eNote 21

Taylor's Limit Formula for Functions of Two Variables

In eNote 4 we studied Taylor's limit formula for one variable and found approximating polynomials of arbitrarily high degree for given functions. For functions of two variables a similar formula applies, but we will only develop this formula to approximations with quadratic polynomials. We can also use Taylor's limit formula to investigate the functional values at and in the neighbourhood of every point where the function is smooth. In particular we will first use the formula to find and investigate local and global maxima and minima of the functions.

Updated: 31.1.2023, shsp.

21.1 Derivatives of Higher Degree

In this eNote we will investigate the functions $f(x,y)$ of two variables in a set M that is a subset of the domain of $f(x,y)$ in the plane, $M \subset \mathcal{D}(f) \subset \mathbb{R}^2$.

We will assume that the function is partially differentiable to any order, that is, all derivatives exist for every (x,y) in the interior of M: $f'_x(x,y)$, $f'_y(x,y)$, $f''_{xx}(x,y)$, $f''_{xy}(x,y)$, $f''_{yy}(x,y)$, $f'''_{xyx}(x,y)$ etc. We will use these higher-order derivatives to construct (the coefficients in) the approximating polynomials of two variables exactly as we did for functions of one variable.

Definition 21.1

If a function $f(x,y)$ is partially differentiable to an arbitrary order at every point (x,y) in the interior $\mathring{M}$ of a given set M in the plane, then we call the function ***smooth*** in $\mathring{M}$.

Example 21.2 Higher-order Derivatives of Some Elementary Functions

Here are some first- and second-order derivatives of some well-known functions of two variables. Note that $f''_{yx}(x,y)$ cannot be seen explicitly in the table – this is because $f''_{yx}(x,y) = f''_{xy}(x,y)$ for all (x,y) when the functions are smooth, as we assume, see eNote 19. The function $\rho_{(0,0)}(x,y) = \sqrt{x^2+y^2}$ denotes the distance function from $(0,0)$ to (x,y); this is only smooth for $(x,y) \neq (0,0)$ as it appears for the shown first- and second-order partial derivatives of $\rho_{(0,0)}(x,y)$ in the table. The function $\rho^2_{(0,0)}(x,y) = x^2+y^2$, on the other hand, is (as appears from the table) a smooth function over its entire domain, i.e. for all $(x,y) \in \mathbb{R}^2$. Similarly this also applies for the square of the distance from an *arbitrary point* (x_0,y_0) i.e. the function:

$$\rho^2_{(x_0,y_0)}(x,y) = (x-x_0)^2 + (y-y_0)^2 \quad .$$

$f(x,y)$	$f'_x(x,y)$	$f'_y(x,y)$	$f''_{xx}(x,y)$	$f''_{xy}(x,y)$	$f''_{yy}(x,y)$
e^x	e^x	0	e^x	0	0
e^{x+y}	e^{x+y}	e^{x+y}	e^{x+y}	e^{x+y}	e^{x+y}
$e^{x^2+y^2}$	$2\cdot x\cdot e^{x^2+y^2}$	$2\cdot y\cdot e^{x^2+y^2}$	$(2+4\cdot x^2)\cdot e^{x^2+y^2}$	$4\cdot x\cdot y\cdot e^{x^2+y^2}$	$(2+4\cdot y^2)\cdot e^{x^2+y^2}$
$e^{x\cdot y}$	$y\cdot e^{x\cdot y}$	$x\cdot e^{x\cdot y}$	$y^2\cdot e^{x\cdot y}$	$e^{x\cdot y}+x\cdot y\cdot e^{x\cdot y}$	$x^2\cdot e^{x\cdot y}$
$\rho_{(0,0)}(x,y)$	$\frac{x}{\sqrt{x^2+y^2}}$	$\frac{y}{\sqrt{x^2+y^2}}$	$\frac{y^2}{(x^2+y^2)^{3/2}}$	$-\frac{x\cdot y}{(x^2+y^2)^{3/2}}$	$\frac{x^2}{(x^2+y^2)^{3/2}}$
$\rho^2_{(0,0)}(x,y)$	$2\cdot x$	$2\cdot y$	2	0	2
$\rho^2_{(x_0,y_0)}(x,y)$	$2\cdot(x-x_0)$	$2\cdot(y-y_0)$	2	0	2

In Example 21.2 the partial derivatives are shown in the same row in the table. In the following it proves to be smart to *gather* or *pack* the partial derivatives, in particular the two first-order derivatives to a vector, the gradient vector of $f(x,y)$ at (x_0,y_0), that are also introduced and applied in eNote 19:

$$\nabla f(x_0,y_0) = (f'_x(x_0,y_0), f'_y(x_0,y_0)) \quad , \tag{21-1}$$

but also the three second-order derivatives in a matrix $\mathbf{H}f(x_0,y_0)$ – the so-called Hessian matrix of the function $f(x,y)$ – at the point (x_0, y_0), determined by

$$\mathbf{H}f(x_0,y_0) = \begin{bmatrix} f''_{xx}(x_0,y_0) & f''_{xy}(x_0,y_0) \\ f''_{yx}(x_0,y_0) & f''_{yy}(x_0,y_0) \end{bmatrix} \quad . \tag{21-2}$$

Note in particular that this matrix is symmetric because $f''_{xy}(x_0,y_0) = f''_{yx}(x_0,y_0)$. Therefore it has some special properties that are decisive in the geometric analysis of functions to the second order.

|||| Example 21.3 Gradients and Hessians

From the tabel we read the following gradients and Hessian matrices of the given elementary smooth functions at the point (x_0, y_0):

$f(x,y)$	$\nabla f(x,y)$	$\mathbf{H}f(x,y)$
e^x	$(e^x, 0)$	$\begin{bmatrix} e^x & 0 \\ 0 & 0 \end{bmatrix}$
e^{x+y}	(e^{x+y}, e^{x+y})	$\begin{bmatrix} e^{x+y} & e^{x+y} \\ e^{x+y} & e^{x+y} \end{bmatrix}$
$\rho^2_{(0,0)}(x,y) = x^2 + y^2$	$(2 \cdot x, 2 \cdot y)$	$\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$
$(x - x_0) \cdot \rho^2_{(x_0,y_0)}(x,y)$	$(3 \cdot (x-x_0)^2 + (y-y_0)^2,$ $2 \cdot (x-x_0) \cdot (y-y_0))$	$\begin{bmatrix} 6 \cdot (x-x_0) & 2 \cdot (y-y_0) \\ 2 \cdot (y-y_0) & 2 \cdot (x-x_0) \end{bmatrix}$

Note that the function $f(x,y) = (x - x_0) \cdot \rho^2_{(x_0,y_0)}(x,y)$ has the gradient vector $\mathbf{0}$ and the Hessian matrix $\mathbf{0}$ when they are computed at the basis point (x_0, y_0) of the distance function $\rho_{(x_0,y_0)}(x,y)$.

|||| Exercise 21.4

Determine the gradient $\nabla f(x,y)$ and the Hessian matrix $\mathbf{H}f(x,y)$ at every point $(x,y) \in \mathbb{R}^2$ of each of the following functions:

$$f(x,y) = \cos(x+y) \quad , \quad f(x,y) = \cos(x \cdot y) \quad , \quad f(x,y) = x^3 + y^2 + x \cdot y \quad . \tag{21-3}$$

21.2 Approximation of Functions of Two Variables

Now we will compare the function $f(x,y)$ with an arbitrary quadratic polynomial with the two variables x and y and determine the coefficients in the polynomial by demanding that the difference between the two shall be small close to the point x_0.

