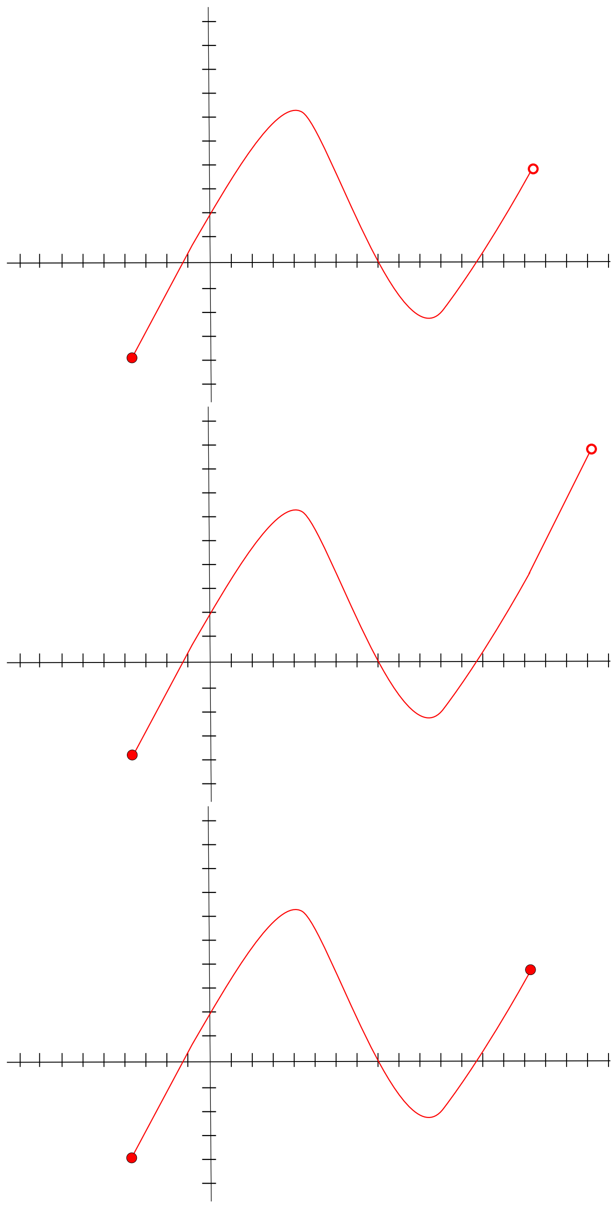


Figure 1: The first two functions are defined on an interval closed on the left and open on the right, whereas the third function is defined on a closed interval (closed on both sides). All three functions have the same minimum and it is attained at the left most point of the interval. In the first function, the maximum occurs at the left most stationary point. The second never attains its maximum and the third function has its maximum at the first stationary point. Notice, only in the third case does Theorem 2.12 guarantee that the global maximum and minimum values are attained.

Theorem 2.12 (Extreme Value Theorem, one variable case)**.** *If $[a, b]$ is a closed finite interval, and $f : [a, b] \longrightarrow \mathbb{R}$ is continuous, then f has a global maximum value and a global minimum value on its domain $[a, b]$.*

In fact, for a function $f : [a, b] \longrightarrow \mathbb{R}$ it is often easy to find where the global extrema occur. Find all stationary points of f in (a, b) and all points where f' is not defined (if f is everywhere differentiable, we don't need to do this part), and compare the values of f at these points with the values of f at the end points (i.e. at a and b). Of all these values, the smallest is the global minimum value of f on the interval and the largest is the global maximum value.

Example 2.13. Find the global maxima and minima of $f(x) = x^3 - x$ on the intervals (a) $[-1, 5]$ and (b) $(-1, 5)$.

The function f is both continuous and differentiable everywhere. The first interval is a closed finite interval, so we are guaranteed to find a global maximum and minimum. We first find the stationary points. We see that $f'(x) = 3x^2 - 1$, so $f'(x) = 0$ if and only if $x = \pm\frac{1}{\sqrt{3}}$. Evaluating f at those points gives $f(-\frac{1}{\sqrt{3}}) = \frac{2}{3\sqrt{3}} \approx 0.38$ and $f(\frac{1}{\sqrt{3}}) = -\frac{2}{3\sqrt{3}} \approx -0.38$. Evaluating f on the endpoints of the interval gives $f(-1) = 0$ and $f(5) = 120$. Thus, the minimum of f on the interval $[-1, 5]$ is $-\frac{2}{3\sqrt{3}}$, which occurs at the stationary point $x = \frac{1}{\sqrt{3}}$, and the maximum is 120, which occurs at the endpoint $x = 5$.

For the second interval, the minimum of f is still $-\frac{2}{3\sqrt{3}}$. However, since 5 is not in the second interval, there is no maximum. Thus, on the second interval, f achieves its global minimum, but not its global maximum. That is, there is no $c \in (-1, 5)$ such that $f(c) \geq f(x)$ for all $x \in (-1, 5)$.

The extreme value theorem in single variable calculus has a generalisation to multivariable calculus, which we will discuss only in the case of functions of two variables. First we need the notions of a closed and bounded set in $\mathbb{R}^2$. Consider a set $S \subseteq \mathbb{R}^2$. A point $b \in \mathbb{R}^2$ is a **boundary point** of S if every disk around b, no matter how small, contains both points in S and outside S. The **boundary** of a set S is the set of all its boundary points, and S is **closed** if it contains its boundary.

Example 2.14.

- The boundary of the set $\{(x, y) : x^2 + y^2 < 1\}$ is $\{(x, y) : x^2 + y^2 = 1\}$. Thus, the set $\{(x, y) : x^2 + y^2 < 1\}$ does not contain its boundary and is therefore not closed.

- The set $\{(x, y) : x^2 + y^2 \leq 1\}$ also has the set $\{(x, y) : x^2 + y^2 = 1\}$ as its boundary. Hence $\{(x, y) : x^2 + y^2 \leq 1\}$ contains its boundary and is therefore closed. More generally, a disk $D_r = \{(x, y) : x^2 + y^2 \leq r^2\}$ of any finite radius r is closed.

- The graph of a function with domain $\mathbb{R}$ (like $\{(x, y) : y = 2x - 1, x \in \mathbb{R}\}$ or $\{(x, y) : y = x^2, x \in \mathbb{R}\}$) is closed.

A set is **bounded** if it is contained in some finite disk; that is, it fully lies in some disk $D_r = \{(x, y) : x^2 + y^2 \leq r^2\}$ for some constant r. For example, the infinite line $\{(x, y) : y = 2x - 1, x \in \mathbb{R}\}$ is not bounded (but it is closed), but the finite line segment $\{(x, y) : y = 2x - 1, x \in [0, 10]\}$ is bounded (and closed, since $[0, 10]$ is a closed interval). The set given by the circle $\{(x, y) : x^2 + y^2 = 4\}$ is bounded (and closed), as is the disk $\{(x, y) : x^2 + y^2 \leq 1\}$. The extreme value theorem is as follows.