The polynomial that we will test and use for the comparison is the following:

$$\begin{aligned} & a + b_1 \cdot (x - x_0) + b_2 \cdot (y - y_0) \\ & + \left(\frac{h_{11}}{2}\right) \cdot (x - x_0)^2 + h_{12} \cdot (x - x_0) \cdot (y - y_0) + \left(\frac{h_{22}}{2}\right) \cdot (y - y_0)^2 \quad . \end{aligned} \tag{21-4}$$

We will choose the coefficients a, b_1, b_2, h_{11}, h_{12}, and h_{22}, so that the residual $R(x,y) = R_{2,(x_0,y_0)}(x,y)$ in relation to $f(x,y)$ becomes (numerically) as small as possible close to the point (x_0, y_0) – just as we did when we motivated Taylor's formulae for functions of one variable in eNote 4:

$$\begin{aligned} f(x,y) = {} & a + b_1 \cdot (x - x_0) + b_2 \cdot (y - y_0) \\ & + \left(\frac{h_{11}}{2}\right) \cdot (x - x_0)^2 + h_{12} \cdot (x - x_0) \cdot (y - y_0) + \left(\frac{h_{22}}{2}\right) \cdot (y - y_0)^2 \\ & + R(x,y) \quad . \end{aligned} \tag{21-5}$$

The quadratic polynomial that we use for the approximation in (21-5) is a totally general quadratic polynomial; *every* quadratic polynomial can be written in this form even if the values of x_0 and y_0 are given beforehand. Note that there are 6 coefficients that are not yet determined: a, b_1, b_2, h_{11}, h_{12}, og h_{22}. And the set of quadratic polynomials is a vector space of exactly 6 dimensions!

The trick is to find the best values of these 6 coefficients, so that $R(x,y)$ becomes numerically small at and close to the point (x_0, y_0).

The residual function $R(x,y)$ can be written as the difference between $f(x,y)$ and the polynomial:

$$\begin{aligned} R(x,y) = {} & f(x,y) - a - b_1 \cdot (x - x_0) - b_2 \cdot (y - y_0) \\ & - \left(\frac{h_{11}}{2}\right) \cdot (x - x_0)^2 - h_{12} \cdot (x - x_0) \cdot (y - y_0) - \left(\frac{h_{22}}{2}\right) \cdot (y - y_0)^2 \quad . \end{aligned} \tag{21-6}$$

We will now state the natural requirements to the residual function – requirements that secure that the residual function is very small at and in the neighborhood of (x_0, y_0) and thus that the *polynomial* we find in this way is close to the function $f(x, y)$.

Naturally the first requirement is that the residual function is 0 at (x_0, y_0). We substitute this requirement into (21-6) and get:

$$R(x_0, y_0) = 0 \quad \text{yielding} \quad 0 = f(x_0, y_0) - a \quad , \tag{21-7}$$

whereby the constant term in the approximating polynomial is already determined:

$$a = f(x_0, y_0) \quad . \tag{21-8}$$

Of course the next requirement is $R'_x(x_0, y_0) = R'_y(x_0, y_0) = 0$ such that $\nabla R(x_0, y_0) = (0, 0)$, which geometrically means that the graph of the residual function has a horisontal tangent plane – identical to the (x, y) plane – at (x_0, y_0). If we substitute these requirements into $R(x, y)$ in Equation (21-6) we get

$$\begin{aligned} R'_x(x_0, y_0) = 0 \quad &\text{which yields} \quad 0 = f'_x(x_0, y_0) - b_1 \quad , \\ R'_y(x_0, y_0) = 0 \quad &\text{which yields} \quad 0 = f'_y(x_0, y_0) - b_2 \quad , \end{aligned} \tag{21-9}$$

whereby the coefficients $b_1 = f'_x(x_0, y_0)$ and $b_2 = f'_x(x_0, y_0)$ are also determined. We have:

$$(b_1, b_2) = \nabla f(x_0, y_0) \quad . \tag{21-10}$$

The last requirement on $R(x, y)$ we use is that all three second-order partial derivatives of $R(x, y)$ are also 0 at (x_0, y_0):

Substituting these 3 requirements into $R(x, y)$ in Equation (21-6) we get

$$\begin{aligned} R''_{xx}(x_0, y_0) = 0 \quad &\text{which yields} \quad 0 = f''_{xx}(x_0, y_0) - h_{11} \quad , \\ R''_{xy}(x_0, y_0) = 0 \quad &\text{which yields} \quad 0 = f''_{xy}(x_0, y_0) - h_{12} \quad , \\ R''_{yy}(x_0, y_0) = 0 \quad &\text{which yields} \quad 0 = f''_{yy}(x_0, y_0) - h_{22} \quad . \end{aligned} \tag{21-11}$$

By this the last 3 coefficients in the approximating polynomial are found.

We can write the result like this – while keeping in mind that $f''_{xy}(x_0, y_0) = f''_{xy}(x_0, y_0) = h_{12}$:

$$\begin{bmatrix} h_{11} & h_{12} \\ h_{12} & h_{22} \end{bmatrix} = \begin{bmatrix} f''_{xx}(x_0, y_0) & f''_{xy}(x_0, y_0) \\ f''_{yx}(x_0, y_0) & f''_{yy}(x_0, y_0) \end{bmatrix} = \mathbf{H}f(x_0, y_0) \quad . \tag{21-12}$$

Thus we have seen that the requirements posed to the residual function *can be fulfilled*, and that the resulting quadratic polynomial that fulfills the requirements, precisely therefore, deserves to be called *the approximating polynomial of the second degree* of $f(x, y)$ with the point of development (x_0, y_0):

|||| Definition 21.5 Approximating Quadratic Polynomial

Let $f(x, y)$ denote a smooth function in a set M in $\mathbb{R}^2$ that contains the point $(x_0, y_0) \in \mathring{M}$. The polynomial

$$\begin{aligned} P_{2,(x_0,y_0)}(x, y) = f(x_0, y_0) &+ f'_x(x_0, y_0) \cdot (x - x_0) + f'_y(x_0, y_0) \cdot (y - y_0) \\ &+ \frac{1}{2} \cdot f''_{xx}(x_0, y_0) \cdot (x - x_0)^2 \\ &+ f''_{xy}(x_0, y_0) \cdot (x - x_0) \cdot (y - y_0) \\ &+ \frac{1}{2} \cdot f''_{yy}(x_0, y_0) \cdot (y - y_0)^2 \end{aligned} \tag{21-13}$$

is called the ***approximating polynomial*** of the second degree of the function $f(x, y)$ with the point of development (x_0, y_0).

Now – just like for functions of one variable – it is reasonable to expect that with the requirements used on the residual function $R(x, y) = R_{2,(x_0,y_0)}(x, y)$ it must be numerically comparable to a *power of the distance function* $\rho_{(x_0,y_0)}(x, y)$ to the point of development (x_0, y_0). E.g. from the last example in 21.3 we can see that all requirements that we have posed on $R(x, y)$ are fulfilled for the function $(x - x_0) \cdot \rho^2_{(x_0,y_0)}(x, y)$.

This is precisely the content of the following theorem that expresses this by the use of an epsilon function, this of course not normally being as simple as $(x - x_0)$:

|||| Lemma 21.6 The Residual Function with an Epsilon Term

Given a smooth function $f(x,y)$. Then the residual function $R_{2,(x_0,y_0)}(x,y)$ can be written in the following form:

$$R_{2,(x_0,y_0)}(x,y) = \rho^2_{(x_0,y_0)}(x,y) \cdot \varepsilon_f(x-x_0, y-y_0) \quad , \tag{21-14}$$

where $\varepsilon_f(x-x_0, y-y_0)$ is an epsilon function of $(x-x_0, y-y_0)$, i.e.

$$\varepsilon_f(x-x_0, y-y_0) \to 0 \quad \text{for} \quad (x,y) \to (x_0,y_0) \quad . \tag{21-15}$$

We now have all the ingredients to the main theorem of this eNote:

|||| Theorem 21.7 Taylor's Limit Formula for Functions of Two Variables

Every smooth function $f(x,y)$ of two variables can be split into an approximating quadratic polynomial at a given point of development (x_0,y_0) and a residual function with an epsilon term in this way:

$$\begin{aligned} f(x,y) &= P_{2,(x_0,y_0)}(x,y) + R_{2,(x_0,y_0)}(x,y) \\ &= f(x_0,y_0) + f'_x(x_0,y_0)\cdot(x-x_0) + f'_y(x_0,y_0)\cdot(y-y_0) \\ &\qquad + \frac{1}{2}\cdot f''_{xx}(x_0,y_0)\cdot(x-x_0)^2 \\ &\qquad + f''_{xy}(x_0,y_0)\cdot(x-x_0)\cdot(y-y_0) \\ &\qquad + \frac{1}{2}\cdot f''_{yy}(x_0,y_0)\cdot(y-y_0)^2 \\ &\quad + \rho^2_{(x_0,y_0)}(x,y)\cdot\varepsilon_f(x-x_0,y-y_0) \quad . \end{aligned} \tag{21-16}$$

By using the gradient of $f(x,y)$ and the Hessian matrix of $f(x,y)$ we can write Taylor's limit formula (21-16) in a more compact form that we will use in what follows with respect to a geometric analysis of $f(x,y)$ at and close to the point of development (x_0,y_0):

|||| Theorem 21.8 Taylor's Limit Formula, Compact, Geometric

Every smooth function $f(x,y)$ of two variables is the sum of an approximating quadratic polynomial with a given point of development (x_0, y_0) and a residual function like this:

$$\begin{aligned} f(x,y) = f(x_0,y_0) + \nabla f(x_0,y_0) \cdot (x - x_0, y - y_0) \\ + \frac{1}{2} \cdot \begin{bmatrix} x - x_0 & y - y_0 \end{bmatrix} \cdot \boldsymbol{H} f(x_0,y_0) \cdot \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} \\ + \rho^2_{(x_0,y_0)}(x,y) \cdot \varepsilon_f(x - x_0, y - y_0) \quad , \end{aligned} \tag{21-17}$$

where

$$\nabla f(x_0,y_0) = (f'_x(x_0,y_0), f'_y(x_0,y_0)) \quad \text{and}$$

$$\mathbf{H} f(x_0,y_0) = \begin{bmatrix} f''_{xx}(x_0,y_0) & f''_{xy}(x_0,y_0) \\ f''_{yx}(x_0,y_0) & f''_{yy}(x_0,y_0) \end{bmatrix} \quad . \tag{21-18}$$

|||| Proof

A direct computation of the following matrix product shows that (21-17) is equivalent to (21-16), and this is the only new thing in the theorem:

$$\begin{aligned} &\begin{bmatrix} x - x_0 & y - y_0 \end{bmatrix} \cdot \mathbf{H} f(x_0,y_0) \cdot \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} \\ &= \begin{bmatrix} x - x_0 & y - y_0 \end{bmatrix} \cdot \begin{bmatrix} f''_{xx}(x_0,y_0) & f''_{xy}(x_0,y_0) \\ f''_{yx}(x_0,y_0) & f''_{yy}(x_0,y_0) \end{bmatrix} \cdot \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} \\ &= \quad f''_{xx}(x_0,y_0) \cdot (x - x_0)^2 \\ &\quad + 2 \cdot f''_{xy}(x_0,y_0) \cdot (x - x_0) \cdot (y - y_0) \\ &\quad + f''_{yy}(x_0,y_0) \cdot (y - y_0)^2 \quad . \end{aligned} \tag{21-19}$$

■

The approximating quadratic polynomial of $f(x,y)$ with the point of development (x_0, y_0)

is therefore in compact form :

$$P_{2,(x_0,y_0)}(x,y) = f(x_0,y_0) + \nabla f(x_0,y_0)\cdot(x-x_0, y-y_0) + \frac{1}{2}\cdot\begin{bmatrix} x-x_0 & y-y_0 \end{bmatrix}\cdot \mathbf{H}f(x_0,y_0)\cdot\begin{bmatrix} x-x_0 \\ y-y_0 \end{bmatrix} \quad . \tag{21-20}$$

Note that the approximating linear polynomial of $f(x,y)$ is exactly the *linear part* (linearization) of the approximating quadratic polynomial:

$$P_{1,(x_0,y_0)}(x,y) = f(x_0,y_0) + \nabla f(x_0,y_0)\cdot(x-x_0, y-y_0). \tag{21-21}$$

In eNote 20 we saw that the geometric interpretation of $P_{1,(x_0,y_0)}(x,y)$ is that the graph of the approximating linear polynomial $P_{1,(x_0,y_0)}(x,y)$ (i.e. the plane $z = P_{1,(x_0,y_0)}(x,y)$ in (x,y,z) space) exactly is the *tangent plane to the graph surface* of $f(x,y)$ at the point $(x_0, y_0, f(x_0,y_0))$.

|||| Definition 21.9 Stationary Points of Functions of Two Variables

Let $f(x,y)$ denote a smooth function in a set M in $\mathcal{D}(f) \subset \mathbb{R}^2$. By a ***stationary point*** of $f(x,y)$ we will understand any point (x_0,y_0) in $\mathring{M}$ where the gradient of $f(x,y)$ is $(0,0)$.

Thus at a stationary point of $f(x,y)$ we have:

$$\nabla f(x_0,y_0) = \Big(f'_x(x_0,y_0), f'_y(x_0,y_0)\Big) = (0,0) \quad . \tag{21-22}$$

Geometrically the definition 21.9 means that (x_0,y_0) is a stationary point in $\mathring{M}$ of $f(x,y)$ if the tangent plane for the graph of $f(x,y)$ through the point $(x_0, y_0, f(x_0,y_0))$ is horisontal, i.e. parallel to the (x,y) plane, such that the unit normal vector to the graph surface at the point $(x_0, y_0, f(x_0,y_0))$ is $\mathbf{N}_{(x_0,y_0)}(f) = (0,0,1)$.

21.3 Functional Investigations

Just like for continuous functions of one variable the functional values $f(x,y)$ are bounded if the region of investigation is sufficiently nice:

|||| Theorem 21.10 Main Theorem for Continuous Functions of Two Variables

Let $f(x,y)$ denote a function that is continuous in all of its domain $\mathcal{D}(f) \subset \mathbb{R}^2$. Let M be a bounded, closed, and connected set in $\mathcal{D}(f)$.

Then the range of the function $f(x,y)$ in the set M is a bounded, closed, and connected interval $[\,A\,,\,B\,] \subset \mathbb{R}$, in short:

$$\mathcal{R}(f_{|_M}) = f(M) = \{f(x)\,|\,x \in M\} = [\,A\,,\,B\,] \quad , \tag{21-23}$$

where we still allow for the possibility of $A = B$ and this happens naturally exactly when $f(x,y)$ is constant in all of M.