Theorem 2.15 (Extreme Value Theorem, two variables case)**.** *If U is a closed, bounded subset of $\mathbb{R}^2$, and $f : U \to \mathbb{R}$ is continuous, then f has a global maximum value and a global minimum value on its domain U.*

Notice that we haven't discussed continuity in $\mathbb{R}^2$, but the well-behaved functions we have seen so far are continuous. In order to find extreme values of a continuous function f on a closed, bounded set U we do the following:

1. Find all stationary points of f in the interior of U (i.e. the stationary points of f that are in U but not on the boundary of U).
2. Find all points in the interior of U where the tangent plane does not exist. (This step can be omitted in our examples because we only consider well-behaved functions that have tangent planes everywhere.)
3. Find all extreme values on the boundary of U. (We will study one method for doing this in the next section.)
4. Compare all the values you have found (i.e. the values of f at the points found in 1. and 2. and the extreme values found in 3.). The largest one is the global maximum value of f on U and the smallest one is the global minimum value.

Example 2.16. Find the maximum and minimum values of $f(x, y) = 3x^2+2xy+3y^2$ on the set $U = \{(x, y) : x^2 + y^2 \leq 1\}$.

Notice that the set U is closed and bounded and hence f has global extreme values. We first find the stationary points of f. From $f_x = 6x+2y = 0$ and $f_y = 2x+6y = 0$ we deduce that the only stationary point of f is $(0, 0)$ (which does lie in the interior of U). Note that $f(0, 0) = 0$.

Next we want to find the extreme values of f on the boundary of U. Since this boundary is the circle $\{(x, y) : x^2 + y^2 = 1\}$, we can parametrize it as $(x(t), y(t)) = (\cos t, \sin t)$ with $t \in [0, 2\pi]$ (for $t = 0$ and $t = 2\pi$ we will get the same point on the circle, however this won't be a problem). Let $F(t) = f(x(t), y(t))$. Then

$$\begin{aligned} F(t) &= 3\cos^2 t + 2\cos t \sin t + 3\sin^2 t \\ &= 3 + \sin(2t) \end{aligned}$$

(using the identities $\cos^2 t+\sin^2 t = 1$ and $2\cos t \sin t = \sin(2t)$). The extreme values of f on the boundary of U are the same as the extreme values of F on the interval

$[0, 2\pi]$. The stationary points t of F in this interval can be found as follows:

$$\begin{aligned} F'(t) = 0 \text{ and } t \in [0, 2\pi] &\Leftrightarrow 2\cos(2t) = 0 \text{ and } t \in [0, 2\pi] \\ &\Leftrightarrow 2t = \pm\frac{\pi}{2}, \pm\frac{3\pi}{2}, \pm\frac{5\pi}{2}, \ldots \text{ and } t \in [0, 2\pi] \\ &\Leftrightarrow t = \frac{\pi}{4}, \frac{3\pi}{4}, \frac{5\pi}{4}, \frac{7\pi}{4}. \end{aligned}$$

The value of F at these points is $F(\frac{\pi}{4}) = F(\frac{5\pi}{4}) = 4$ and $F(\frac{3\pi}{4}) = F(\frac{7\pi}{4}) = 2$, and the value of F at the end points of the interval $[0, 2\pi]$ is $F(0) = F(2\pi) = 3$. Thus the minimum value of F on $[0, 2\pi]$ is 2 and the maximum value is 4.

Comparing the local extreme values f in the interior of U and the extreme values of f on the boundary of U, we deduce that the maximum value of f on U is 4 (attained at the two points $(\cos\frac{\pi}{4}, \sin\frac{\pi}{4}) = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$ and $(\cos\frac{5\pi}{4}, \sin\frac{5\pi}{4}) = (-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}))$ and the minimum value is 0 (attained at the point $(0, 0)$).

Exercises 2.3.

1. Find the global maximum and minimum of $f(x) = x + \sin x$ in the intervals (a) $[0, 5]$ and (b) $(0, 5]$.

2. Find the boundary of the set $U = \{(x, y) : x > 0 \text{ and } y > 0 \text{ and } xy \leq 1\} \subseteq \mathbb{R}^2$. Is U closed? Is U bounded?

2.4 Lagrange multipliers

In the previous section we encountered the problem of finding the extrema of $f(x, y) = 3x^2 + 2xy + 3y^2$ on the set $\{(x, y) : x^2 + y^2 = 1\}$. Note that the set $\{(x, y) : x^2 + y^2 = 1\}$ can be described as the set of points (x, y) satisfying the condition $g(x, y) = 0$ where $g(x, y) = x^2 + y^2 - 1$.

More generally, let $U \subseteq \mathbb{R}^2$ and $f, g : U \to \mathbb{R}$. Suppose we want to find the maximum and minimum of $f(x, y)$ subject to the **constraint** $g(x, y) = 0$. In the previous section we did this by parameterizing the set $\{(x, y) : g(x, y) = 0\}$ and thus reducing the problem to finding the maximum and minimum of a function of one variable. However it is often difficult (or just algebraically messy) to parametrize this set, and in such cases the **method of Lagrange multipliers** can be useful to find the extrema of $f(x, y)$ subject to the constraint $g(x, y) = 0$.

Given $f(x, y)$ and $g(x, y)$, we define the **Lagrange function** $L(x, y, \lambda)$ by

$$L(x, y, \lambda) = f(x, y) - \lambda g(x, y),$$

where λ is a new variable. One can show that if a point (x_0, y_0) is a local maximum or minimum of $f(x, y)$ subject to the constraint $g(x, y) = 0$, then there exists a λ_0

such that (x_0, y_0, λ_0) is a stationary point of $L(x, y, \lambda)$.[4] The variable λ is called a **Lagrange multiplier**.

Using the Lagrange function we can find the maximum and minimum of $f(x, y)$ subject to $g(x, y) = 0$ as follows. We first find all stationary points (x_0, y_0, λ_0) of $L(x, y, \lambda)$ and form the points (x_0, y_0) by omitting the last coordinate (so we don't even have to compute the last coordinate λ_0 when finding the stationary points). For each of these points compute $f(x_0, y_0)$. The largest of these values will be the maximum and the smallest will be the minimum.

Example 2.17. Find the maximum and minimum values of $f(x, y) = 3x^2+2xy+3y^2$ subject to the constraint $g(x, y) = 0$ where $g(x, y) = x^2 + y^2 - 1$.