Exactly as for functions of one variable we now define:

|||| Definition 21.11 Global Minimum and Global Maximum

When a function $f(x,y)$ has the range $\mathcal{R}(f_{|_M}) = f(M) = [\,A\,,\,B\,]$ in a set M we say that

1. A is the ***global minimum*** of $f(x,y)$ in M, and if $f(x_0,y_0) = A$ for $(x_0,y_0) \in M$ then (x_0,y_0) is a ***global minimum point*** of $f(x,y)$ in M.
2. B is the ***global maksimum*** of $f(x,y)$ in M, and if $f(x_0,y_0) = B$ for $(x_0,y_0) \in M$ then (x_0,y_0) is a ***global maximum point*** of $f(x,y)$ in M.

An entirely natural and very frequent problem is to find the global maxima and minima of given functions $f(x,y)$ in given sets M in $\mathbb{R}^2$ and at the same time determine the (x,y) points in M for which these maxima and minima are *found* – that is, to determine the minimum and maximum points.

To solve this problem the following is a great help (just as in the investigation of functions of one variable):

|||| Lemma 21.12 Maxima and Minima at Stationary Points

Let (x_0, y_0) be a global maximum or minimum point of $f(x, y)$ in M.

Assume that (x_0, y_0) is not a point on the boundary of the set M, i.e. we assume that (x_0, y_0) is an interior point in M. Furthermore assume that $f(x, y)$ is differentiable at the point (x_0, y_0).

Then (x_0, y_0) is a ***stationary point*** of $f(x, y)$ at $(x_0, y_0) \in \mathring{M}$:

$$(f'_x(x_0, y_0), f'_y(x_0, y_0)) = \boldsymbol{\nabla} f(x_0, y_0) = \mathbf{0} = (0, 0) \quad . \tag{21-24}$$

|||| Proof

The proof follows a line of argument similar to that for functions of one variable. In short: If $\boldsymbol{\nabla} f(x_0, y_0) \neq (0, 0)$ then the functional values *increase* effectively in the direction $\boldsymbol{\nabla} f(x_0, y_0)$ from (x_0, y_0) and the functional values *decrease* effectively in the opposite direction $-\boldsymbol{\nabla} f(x_0, y_0)$ from (x_0, y_0) and thus (x_0, y_0) can neither be a minimum point nor a maximum point.

■

Thus we have the method of investigation for functions of two variables:

|||| Method 21.13 Method of Investigation

Let $f(x,y)$ be a continuous function and M a closed, bounded, and connected region in the domain $\mathcal{D}(f) \subset \mathbb{R}^2$.

Maximum and minimum values of the function $f(x,y)$, i.e. A and B of the range $[A,B]$ of $f(x,y)$ in M, are found by computing and comparing the functional values at the following points:

1. The boundary points of the set M.
2. The exceptional points, i.e. the points in the interior of M where the function $f(x,y)$ is *not* differentiable.
3. The stationary points, i.e. all the points (x_0, y_0) in the interior of M where the function $f(x,y)$ *is* differentiable and $(f'_x(x_0,y_0), f'_y(x_0,y_0)) = \nabla f(x_0,y_0) = \mathbf{0} = (0,0)$.

By this method of investigation we find not only the global maximum and minimum values of A and B but also the (x,y) points in $M \subset \mathbb{R}^2$ for which the function is equal to the the global maximum B and the global minimum A, respectively, i.e the maximum and the minimum points in the actual region.

|||| Example 21.14 Functional Investigation, Range

Let $f(x,y) = 1 + x^2 - x \cdot y^2$. We will investigate the function in the closed and bounded circular region M in the plane that is bounded by the circle with radius 1 and centre at $(0,0)$, see Figure 21.1.

$$\nabla f(x,y) = (2 \cdot x + y^2, 2 \cdot y \cdot x) \quad . \tag{21-25}$$

The function is differentiable in all of its domain $\mathcal{D}(f) = \mathbb{R}^2$. In the interior of M, i.e. the open circular disc with radius 1 and centre at $(0,0)$ the function has exactly one stationary point $(x_0, y_0) = (0,0)$ and the value of $f(x_0,y_0)$ is $f(0,0) = 1$. Since there are no exceptional points of $f(x,y)$ in the open circular disc it only remains for us to investigate the function on the boundary of the region, i.e. the functional values on the circle with radius 1 and centre at $(0,0)$. We parametrize the actual circle in the plane in the ordinary way:

$$\mathbf{r}(u) = (\cos(u), \sin(u)) \quad , \quad u \in [-\pi, \pi] \quad , \tag{21-26}$$

and substitute the points on the circle into the functional expression and thus obtain the following composite function that expresses the functional values on the circle when u traverses the interval $[-\pi, \pi]$:

$$h(u) = f(\mathbf{r}(u)) = 1 + \cos^2(u) - \cos(u) \cdot \sin^2(u) \quad . \tag{21-27}$$

We differentiate $h(u)$ and find:

$$h'(u) = -\sin(u) \cdot (2 \cdot \cos(u) + 3 \cdot \cos^2(u) - 1) \tag{21-28}$$

that gives $h'(u) = 0$ for $u = 0$, $u = \pm\pi$, and $u = \pm\arccos(1/3)$. At these stationary points of the function $h(u)$, $h(u)$ assumes the values : $h(0) = 2$, $h(\pm\pi) = 2$, and $h(\pm\arccos(1/3)) = 22/27$, respectively.

By comparing all the possible values for maximum and minimum values of $f(x, y)$ given by the method 21.13 we conclude that the range of $f(x, y)$ on the given circular disc is:

$$f(M) = \left[\frac{22}{27}, 2\right] \quad , \tag{21-29}$$

and the maximum points and minimum points all exist on boundary of the circular region and are respectively: maximum points $(\cos(0), \sin(0)) = (1, 0)$ and $(\cos(\pm\pi), \sin(\pm\pi)) = (-1, 0)$; minimum points, $(\cos(\pm\arccos(1/3)), \sin(\pm\arccos(1/3))) = (1/3, \pm 2\sqrt{2}/3)$.

|||| Definition 21.15 Local Minima and Local Maxima

Let $f(x, y)$ denote a function in a region M that contains a given interior point $(x_0, y_0) \in \mathring{M}$.

1. If $f(x, y) \geq f(x_0, y_0)$ for all (x, y) in a (generally small) neighborhood around (x_0, y_0) i.e. for all (x, y) sufficiently close to (x_0, y_0), then $f(x_0, y_0)$ is called a ***local minimum value*** of $f(x, y)$ in M, and (x_0, y_0) is a ***local minimum point*** of $f(x, y)$ in M. If actually $f(x, y) > f(x_0, y_0)$ for all (x, y) in the neighborhood except the point (x_0, y_0) itself, then $f(x_0, y_0)$ is a ***proper local minimum value***.

2. If $f(x, y) \leq f(x_0, y_0)$ for all (x, y) in a (generally small) neighborhood around (x_0, y_0), then $f(x_0, y_0)$ is a ***local maximum value*** of $f(x, y)$ in M, and (x_0, y_0) is a ***local maximum point*** of $f(x, y)$ in M. If actually $f(x, y) < f(x_0, y_0)$ for all (x, y) in the neighborhood (except the point (x_0, y_0) itself), then $f(x_0, y_0)$ is a ***proper local maximum value***.

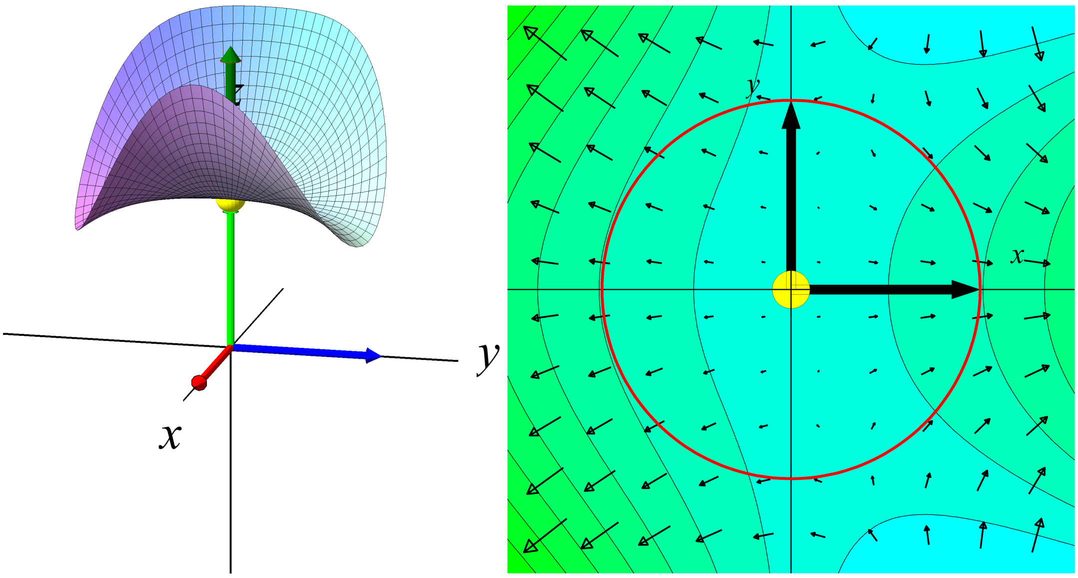

Figure 21.1: The graph of the function $f(x,y) = 1 + x^2 - x \cdot y^2$. In the right part of the figure the level curves and the gradient vector field of the function can be inspected. The red circle is the boundary of the circular region that is considered in Example 21.14. It is seen that the gradient vector field of $f(x,y)$ is *orthogonal to the circle* precisely at the 4 points where the 'height'-functon $h(u)$ has $h'(u) = 0$ viz. $(\pm 1, 0)$ and $(1/3, \pm 2\sqrt{2}/3)$.

21.4 Special Investigation at Stationary Points

If the function we wish to investigate is smooth at its stationary points, then we can qualify the method 21.13 even better, because the approximating quadratic polynomial with point of development at a stationary point (x_0, y_0) can be used to decide whether the value of $f(x,y)$ at the stationary point is a candidate to be a global maximum value or a global minimum value.

A proper local minimum point in the interior of M cannot of course be a global maximum point of $f(x,y)$ in M and a proper maximum point in the interior of M cannot be a global minimum point of $f(x,y)$ in M.

For the local analysis of $f(x,y)$ around a stationary point we will use the *eigenvalues* of the Hessian matrix $\mathbf{H}f(x_0, y_0)$ of $f(x,y)$ at the stationary point. As mentioned previously, the symmetry of $\mathbf{H}f(x_0, y_0)$ is an important property, for as we shall now see it means that $\mathbf{H}f(x_0, y_0)$ can *be diagonalized by a similarity transformation*, which we have introduced and treated in eNote 10.

The following theorem about symmetric matrices is also important for many other applications. As is shown in eNote 19 the result can be generalized to arbitrary symmetric matrices. For 2×2-matrices the proof is relatively simple, so we will go through the reasoning below – as an appetizer for the general result.

|||| Theorem 21.16 Diagonalization of Symmetric 2×2-Matrices

Every symmetric 2×2 matrix $\mathbf{A}$ has real eigenvalues λ_1 and λ_2 and $\mathbf{A}$ can be diagonalized by a 2×2-similarity transformation $\mathbf{D}$ (change-of-base matrix) in which the two column vectors are orthogonal unit eigenvectors of $\mathbf{A}$.

This means that $\mathbf{D}^{-1} = \mathbf{D}^\top$ and

$$\begin{aligned} \mathbf{\Lambda} &= \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \quad , \\ \mathbf{A} &= \mathbf{D} \cdot \mathbf{\Lambda} \cdot \mathbf{D}^{-1} = \mathbf{D} \cdot \mathbf{\Lambda} \cdot \mathbf{D}^\top \quad , \text{and} \\ \mathbf{\Lambda} &= \mathbf{D}^{-1} \cdot \mathbf{A} \cdot \mathbf{D} = \mathbf{D}^\top \cdot \mathbf{A} \cdot \mathbf{D} \quad . \end{aligned} \tag{21-30}$$

|||| Proof

First we can directly see that every symmetric 2×2 matrix has *real eigenvalues*: The characteristic polynomial of the symmetric matrix

$$\mathbf{A} = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \tag{21-31}$$

is

$$\mathbf{K}_\mathbf{A}(\lambda) = (a - \lambda) \cdot (c - \lambda) - b^2 = \lambda^2 - \lambda \cdot (a + c) + (a \cdot c - b^2) \quad , \tag{21-32}$$

that has the roots (the eigenvalues)

$$\begin{aligned} \lambda_1 &= \frac{1}{2}\left(a + c + \sqrt{(a-c)^2 + 4 \cdot b^2}\right) \quad \text{and} \\ \lambda_2 &= \frac{1}{2}\left(a + c - \sqrt{(a-c)^2 + 4 \cdot b^2}\right) \quad , \end{aligned} \tag{21-33}$$

that both are real numbers.

The two eigenvalues are equal exactly when $a = c$ and $b = 0$, that is when $\mathbf{A}$ *is* already in diagonal form and is a product of the unit matrix – possibly the $\mathbf{0}$-matrix. In these cases we can choose $\mathbf{D} = \mathbf{E}$.

If $\lambda_1 \neq \lambda_2$ we also get directly from eNote 10,Theorem 10.4, that there *exists* a similarity transformation $\mathbf{D}$ that diagonalizes $\mathbf{A}$:

$$\mathbf{\Lambda} = \mathbf{D}^{-1} \cdot \mathbf{A} \cdot \mathbf{D} \quad . \tag{21-34}$$

The column vectors $\mathbf{D}$ are eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$ corresponding to λ_1 and λ_2, respectively.

We will show that $\mathbf{v}_1$ and $\mathbf{v}_2$ are orthogonal vectors. Since $\mathbf{A}$ is symmetric we have

$$\mathbf{v}_1 \cdot (\mathbf{A}\mathbf{v}_2) = \mathbf{v}_2 \cdot (\mathbf{A}\mathbf{v}_1) \quad ; \tag{21-35}$$

this can be seen by a direct calculation using arbitrary vectors $\mathbf{v}_1$ and $\mathbf{v}_2$.