This is the same problem as in the previous section, but now we want to use the method of Lagrange multipliers to solve it. The Lagrange function is

$$L(x, y, \lambda) = 3x^2 + 2xy + 3y^2 - \lambda(x^2 + y^2 - 1).$$

To compute the stationary point of $L(x, y, \lambda)$ we need to compute the partial derivatives:

$$\begin{aligned} L_x &= 6x + 2y - 2\lambda x, \\ L_y &= 2x + 6y - 2\lambda y, \\ L_\lambda &= -(x^2 + y^2 - 1). \end{aligned}$$

Setting $L_x = L_y = L_\lambda = 0$ gives

$$\begin{aligned} 6x + 2y - 2\lambda x &= 0, \\ 2x + 6y - 2\lambda y &= 0, \\ -(x^2 + y^2 - 1) &= 0. \end{aligned}$$

Multiplying the first equation by y and the second equation by x and then taking the difference gives $2y^2 - 2x^2 = 0$ and hence $x^2 = y^2$. Now from $x^2 = y^2$ and $x^2 + y^2 - 1 = 0$ we deduce $2x^2 = 1$ and hence $x = \pm\frac{\sqrt{2}}{2}$. For each of these values of x we obtain $y^2 = x^2 = \frac{1}{2}$ and hence $y = \pm\frac{\sqrt{2}}{2}$. Thus the stationary points are

$$(\tfrac{\sqrt{2}}{2}, \tfrac{\sqrt{2}}{2}), (-\tfrac{\sqrt{2}}{2}, -\tfrac{\sqrt{2}}{2}), (-\tfrac{\sqrt{2}}{2}, \tfrac{\sqrt{2}}{2}), (\tfrac{\sqrt{2}}{2}, -\tfrac{\sqrt{2}}{2}).$$

Note that we did not even compute the λ-coordinate for these stationary points as it is not required. Evaluating f at those four points gives

$$f(\tfrac{\sqrt{2}}{2}, \tfrac{\sqrt{2}}{2}) = f(-\tfrac{\sqrt{2}}{2}, -\tfrac{\sqrt{2}}{2}) = 4, \quad f(-\tfrac{\sqrt{2}}{2}, \tfrac{\sqrt{2}}{2}) = f(\tfrac{\sqrt{2}}{2}, -\tfrac{\sqrt{2}}{2}) = 2.$$

Hence the maximum value is 4 and the minimum value is 2.

[4] To make this statement precise, f and g also need to satisfy some further technical assumptions, however for our purposes these details are not important.

Example 2.18. Find the point on the line $3x+2y=5$ that is closest to the point $(3,1)$.

The distance between a general point (x,y) and the point $(3,1)$ is

$$\sqrt{(x-3)^2+(y-1)^2}.$$

We want to find the minimum value of this distance subject to the constraint $3x+2y-5=0$. In fact, it is easier to minimize the square of the distance, and so we minimize $f(x,y)=(x-3)^2+(y-1)^2$, subject to the given constraint.

Let $L(x,y,\lambda)=(x-3)^2+(y-1)^2-\lambda(3x+2y-5)$. Then

$$\begin{aligned}\frac{\partial L}{\partial x} &= 2(x-3)-3\lambda,\\ \frac{\partial L}{\partial y} &= 2(y-1)-2\lambda,\\ \frac{\partial L}{\partial \lambda} &= -3x-2y+5.\end{aligned}$$

Putting $\frac{\partial L}{\partial x}=\frac{\partial L}{\partial y}=\frac{\partial L}{\partial \lambda}=0$ gives

$$\begin{aligned}2(x-3)-3\lambda &= 0, &(12)\\ 2(y-1)-2\lambda &= 0, &(13)\\ 3x+2y &= 5. &(14)\end{aligned}$$

Multiplying equation (12) by 2, and equation (13) by 3 gives

$$\begin{aligned}4(x-3)-6\lambda &= 0,\\ 6(y-1)-6\lambda &= 0.\end{aligned}$$

Thus $4(x-3)=6(y-1)$. Multiplying out the brackets in this equation, and simplifying the result gives

$$2x-3y = 3. \qquad (15)$$

Multiplying equation (14) by 3, and equation (15) by 2 gives

$$\begin{aligned}9x+6y &= 15,\\ 4x-6y &= 6.\end{aligned}$$

Adding the last two equations we get $13x=21$, and so $x=\frac{21}{13}$. Using equation (15) with $x=\frac{21}{13}$ we get $y=\frac{1}{13}$. Thus the point $(\frac{21}{13},\frac{1}{13})$ on the line $3x+2y=5$ is closest to the point $(3,1)$.

Although not required, the distance of the point $(\frac{21}{13},\frac{1}{13})$ to the point $(3,1)$ is

$$\begin{aligned}\sqrt{f\left(\frac{21}{13},\frac{1}{13}\right)} &= \sqrt{\left(\frac{21}{13}-3\right)^2+\left(\frac{1}{13}-1\right)^2}\\ &= \sqrt{\frac{36}{13}}=\frac{6}{\sqrt{13}}.\end{aligned}$$

Finally we remark that there is no point on the line $3x+2y=5$ that has the largest distance from the point $(3,1)$. In other words, the function $f(x,y)$ has no maximum value subject to the constraint $3x+2y-5=0$. However this is not a contradiction to the extreme value theorem because the set $\{(x,y) : 3x+2y-5=0\}$ is not bounded.

Exercises 2.4.

1. Find the maximum and minimum values of $f(x,y)=xy$ on the ellipse $18x^2+2y^2=25$.

2. Find the maximum and minimum values of $2x+y$ on the ellipse $x^2+xy+4y^2+2x+16y+7=0$.

3. Find the maximum and minimum values of xy^2 on the circle $x^2+y^2=1$.

2.5 Lagrange multipliers with more than one constraint

Let $U \subseteq \mathbb{R}^n$, and let $f, g_1, g_2, \ldots, g_m : U \to \mathbb{R}$. To find the maximum and minimum values of $f(x_1, x_2, \ldots, x_n)$, subject to the constraints

$$\begin{aligned} g_1(x_1, x_2, \ldots, x_n) &= 0, \\ g_2(x_1, x_2, \ldots, x_n) &= 0, \\ &\vdots \\ g_m(x_1, x_2, \ldots, x_n) &= 0, \end{aligned}$$

we find the stationary points of

$$\begin{aligned} L(x_1, x_2, \ldots, x_n, \lambda_1, \lambda_2, \ldots, \lambda_m) = f(x_1, x_2, \ldots, x_n) &- \lambda_1 g_1(x_1, x_2, \ldots, x_n) \\ &- \lambda_2 g_2(x_1, x_2, \ldots, x_n) - \cdots - \lambda_m g_m(x_1, x_2, \ldots, x_n). \end{aligned}$$

The introduced variables $\lambda_1, \lambda_2, \ldots, \lambda_m$ are called **Lagrange multipliers**.

Example 2.19. Find the maximum and minimum values of $x+z$ at the points where the plane $4x+y+z=34$ and the cylinder $x^2+y^2=40$ intersect.