Since $\mathbf{v}_1$ is a proper eigenvector of $\mathbf{A}$ corresponding to λ_1 and similarly $\mathbf{v}_2$ is a proper eigenvector of $\mathbf{A}$ corresponding to λ_2 we get:

$$\begin{aligned} \mathbf{A}\mathbf{v}_2 &= \lambda_2 \cdot \mathbf{v}_2 \\ \mathbf{A}\mathbf{v}_1 &= \lambda_1 \cdot \mathbf{v}_1 \quad , \end{aligned} \tag{21-36}$$

which substituted into (21-35) yields:

$$\lambda_2 \cdot \mathbf{v}_1 \cdot \mathbf{v}_2 = \lambda_1 \cdot \mathbf{v}_2 \cdot \mathbf{v}_1 \quad , \tag{21-37}$$

so that

$$(\mathbf{v}_1 \cdot \mathbf{v}_2) \cdot (\lambda_1 - \lambda_2) = 0 \quad . \tag{21-38}$$

Since $\lambda_1 \neq \lambda_2$ this equation can only be fulfilled for $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$. The two eigenvectors are therefore orthogonal as postulated in the theorem.

Next we can assume that the two eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are unit vectors – if not we just divide by their respective lengths and this will not change their "eigenvector properties".

The two unit eigenvectors are used as column vectors in the similarity transformation matrix D, which thereby fulfills the following equation :

$$\mathbf{D}^{\top} \cdot \mathbf{D} = \mathbf{E} \quad , \tag{21-39}$$

and thus

$$\mathbf{D}^{\top} = \mathbf{D}^{-1} \tag{21-40}$$

as wished for and postulated in the theorem .

■

In particular the eigenvalues of the Hessian matrix are always two real numbers and it is these eigenvalues we now will use in order to decide whether a given stationary point is a proper maximum or minimum point.

|||| Lemma 21.17 Local Analysis About a Stationary Point

Let $f(x,y)$ be a smooth function and assume that (x_0, y_0) is a stationary point of $f(x,y)$ in an open region $\mathring{M} \subset \mathcal{D}(f) \subset \mathbb{R}^2$. Let $\lambda_1(x_0,y_0)$ and $\lambda_2(x_0,y_0)$ denote the eigenvalues of the Hessian matrix of $f(x,y)$ at (x_0,y_0). The following applies:

1. If $\lambda_1(x_0,y_0) > 0$ and $\lambda_2(x_0,y_0) > 0$ then $f(x_0,y_0)$ is a proper local minimum value of $f(x,y)$.
2. If $\lambda_1(x_0,y_0) < 0$ and $\lambda_2(x_0,y_0) < 0$ then $f(x_0,y_0)$ is a proper local maximum value of $f(x,y)$.
3. If $\lambda_1(x_0,y_0) \cdot \lambda_2(x_0,y_0) < 0$ then $f(x_0,y_0)$ is neither a local minimum value nor a local maximum value of $f(x,y)$.
4. If $\lambda_1(x_0,y_0) \cdot \lambda_2(x_0,y_0) = 0$ then there is not enough information to decide whether $f(x_0,y_0)$ is either a local minimum value or a local maximum value of $f(x,y)$.

|||| Proof

From Taylor's limit formula we have at the stationary point the following representation of $f(x,y)$:

$$f(x,y) = f(x_0,y_0) + \frac{1}{2} \cdot \begin{bmatrix} x-x_0 & y-y_0 \end{bmatrix} \cdot \mathbf{H}f(x_0,y_0) \cdot \begin{bmatrix} x-x_0 \\ y-y_0 \end{bmatrix} + R_2(x,y) \,. \qquad (21\text{-}41)$$

We diagonalize $\mathbf{H}f(x_0,y_0)$ with the similarity transformation $\mathbf{D}$, cf. Theorem 21.16, so that

$$\begin{aligned} f(x,y) &= f(x_0,y_0) + \frac{1}{2} \cdot \begin{bmatrix} x-x_0 & y-y_0 \end{bmatrix} \cdot \mathbf{D} \cdot \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \cdot \mathbf{D}^{-1} \cdot \begin{bmatrix} x-x_0 \\ y-y_0 \end{bmatrix} + R_2(x,y) \\ &= f(x_0,y_0) + \frac{1}{2} \cdot \left(\begin{bmatrix} x-x_0 & y-y_0 \end{bmatrix} \cdot \mathbf{D} \right) \cdot \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \cdot \left(\mathbf{D}^\top \cdot \begin{bmatrix} x-x_0 \\ y-y_0 \end{bmatrix} \right) + R_2(x,y) \quad , \end{aligned} \qquad (21\text{-}42)$$

where we note that the following two vectors are equal (apart from transposition)

$$\left(\left[\begin{array}{cc} x-x_0 & y-y_0 \end{array}\right]\cdot \mathbf{D}\right)^\top = \left(\mathbf{D}^\top \cdot \left[\begin{array}{c} x-x_0 \\ y-y_0 \end{array}\right]\right) = \left[\begin{array}{c} \alpha \\ \beta \end{array}\right] \quad . \tag{21-43}$$

I.e. we can write in short form:

$$\begin{aligned} f(x,y) &= f(x_0,y_0) + \frac{1}{2}\cdot\left[\begin{array}{cc} \alpha & \beta \end{array}\right]\cdot\left[\begin{array}{cc} \lambda_1 & 0 \\ 0 & \lambda_2 \end{array}\right]\cdot\left[\begin{array}{c} \alpha \\ \beta \end{array}\right] + R_2(x,y) \\ &= f(x_0,y_0) + \frac{1}{2}\cdot\left(\lambda_1\cdot\alpha^2 + \lambda_2\cdot\beta^2\right) + R_2(x,y) \end{aligned} \tag{21-44}$$

If both λ_1 and λ_2 are positive and $\lambda_1 \geq \lambda_2 > 0$ we get directly from this:

$$f(x,y) \geq f(x_0,y_0) + \lambda_2\cdot\left(\alpha^2+\beta^2\right) + R_2(x,y) \quad . \tag{21-45}$$

But we also have that $\alpha^2+\beta^2 = (x-x_0)^2 + (y-y_0)^2$ because $\mathbf{D}$ is so harmless:

$$\begin{aligned} \alpha^2+\beta^2 &= \left(\left[\begin{array}{cc} \alpha & \beta \end{array}\right]\cdot\mathbf{D}^\top\right)\cdot\left(\mathbf{D}\cdot\left[\begin{array}{c} \alpha \\ \beta \end{array}\right]\right) \\ &= \left[\begin{array}{cc} x-x_0 & y-y_0 \end{array}\right]\cdot\left[\begin{array}{c} x-x_0 \\ y-y_0 \end{array}\right] \\ &= (x-x_0)^2 + (y-y_0)^2 \\ &= \rho^2_{(x_0,y_0)}(x,y) \quad . \end{aligned} \tag{21-46}$$

If we next introduce the residual $R_2(x,y)$ with an epsilon term, we finally get:

$$f(x,y) \geq f(x_0,y_0) + \rho^2_{(x_0,y_0)}(x,y)\cdot\left(\lambda_2 + \varepsilon_f(x-x_0,y-y_0)\right) \quad . \tag{21-47}$$

We can hereby conclude that for all (x,y) sufficiently close to (but different from) the point of development (x_0,y_0) $f(x,y) > f(x_0,y_0)$ because $\lambda_2 > 0$ and $\varepsilon_f(x-x_0,y-y_0) \to 0$ for $(x,y) \to (x_0,y_0)$.

Similarly it can be concluded that $f(x,y) < f(x_0,y_0)$ for all points (x,y) sufficiently close to (but different from) (x_0,y_0) if both eigenvalues are negative.

If $\lambda_1 > 0$ and $\lambda_2 < 0$ we can use (21-44) directly with $\beta = 0$ and $\alpha = 0$, respectively, and by this obtain values of $f(x,y)$ that are larger and smaller than $f(x_0,y_0)$, respectively, for the corresponding points (x,y) sufficiently close to (but different from) (x_0,y_0) such that $f(x_0,y_0)$ can be neither a local minimum value nor a local maximum value. A similar conclusion is obtained when $\lambda_1 < 0$ and $\lambda_2 > 0$.

■

One could be led to believe that if only all second-order derivatives of $f(x,y)$ are positive at a stationary point (x_0, y_0) then the function must have a proper minimum point at (x_0, y_0). But it is not *that* simple. E.g. the following Hessian matrix does not have positive eigenvalues even though all elements are positive:

$$\mathbf{H}f(x,y) = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \quad . \tag{21-48}$$

The eigenvalues are $\lambda_1 = 3$ and $\lambda_2 = -1$, respectively, so according to Lemma 21.17 we can conclude that possible stationary points of $f(x,y)$ are neither a local maximum point nor minimum point of $f(x,y)$.

The matrix shown is the *constant* Hessian matrix of the function

$$f(x,y) = \frac{1}{2} \cdot \left(x^2 + y^2 + 4 \cdot x \cdot y\right) \quad . \tag{21-49}$$

The function has an obvious stationary point at $(x_0, y_0) = (0,0)$ but obviously no proper minimum point at $(0,0)$, which can also be seen in Figur 21.2. The function considered is its own approximating quadratic polynomial with development point at $(0,0)$.

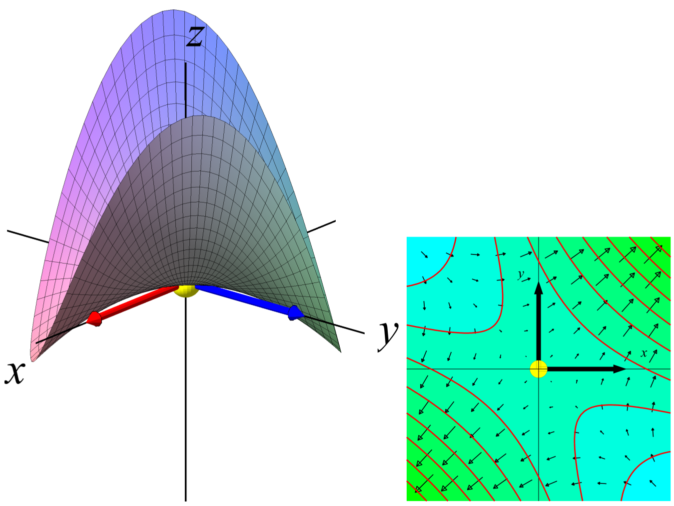

Figure 21.2: The graph of the function $f(x,y) = \frac{1}{2} \cdot \left(x^2 + y^2 + 4 \cdot x \cdot y\right)$ is shown to the left and the level curves of the function together with the gradient vector field to the right.

|||| Example 21.18 Inspection of a Stationary Point

Let the function $f(x,y)$ be given by

$$f(x,y) = x^3 + y^2 + x \cdot y \quad , \quad \nabla f(x,y) = (3 \cdot x^2 + y, x + 2 \cdot y) \quad , \quad \mathbf{H}f(x,y) = \begin{bmatrix} 6 \cdot x & 1 \\ 1 & 2 \end{bmatrix} \quad .$$

Then $(0,0)$ is a stationary point of $f(x,y)$ and the eigenvalues of $\mathbf{H}f(0,0)$ are $\lambda_1(x_0,y_0) = 1 + \sqrt{2} > 0$ and $\lambda_2(x_0,y_0) = 1 - \sqrt{2} < 0$, respectively. See Figure 21.3. There is obviously neither a local minimum nor a local maximum at the stationary point in accordance with Lemma 21.17.

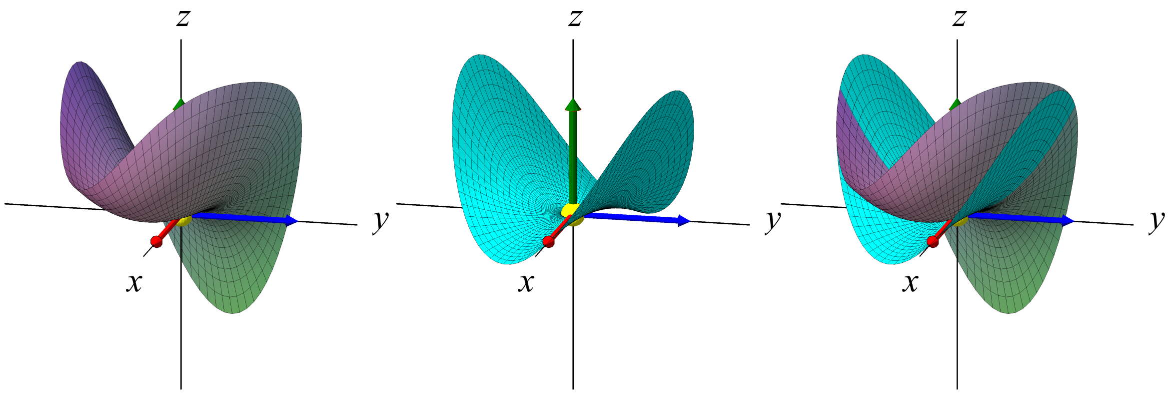

Figure 21.3: The graph of the function $f(x,y) = x^3 + y^2 + x \cdot y$ is shown to the left, the graph of the approximating quadratic polynomial with the point of development $(0,0)$ is shown in the middle, and both in the same plot to the right.

|||| Exercise 21.19

The function $f(x,y) = x^3 + y^2 + x \cdot y$ which we investigated in Example 21.18 also has a stationary point at $(x_0,y_0) = (1/12, -1/24)$. Investigate whether this stationary point is a local maximum or a local minimum or neither of these of $f(x,y)$.

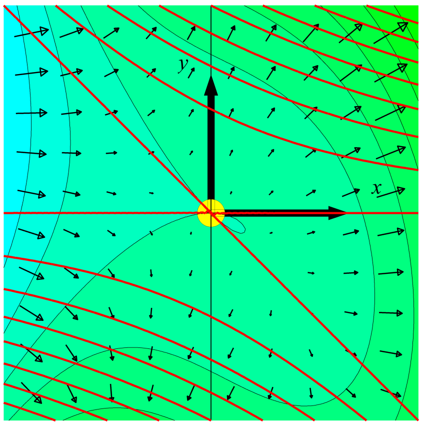

Figure 21.4: Level curves (black) and the gradient vector field of the function $f(x,y) = x^3 + y^2 + x \cdot y$ together with some level curves (red) for the approximating quadratic polynomial of $f(x,y)$ around the stationary point $(x_0, y_0) = (0,0)$.

|||| Example 21.20 Inspection of a Stationary Point

Let the function $f(x,y)$ be given by the following data in $\mathbb{R}^2$:

$$\begin{aligned} f(x,y) &= \sin(x) \cdot \cos(x \cdot y) \quad , \\ \nabla f(x,y) &= (\cos(x)\cos(x \cdot y) - y \cdot \sin(x) \cdot \sin(x \cdot y), -x \cdot \sin(x) \cdot \sin(x \cdot y)) \quad . \end{aligned} \tag{21-50}$$