Here we want to find the maximum and minimum values of $f(x,y,z)=x+z$, subject to the constraints $g_1(x,y,z)=0$ and $g_2(x,y,z)=0$, where $g_1(x,y,z)=4x+y+z-34$ and $g_2(x,y,z)=x^2+y^2-40$. Let

$$L(x,y,z,\lambda,\mu) = x+z-\lambda(4x+y+z-34)-\mu(x^2+y^2-40).$$

Then

$$\begin{aligned}
\frac{\partial L}{\partial x} &= 1-4\lambda-2\mu x,\\
\frac{\partial L}{\partial y} &= -\lambda-2\mu y,\\
\frac{\partial L}{\partial z} &= 1-\lambda,\\
\frac{\partial L}{\partial \lambda} &= -(4x+y+z-34),\\
\frac{\partial L}{\partial \mu} &= -(x^2+y^2-40).
\end{aligned}$$

Putting $\frac{\partial L}{\partial x}=\frac{\partial L}{\partial y}=\frac{\partial L}{\partial z}=\frac{\partial L}{\partial \lambda}=\frac{\partial L}{\partial \mu}=0$ gives

$$\begin{aligned}
4\lambda+2\mu x &= 1, && (16)\\
\lambda+2\mu y &= 0, && (17)\\
\lambda &= 1, && (18)\\
4x+y+z &= 34, && (19)\\
x^2+y^2 &= 40. && (20)
\end{aligned}$$

From equation (18), $\lambda=1$. Substituting this value of λ into equations (16) and (17) gives $2\mu x=-3$ and $2\mu y=-1$, respectively. Multiplying the second of these equations by 3 gives $6\mu y=-3$, and so

$$2\mu x=-3=6\mu y.$$

Now $\mu\neq 0$ (otherwise equation (17) gives $\lambda=0$, contrary to $\lambda=1$) and so $x=3y$. Substituting $x=3y$ into equation (20) gives $10y^2=40$. Thus $y^2=4$, and so $y=\pm 2$. For $y=2$ we get $x=6$, and for $y=-2$ we get $x=-6$. From equation (19),

$$z=34-4x-y.$$

Thus when $y=2$ and $x=6$ then $z=8$, and when $y=-2$ and $x=-6$ then $z=60$. Now $f(6,2,8)=14$ and $f(-6,-2,60)=54$. Thus the maximum and minimum values of $x+z$ at the points where the plane $4x+y+z=34$ and the cylinder $x^2+y^2=40$ intersect are 54 and 14, respectively.

Example 2.20. Find the maximum and minimum values of $f(x,y,z)=xy+4z$ subject to the constraints $x+y+z=0$ and $x^2+y^2+z^2=24$.

Let

$$L(x,y,z,\lambda,\mu)=xy+4z-\lambda(x+y+z)-\mu(x^2+y^2+z^2-24).$$

Then

$$\begin{aligned}
\frac{\partial L}{\partial x} &= y - \lambda - 2\mu x, \\
\frac{\partial L}{\partial y} &= x - \lambda - 2\mu y, \\
\frac{\partial L}{\partial z} &= 4 - \lambda - 2\mu z, \\
\frac{\partial L}{\partial \lambda} &= -(x + y + z), \\
\frac{\partial L}{\partial \mu} &= -(x^2 + y^2 + z^2 - 24).
\end{aligned}$$

Putting $\frac{\partial L}{\partial x} = \frac{\partial L}{\partial y} = \frac{\partial L}{\partial z} = \frac{\partial L}{\partial \lambda} = \frac{\partial L}{\partial \mu} = 0$ gives

$$\begin{aligned}
y - 2\mu x &= \lambda, && (21) \\
x - 2\mu y &= \lambda, && (22) \\
4 - 2\mu z &= \lambda, && (23) \\
x + y + z &= 0, && (24) \\
x^2 + y^2 + z^2 &= 24. && (25)
\end{aligned}$$

From equations (21) and (22) we get

$$y - 2\mu x = \lambda = x - 2\mu y.$$

Thus

$$y - x = 2\mu(x - y),$$

and so either $x = y$ or $\mu = -\frac{1}{2}$. We now consider these two possibilities separately.

If $x = y$ then equation (24) gives $2x + z = 0$, and so $z = -2x$. Then equation (25) gives $x^2 + x^2 + 4x^2 = 24$, and so $x^2 = 4$. Hence $x = \pm 2$, and so

$$x = \pm 2, y = \pm 2, z = \mp 4.$$

If $\mu = -\frac{1}{2}$ then equations (21) and (23) give $x + y = \lambda$ and $z + 4 = \lambda$, respectively. Thus

$$x + y = \lambda = z + 4,$$

and so

$$x + y - z = 4. \qquad (26)$$

Subtracting equation (26) from equation (24) gives $2z = -4$, and so $z = -2$. Adding equations (26) and (24) gives $2(x + y) = 4$, and so $y = 2 - x$. Putting $y = 2 - x$ and $z = -2$ into equation (25) we get

$$\begin{aligned}
x^2 + (2 - x)^2 + (-2)^2 = 24 &\Leftrightarrow x^2 + (4 - 4x + x^2) + 4 = 24 \\
&\Leftrightarrow x^2 - 2x - 8 = 0 \\
&\Leftrightarrow (x - 4)(x + 2) = 0 \\
&\Leftrightarrow x = -2 \text{ or } 4.
\end{aligned}$$

When $x = -2$ then $y = 4$ and $z = -2$, and when $x = 4$ then $y = -2$ and $z = -2$.

From the preceding calculations we have found four stationary points of L, namely

$$(2, 2, -4), (-2, -2, 4), (-2, 4, -2) \text{ and } (4, -2, -2).$$

Now $f(2, 2, -4) = -12$, $f(-2, -2, 4) = 20$, $f(-2, 4, -2) = -16$ and $f(4, -2, -2) = -16$. Thus the maximum and minimum values of $f(x, y, z)$, subject to the given constraints, are 20 and -16, respectively.

Exercises 2.5.

1. Find the minimum distance from the origin to a point on the curve where the plane $2y + 4z = 15$ and the surface $z^2 = 4(x^2 + y^2)$ intersect.

2.6 The chain rule

In this section we generalize the chain rule to functions of more than one variable. First we extend our idea of a real function.

Multivariable vector-valued functions

Let $U \subseteq \mathbb{R}^n$. A **(multivariable vector-valued) function** $f : U \to \mathbb{R}^m$ determines for each point $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ of U a unique point $\mathbf{y} = (y_1, y_2, \ldots, y_m)$ of $\mathbb{R}^m$. We write $f(x_1, x_2, \ldots, x_n) = (y_1, y_2, \ldots, y_m)$ or $f(\mathbf{x}) = \mathbf{y}$.

Each coordinate of $\mathbf{y}$ is uniquely determined by $(x_1, x_2, \ldots, x_n)$. That is, each coordinate of $\mathbf{y}$ can be thought of as a function of $(x_1, x_2, \ldots, x_n)$. So there are functions $f_1, f_2, \ldots, f_m : U \to \mathbb{R}$ such that

$$f(x_1, x_2, \ldots, x_n) = (f_1(x_1, x_2, \ldots, x_n), f_2(x_1, x_2, \ldots, x_n), \ldots, f_m(x_1, x_2, \ldots, x_n)),$$

or $f(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_m(\mathbf{x}))$.

Example 2.21. Let $f : \mathbb{R}^2 \to \mathbb{R}^2$, with $f(x, y) = (x^2y, x + 3y)$. Here $f_1(x, y) = x^2y$ and $f_2(x, y) = x + 3y$. An alternative way of defining f is to write $f(x, y) = (f_1, f_2)$ where $f_1(x, y) = x^2y$ and $f_2(x, y) = x + 3y$.

Example 2.22. Let $f : \mathbb{R} \to \mathbb{R}^3$, with $f(t) = (t, \cos t, \sin t)$. Functions of one variable from $\mathbb{R}$ to $\mathbb{R}^m$, like this one (that has $m = 3$), define a **parametric** curve in $\mathbb{R}^m$. If we plot the function f as t varies we obtain a spiral going round the x-axis.

Example 2.23. Let $f : \mathbb{R}^3 \to \mathbb{R}^2$, with $f(x, y, z) = (u, v)$, where $u = x + y - z$ and $v = 2x - y - 2z$.

Let $U \subseteq \mathbb{R}^n$, and let $f : U \to \mathbb{R}^m$ with $f(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_m(\mathbf{x}))$. Then the **derivative** of f at $\mathbf{x}$, denoted by $f'(\mathbf{x})$ or $\dfrac{\mathrm{d}f}{\mathrm{d}\mathbf{x}}$, is the $m \times n$ matrix whose (i,j)th entry is $\dfrac{\partial f_i}{\partial x_j}$. That is,

$$f'(\mathbf{x}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}.$$

Example 2.24. Let $f : \mathbb{R}^2 \to \mathbb{R}^2$, with $f(x,y) = (x^2y, x+3y)$. Here $f_1(x,y) = x^2y$ and $f_2(x,y) = x+3y$, and

$$f'(x,y) = \begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{pmatrix} = \begin{pmatrix} 2xy & x^2 \\ 1 & 3 \end{pmatrix}.$$

Example 2.25. Let $f : \mathbb{R} \to \mathbb{R}^3$, with $f(t) = (t, \cos t, \sin t)$. Then

$$f'(t) = \begin{pmatrix} \frac{\mathrm{d}}{\mathrm{d}t}(t) \\ \frac{\mathrm{d}}{\mathrm{d}t}(\cos t) \\ \frac{\mathrm{d}}{\mathrm{d}t}(\sin t) \end{pmatrix} = \begin{pmatrix} 1 \\ -\sin t \\ \cos t \end{pmatrix}.$$

Example 2.26. Let $f : \mathbb{R}^3 \to \mathbb{R}^2$, with $f(x,y,z) = (u,v)$, where $u = x+y-z$ and $v = 2x - y - 2z$. Then

$$f'(x,y,z) = \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 \\ 2 & -1 & -2 \end{pmatrix}.$$

The aim of the chain rule

Recall the chain rule for single variable functions. Let $g : \mathbb{R} \to \mathbb{R}$ and $f : \mathbb{R} \to \mathbb{R}$ be functions, and let $F = f \circ g$ be their composition, i.e. $F : \mathbb{R} \to \mathbb{R}$ is the function

$$F(x) = f(g(x)).$$

Then the chain rule expresses the derivative of F in terms of the derivatives of f and g:

$$F'(x) = f'(g(x))g'(x).$$

We want to generalize this to multivariable vector-valued functions. Let $g : \mathbb{R}^n \to \mathbb{R}^m$ and $f : \mathbb{R}^m \to \mathbb{R}^k$ be functions, and let $F = f \circ g$ be their composition, i.e. $F : \mathbb{R}^n \to \mathbb{R}^k$ is the function

$$F(\mathbf{x}) = f(g(\mathbf{x})).$$

The aim of the chain rule is to express the partial derivatives of F in terms of the partial derivatives of f and g. Before stating the general result, we consider some important special cases.

To simplify the notation we state all results in this section only for functions with domain $\mathbb{R}^n$ for some n. However everything remains true if we consider functions $g : U \to \mathbb{R}^m$ where $U \subseteq \mathbb{R}^n$ and $f : V \to \mathbb{R}^k$ where $V \subseteq \mathbb{R}^m$ as long as $g(\mathbf{x}) \in V$ for all $\mathbf{x} \in U$.

Some special cases of the chain rule

In this section we consider three simple special cases of the chain rule, and illustrate them with examples.

Case 1: $\mathbb{R}^2 \xrightarrow{g} \mathbb{R} \xrightarrow{f} \mathbb{R}$

Let $g : \mathbb{R}^2 \to \mathbb{R}$ and $f : \mathbb{R} \to \mathbb{R}$, and let $F : \mathbb{R}^2 \to \mathbb{R}$ with

$$F(x, y) = (f \circ g)(x, y) = f(g(x, y)).$$

Then

$$\begin{aligned} \frac{\partial F}{\partial x} &= \frac{\mathrm{d}f}{\mathrm{d}g}\frac{\partial g}{\partial x}, \\ \frac{\partial F}{\partial y} &= \frac{\mathrm{d}f}{\mathrm{d}g}\frac{\partial g}{\partial y}. \end{aligned}$$

Example 2.27. Let $F(x, y) = (x^2 + 3xy - y^2)^4$. Then $F(x, y) = f(g)$ where $g = x^2 + 3xy - y^2$ and $f = g^4$. Thus

$$\begin{aligned} \frac{\partial F}{\partial x} &= \frac{\mathrm{d}f}{\mathrm{d}g}\frac{\partial g}{\partial x} \\ &= 4g^3(2x + 3y) \\ &= 4(x^2 + 3xy - y^2)^3(2x + 3y), \\ \frac{\partial F}{\partial y} &= \frac{\mathrm{d}f}{\mathrm{d}g}\frac{\partial g}{\partial y} \\ &= 4g^3(3x - 2y) \\ &= 4(x^2 + 3xy - y^2)^3(3x - 2y). \end{aligned}$$

Example 2.28. Let $f : \mathbb{R} \to \mathbb{R}$ and let $z = f(x^2 + y^2)$. Show that

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

Let $g(x, y) = x^2 + y^2$. Then $z(x, y) = f(g(x, y))$. Thus

$$\begin{aligned} \frac{\partial z}{\partial x} &= \frac{\mathrm{d}f}{\mathrm{d}g}\frac{\partial g}{\partial x} = 2x\frac{\mathrm{d}f}{\mathrm{d}g}, \\ \frac{\partial z}{\partial y} &= \frac{\mathrm{d}f}{\mathrm{d}g}\frac{\partial g}{\partial y} = 2y\frac{\mathrm{d}f}{\mathrm{d}g}. \end{aligned}$$

Hence

$$y\frac{\partial z}{\partial x} = 2xy\frac{\mathrm{d}f}{\mathrm{d}g} = x\frac{\partial z}{\partial y}.$$

Case 2: $\mathbb{R} \xrightarrow{g} \mathbb{R}^2 \xrightarrow{f} \mathbb{R}$

Let $g : \mathbb{R} \to \mathbb{R}^2$, with $g(t) = (u(t), v(t))$, and $f : \mathbb{R}^2 \to \mathbb{R}$, and let $F : \mathbb{R} \to \mathbb{R}$ with

$$F(t) = (f \circ g)(t) = f(g(t)).$$

Then

$$\frac{\mathrm{d}F}{\mathrm{d}t} = \frac{\partial f}{\partial u}\frac{\mathrm{d}u}{\mathrm{d}t} + \frac{\partial f}{\partial v}\frac{\mathrm{d}v}{\mathrm{d}t}.$$

Example 2.29. Let $F(t) = u^2 + v^2$, where $u = t\cos t$ and $v = \sin t$. Then $F(t) = f(g(t))$ where $g(t) = (u(t), v(t))$ and $f(u, v) = u^2 + v^2$. Thus

$$\begin{aligned}
\frac{\mathrm{d}F}{\mathrm{d}t} &= \frac{\partial f}{\partial u}\frac{\mathrm{d}u}{\mathrm{d}t} + \frac{\partial f}{\partial v}\frac{\mathrm{d}v}{\mathrm{d}t} \\
&= 2u(\cos t - t\sin t) + 2v\cos t \\
&= 2t\cos t(\cos t - t\sin t) + 2\sin t\cos t \\
&= 2\cos t(t\cos t - t^2\sin t + \sin t).
\end{aligned}$$

Case 3: $\mathbb{R}^2 \xrightarrow{g} \mathbb{R}^2 \xrightarrow{f} \mathbb{R}$

Let $g : \mathbb{R}^2 \to \mathbb{R}^2$, with $g(x, y) = (u(x, y), v(x, y))$, and $f : \mathbb{R}^2 \to \mathbb{R}$, and let $F : \mathbb{R}^2 \to \mathbb{R}$ with

$$F(x, y) = (f \circ g)(x, y) = f(g(x, y)).$$

Then

$$\begin{aligned}
\frac{\partial F}{\partial x} &= \frac{\partial f}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial x}, \\
\frac{\partial F}{\partial y} &= \frac{\partial f}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial y}.
\end{aligned}$$

Example 2.30. Let $u = x^2y$ and $v = xy^3$ and let $F(x, y) = 2u + v^2$. Then $F(x, y) = f(g(x, y))$ where $f(u, v) = 2u + v^2$ and $g(x, y) = (u, v) = (x^2y, xy^3)$. Thus

$$\begin{aligned}
\frac{\partial F}{\partial x} &= \frac{\partial f}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial x} \\
&= 2(2xy) + 2vy^3 \\
&= 4xy + 2xy^6 \\
&= 2xy(2 + y^5), \\
\frac{\partial F}{\partial y} &= \frac{\partial f}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial y} \\
&= 2x^2 + 2v(3xy^2) \\
&= 2x^2 + 6x^2y^5 \\
&= 2x^2(1 + 3y^5).
\end{aligned}$$

Example 2.31. Let f be a function of two variables x and y, with $x = r\cos\theta$ and $y = r\sin\theta$. Find $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ in terms of r, θ, $\frac{\partial f}{\partial r}$ and $\frac{\partial f}{\partial \theta}$.

Using the chain rule gives

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

Similarly,

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

Thus

$$\frac{\partial f}{\partial r} = \cos\theta\frac{\partial f}{\partial x} + \sin\theta\frac{\partial f}{\partial y}, \tag{27}$$

$$\frac{\partial f}{\partial \theta} = -r\sin\theta\frac{\partial f}{\partial x} + r\cos\theta\frac{\partial f}{\partial y}. \tag{28}$$

Multiplying equation (27) by $r\cos\theta$ and equation (28) by $\sin\theta$ gives

$$r\cos\theta\frac{\partial f}{\partial r} = r\cos^2\theta\frac{\partial f}{\partial x} + r\cos\theta\sin\theta\frac{\partial f}{\partial y}, \tag{29}$$

$$\sin\theta\frac{\partial f}{\partial \theta} = -r\sin^2\theta\frac{\partial f}{\partial x} + r\cos\theta\sin\theta\frac{\partial f}{\partial y}. \tag{30}$$

Subtracting equation (30) from equation (29) gives

$$r\cos\theta\frac{\partial f}{\partial r} - \sin\theta\frac{\partial f}{\partial \theta} = r(\cos^2\theta + \sin^2\theta)\frac{\partial f}{\partial x}.$$

Now, since $\cos^2\theta + \sin^2\theta = 1$, we get

$$\frac{\partial f}{\partial x} = \cos\theta\frac{\partial f}{\partial r} - \frac{\sin\theta}{r}\frac{\partial f}{\partial \theta}.$$

Using a similar method, or by substituting for $\frac{\partial f}{\partial x}$ in either equation (27) or equation (28), we get

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

The general case of the chain rule

We now state the general form of the chain rule.

Theorem 2.32. *Let $g : \mathbb{R}^n \to \mathbb{R}^m$, $f : \mathbb{R}^m \to \mathbb{R}^k$ and $F = f \circ g : \mathbb{R}^n \to \mathbb{R}^k$. Then*

$$F'(\mathbf{x}) = f'(g(\mathbf{x}))g'(\mathbf{x}).$$

Here the equation $F'(\mathbf{x}) = f'(g(\mathbf{x}))g'(\mathbf{x})$ is an equation of matrices. In the alternative notation for the derivative it becomes $\frac{\mathrm{d}F}{\mathrm{d}\mathbf{x}} = \frac{\mathrm{d}f}{\mathrm{d}g}\frac{\mathrm{d}g}{\mathrm{d}\mathbf{x}}$. By looking at individual entries of the matrix $F'(\mathbf{x})$ we obtain expressions for the partial derivatives of F in terms of the partial derivatives of f and g.

Application: $\mathbb{R}^m \xrightarrow{(u_1,u_2,\dots,u_n)} \mathbb{R}^n \xrightarrow{h} \mathbb{R}$

Suppose $u_1, u_2, \dots, u_n$ are functions of m variables $x_1, x_2, \dots, x_m$, and h is a function of n variables. Let $H(x_1, x_2, \dots, x_m) = h(u_1, u_2, \dots, u_n)$. Then H is a function of $x_1, x_2, \dots, x_m$ and Theorem 2.32 gives the equation of matrices

$$\begin{pmatrix} \frac{\partial H}{\partial x_1} & \frac{\partial H}{\partial x_2} & \cdots & \frac{\partial H}{\partial x_m} \end{pmatrix} = \begin{pmatrix} \frac{\partial h}{\partial u_1} & \frac{\partial h}{\partial u_2} & \cdots & \frac{\partial h}{\partial u_n} \end{pmatrix} \begin{pmatrix} \frac{\partial u_1}{\partial x_1} & \frac{\partial u_1}{\partial x_2} & \cdots & \frac{\partial u_1}{\partial x_m} \\ \frac{\partial u_2}{\partial x_1} & \frac{\partial u_2}{\partial x_2} & \cdots & \frac{\partial u_2}{\partial x_m} \\ \vdots & \vdots & & \vdots \\ \frac{\partial u_n}{\partial x_1} & \frac{\partial u_n}{\partial x_2} & \cdots & \frac{\partial u_n}{\partial x_m} \end{pmatrix}.$$

Hence, for each i, with $1 \le i \le m$,

$$\frac{\partial H}{\partial x_i} = \frac{\partial h}{\partial u_1}\frac{\partial u_1}{\partial x_i} + \frac{\partial h}{\partial u_2}\frac{\partial u_2}{\partial x_i} + \cdots + \frac{\partial h}{\partial u_n}\frac{\partial u_n}{\partial x_i}.$$

Example 2.33. Let $h = uv+uw+vw$, where $u = y^2$, $v = x^2+2xy$ and $w = x^2-2xy$. Then, using the chain rule, we get

$$\begin{aligned} \frac{\partial h}{\partial x} &= \frac{\partial h}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial h}{\partial v}\frac{\partial v}{\partial x} + \frac{\partial h}{\partial w}\frac{\partial w}{\partial x} \\ &= (v+w)(0) + (u+w)(2x+2y) + (u+v)(2x-2y) \\ &= 2(x^2-2xy+y^2)(x+y) + 2(x^2+2xy+y^2)(x-y) \\ &= 2(x-y)^2(x+y) + 2(x+y)^2(x-y) \\ &= 2(x-y)(x+y)(x-y+x+y) \\ &= 4x(x-y)(x+y). \end{aligned}$$

Similarly,

$$\begin{aligned} \frac{\partial h}{\partial y} &= \frac{\partial h}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial h}{\partial v}\frac{\partial v}{\partial y} + \frac{\partial h}{\partial w}\frac{\partial w}{\partial y} \\ &= (v+w)(2y) + (u+w)(2x) + (u+v)(-2x) \\ &= (2x^2)(2y) + (x^2-2xy+y^2)(2x) + (x^2+2xy+y^2)(-2x) \\ &= 4x^2y + 2x(x-y)^2 - 2x(x+y)^2 \\ &= 4x^2y + 2x\left((x-y)^2 - (x+y)^2\right) \\ &= 4x^2y + 2x\left((2x)(-2y)\right) \qquad \text{(using } A^2-B^2 = (A+B)(A-B)\text{)} \\ &= 4x^2y - 8x^2y \\ &= -4x^2y. \end{aligned}$$

Euler's Theorem on homogeneous functions

As an interesting application of the chain rule we now prove Euler's Theorem on homogeneous functions. A function $f : \mathbb{R}^2 \to \mathbb{R}$ is **homogeneous** of **degree** n if, for all $x, y, \lambda \in \mathbb{R}$,

$$f(\lambda x, \lambda y) = \lambda^n f(x, y).$$

For example, $f(x,y) = xy$ is homogeneous of degree 2 since

$$f(\lambda x, \lambda y) = (\lambda x)(\lambda y) = \lambda^2 xy = \lambda^2 f(x,y).$$

Theorem 2.34 (Euler's Theorem on homogeneous functions)**.** *Let $g : \mathbb{R}^2 \to \mathbb{R}$ be a homogeneous function of degree n. Then*

$$ng(x,y) = x\frac{\partial g}{\partial x} + y\frac{\partial g}{\partial y}.$$

Proof. For $x, y \in \mathbb{R}$, let $G : \mathbb{R} \to \mathbb{R}$ where

$$G(\lambda) = g(\lambda x, \lambda y) = g(u,v),$$

$u = \lambda x$ and $v = \lambda y$. By the chain rule

$$\begin{aligned}\frac{\mathrm{d}G}{\mathrm{d}\lambda} &= \frac{\partial g}{\partial u}\frac{\mathrm{d}u}{\mathrm{d}\lambda} + \frac{\partial g}{\partial v}\frac{\mathrm{d}v}{\mathrm{d}\lambda} \\ &= x\frac{\partial g}{\partial u} + y\frac{\partial g}{\partial v}.\end{aligned}$$

However, since g is homogeneous of degree n, $G(\lambda) = \lambda^n g(x,y)$. Thus

$$\frac{\mathrm{d}G}{\mathrm{d}\lambda} = n\lambda^{n-1} g(x,y).$$

Comparing the two expressions for $\frac{\mathrm{d}G}{\mathrm{d}\lambda}$ gives

$$n\lambda^{n-1} g(x,y) = x\frac{\partial g}{\partial u} + y\frac{\partial g}{\partial v}. \tag{31}$$

Equation (31) is true for all λ. If we put $\lambda = 1$ then $u = x$, $v = y$, and equation (31) gives

$$ng(x,y) = x\frac{\partial g}{\partial x} + y\frac{\partial g}{\partial y},$$

as required. □

Exercises 2.6.

1. Let $g(x,y) = x^2 + 2xy - y^2$, $x = 3t - 1$ and $y = t^2$. Use the chain rule to find $\dfrac{\mathrm{d}g}{\mathrm{d}t}$.
2. Let $f : \mathbb{R} \to \mathbb{R}$, and let $z = f(xy)$.
   (a) Show that $x\dfrac{\partial z}{\partial x} - y\dfrac{\partial z}{\partial y} = 0$.
   (b) Verify the result in part (a) for the particular case when $f(t) = 2t^3 - t^2 + 3t$.

3. Let F be a function of u and v, where $u = x^2 + y^2$ and $v = xy$. Find $\dfrac{\partial F}{\partial u}$ and $\dfrac{\partial F}{\partial v}$ in terms of x, y, $\dfrac{\partial F}{\partial x}$ and $\dfrac{\partial F}{\partial y}$.

4. Let $g(x, y) = \dfrac{3xy^3 - 2x^2y^2}{4x + 5y}$.

   (a) Show that g is homogeneous of degree 3.

   (b) Verify Euler's Theorem on homogeneous functions by evaluating $x\frac{\partial g}{\partial x} + y\frac{\partial g}{\partial y}$ directly.

5. Suppose $u_1, u_2, \ldots, u_n : \mathbb{R} \to \mathbb{R}$ are functions of a single variable t, and f is a function of n variables. Let $F(t) = f\big(u_1(t), u_2(t), \ldots, u_n(t)\big)$. Use the general case of the chain rule (Theorem 2.32) to find an expression for $\frac{\mathrm{d}F}{\mathrm{d}t}$ in terms of the derivatives $\frac{\mathrm{d}u_1}{\mathrm{d}t}, \frac{\mathrm{d}u_2}{\mathrm{d}t}, \ldots, \frac{\mathrm{d}u_n}{\mathrm{d}t}$ and the partial derivatives $\frac{\partial f}{\partial u_1}, \frac{\partial f}{\partial u_2}, \ldots, \frac{\partial f}{\partial u_n}$.

2.7 Taylor series

Taylor series for functions of one variable

Let $U \subseteq \mathbb{R}$, $f : U \to \mathbb{R}$, and $a \in U$. If the derivative, and all the higher derivatives, of f exist at a, we can consider the series

$$f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2!}f''(a) + \cdots + \frac{(x-a)^n}{n!}f^{(n)}(a) + \cdots .$$

This is the **Taylor series** of f **centred** at a. For all "nice" functions f, this series will converge to $f(x)$, i.e. we have

$$f(x) = f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2!}f''(a) + \cdots + \frac{(x-a)^n}{n!}f^{(n)}(a) + \cdots$$

for all $x \in U$, or at least for all x sufficiently close to a. If the Taylor series is truncated after $n+1$ terms we get the **Taylor polynomial** or **Taylor approximation** of f, of degree n, centred at a.

If we replace x with $a + x$ throughout in the Taylor series of f, we get the following alternative way of expressing the Taylor series of f centred at a:

$$f(a+x) = f(a) + xf'(a) + \frac{x^2}{2!}f''(a) + \cdots + \frac{x^n}{n!}f^{(n)}(a) + \cdots .$$

The following are the Taylor series of some standard functions. The first three are centred at 0 and are valid for all $x \in \mathbb{R}$; the other two are centred at 1 and are valid for $x \in \mathbb{R}$ with $|x| < 1$.

$$
\begin{aligned}
e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + \ldots, \\
\sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + \ldots, \\
\cos x &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} \cdots + (-1)^n \frac{x^{2n}}{(2n)!} + \ldots, \\
\ln(1+x) &= x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots + (-1)^{n+1} \frac{x^n}{n} + \cdots, \\
(1+x)^\alpha &= 1 + \alpha x + \frac{\alpha(\alpha-1)x^2}{2!} + \frac{\alpha(\alpha-1)(\alpha-2)x^3}{3!} \\
&\quad + \cdots + \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)x^n}{n!} + \cdots.
\end{aligned}
$$

Taylor series for functions of two variables

Now let $U \subseteq \mathbb{R}^2$, $f : U \to \mathbb{R}$, and $(a, b) \in U$. If the partial derivatives, and all the higher partial derivatives, of f exist at (a, b), and if the function f is "nice", then

$$
\begin{aligned}
&f(a+h, b+k) \\
&\quad = f(a,b) + \big(hf_x(a,b) + kf_y(a,b)\big) \\
&\qquad + \frac{1}{2!}\big(h^2 f_{xx}(a,b) + 2hkf_{xy}(a,b) + k^2 f_{yy}(a,b)\big) \\
&\qquad + \frac{1}{3!}\big(h^3 f_{xxx}(a,b) + 3h^2k f_{xxy}(a,b) + 3hk^2 f_{xyy}(a,b) + k^3 f_{yyy}(a,b)\big) \\
&\qquad + \cdots + \frac{1}{n!}\sum_{i=0}^{n} \binom{n}{i} h^{n-i}k^i \frac{\partial^n f}{\partial^{n-i}x\partial^i y}(a,b) + \cdots,
\end{aligned}
$$

where $\displaystyle\binom{n}{i} = \frac{n!}{i!(n-i)!}$. This is the **Taylor series** of f **centred** at (a, b). If the series is truncated after $n + 1$ terms we get the **Taylor polynomial** or **Taylor approximation** of f, of degree n, centred at (a, b).

Example 2.35. Expand $x^2y + 3y - 2$ in powers of $x - 1$ and $y + 2$.

Let $f(x, y) = x^2y + 3y - 2$. Putting $x = 1 + h$ and $y = -2 + k$ we can obtain the required expansion by finding the Taylor series of f centred at $(1, -2)$. Now

$$
\begin{aligned}
f_x(x,y) &= 2xy, \\
f_y(x,y) &= x^2 + 3, \\
f_{xx}(x,y) &= 2y, \\
f_{xy}(x,y) &= 2x, \\
f_{yy}(x,y) &= 0,
\end{aligned}
$$

$$\begin{aligned}
f_{xxx}(x,y) &= 0,\\
f_{xxy}(x,y) &= 2,\\
f_{xyy}(x,y) &= 0,\\
f_{yyy}(x,y) &= 0,
\end{aligned}$$

and all higher derivatives are 0. Thus

$$\begin{aligned}
f(x,y) &= f(1+h,-2+k)\\
&= f(1,-2) + \big(hf_x(1,-2) + kf_y(1,-2)\big)\\
&\quad + \frac{1}{2!}\big(h^2 f_{xx}(1,-2) + 2hkf_{xy}(1,-2) + k^2 f_{yy}(1,-2)\big)\\
&\quad + \frac{1}{3!}\big(3h^2 k f_{xxy}(1,-2)\big)\\
&= -10 + (-4h+4k) + \frac{1}{2}\big(-4h^2 + 4hk + 0k^2\big) + \frac{1}{6}(6h^2k)\\
&= -10 - 4h + 4k - 2h^2 + 2hk + h^2k\\
&= -10 - 4(x-1) + 4(y+2) - 2(x-1)^2 + 2(x-1)(y+2) + (x-1)^2(y+2),
\end{aligned}$$

since $h = x-1$ and $k = y+2$.

Example 2.36. Find the Taylor approximation of degree 1 for

$$f(x,y) = \frac{12}{x^2+xy+y^2}$$

centred at $(2,2)$. Use your answer to estimate $f(2.1,1.8)$.

The partial derivatives of f are

$$\begin{aligned}
f_x(x,y) &= -\frac{12(2x+y)}{(x^2+xy+y^2)^2},\\
f_y(x,y) &= -\frac{12(x+2y)}{(x^2+xy+y^2)^2}.
\end{aligned}$$

Now, $f(2,2) = 1$, $f_x(2,2) = -\frac{1}{2}$ and $f_y(2,2) = -\frac{1}{2}$. Thus

$$\begin{aligned}
f(2+h,2+k) &\approx f(2,2) + hf_x(2,2) + kf_y(2,2)\\
&= 1 - \frac{1}{2}h - \frac{1}{2}k.
\end{aligned}$$

Using this Taylor approximation for f with $h = 0.1$ and $k = -0.2$ we get

$$\begin{aligned}
f(2.1,1.8) &= f(2+h,2+k)\\
&\approx 1 - \frac{1}{2}\times 0.1 - \frac{1}{2}\times(-0.2)\\
&= 1.05.
\end{aligned}$$

Note that $f(2.1,1.8) \approx 1.04987$, so the approximation given by the Taylor polynomial is quite good in this case.

Exercises 2.7.

1. Find the Taylor polynomial of degree 2 of $g(x, y) = \dfrac{xy + 1}{x + 2}$ centred at the point $(0, 0)$. Use your answer to estimate $g(0.1, 0.2)$.