Then $(x_0, y_0) = (\pi/2, 0)$ is a stationary point of $f(x,y)$ and at this point we have

$$\mathbf{H}f(\pi/2, 0) = \begin{bmatrix} -1 & 0 \\ 0 & -\pi^2/4 \end{bmatrix} \tag{21-51}$$

In the stationary point the Hessian matrix thus has the eigenvalues -1 and $-\pi^2/4$. We conclude that the stationary point is a proper local maximum point of $f(x,y)$ in accordance with the inspection of Figure 21.5 – and in accordance with Lemma 21.17. The approximating quadratic polynomial of $f(x,y)$ with the development point at the stationary point $(x_0, y_0) = (\pi/2, 0)$ is:

$$\begin{aligned} P_{2,(\pi/2,0)}(x,y) &= 1 - \frac{1}{2}\left(x - \frac{\pi}{2}\right)^2 - \frac{1}{2} \cdot \frac{\pi^2}{4} \cdot y^2 \\ &= \left(1 - \frac{\pi^2}{8}\right) + \frac{\pi}{2} \cdot x - \frac{1}{2} \cdot x^2 - \frac{\pi^2}{8} \cdot y^2 \quad . \end{aligned} \tag{21-52}$$

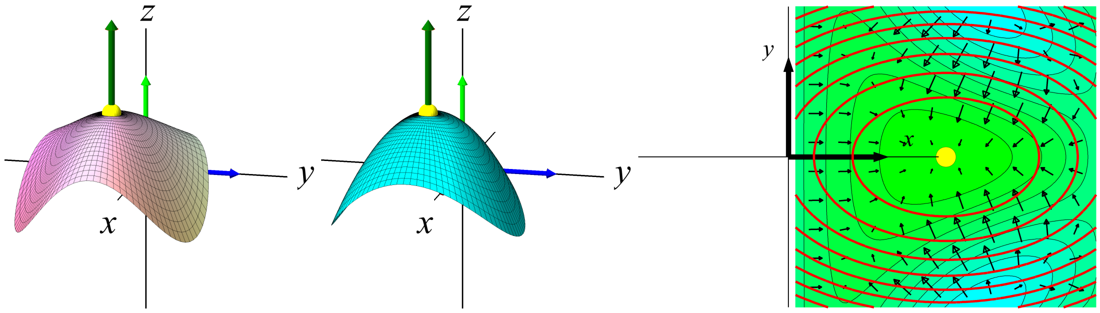

Figure 21.5: To the left: The graph of the function $f(x,y) = \sin(x) \cdot \cos(x \cdot y)$. Centre: The graph of the approximating quadratic polynomial of $f(x,y)$ with the point of development at the stationary point $(\pi/2, 0)$. To the right: The gradient vector field of the function $f(x,y)$, its level curves (black) and the level curves of the approximating quadratic polynomial (red).

|||| Example 21.21 Inspection of a Stationary Point

Let the function $f(x,y)$ be given by the following data in $\mathbb{R}^2$:

$$\begin{aligned} f(x,y) &= 1 + x^2 + (1-y)^2 + x \cdot y \quad , \\ \nabla f(x,y) &= (2 \cdot x + y, 2 \cdot y - 2 + x) \quad . \end{aligned} \tag{21-53}$$

Then $(x_0, y_0) = (-2/3\,,\,4/3)$ is a stationary point of $f(x,y)$ and at this point we have as at any other point

$$\mathbf{H}f(x_0, y_0) = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \tag{21-54}$$

Therefore at the stationary point the Hessian matrix has the eigenvalues 3 and 1. We conclude that the stationary point is a proper local minimum point of $f(x,y)$ in accordance with the inspection of Figure 21.6 and in accordance with Lemma 21.17.

The approximating quadratic polynomial of $f(x,y)$ with the point of development at the stationary point $(x_0, y_0) = (\pi/2, 0)$ is the function itself:

$$P_{2,(-2/3\,,4/3)}(x,y) = f(x,y) \quad . \tag{21-55}$$

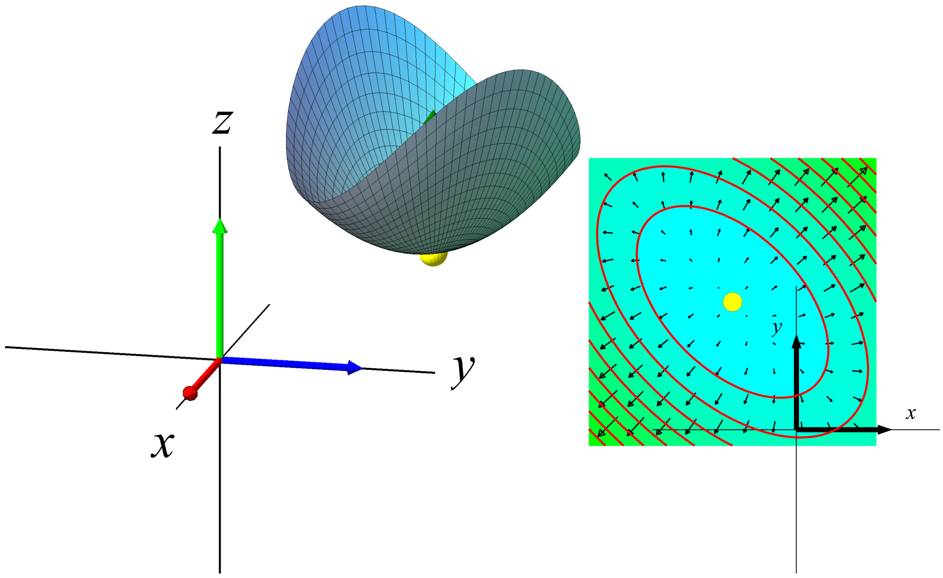

Figure 21.6: To the left: The graph of the function $f(x,y) = 1 + x^2 + (1-y)^2 + x \cdot y$. To the right: The gradient vector field of the function $f(x,y)$, and its level curves (red) around the stationary point $(x_0, y_0) = (-2/3\,,\,4/3)$. Note that the function is its own approximating quadratic polynomial with the point of development at the stationary point.

|||| Example 21.22 Inspection of a Stationary Point

Let the function $f(x,y)$ be given by the following data in $\mathbb{R}^2$:

$$\begin{aligned} f(x,y) &= \frac{1}{2} \cdot \sin(\pi \cdot x^2 + \pi \cdot y^2) \quad , \\ \nabla f(x,y) &= (\pi \cdot x \cdot \cos(\pi \cdot x^2 + \pi \cdot y^2), \pi \cdot y \cdot \cos(\pi \cdot x^2 + \pi \cdot y^2)) \quad . \end{aligned} \tag{21-56}$$

Then $(x_0, y_0) = (0, 1/\sqrt{2})$ is one of many stationary points of $f(x,y)$ and at that point we have

$$\mathbf{H}f(x_0, y_0) = \begin{bmatrix} 0 & 0 \\ 0 & -\pi^2 \end{bmatrix} . \tag{21-57}$$

Therefore at the stationary point the Hessian matrix has the eigenvalues 0 and $-\pi^2$ and therefore Lemma 21.17 cannot decide whether we have a local maximum or minimum at the point. But since $f(0, 1/\sqrt{2}) = 1/2$, which is the maximum value that $f(x,y)$ can assume, then $(0, 1/\sqrt{2})$ is a local maximum point of $f(x,y)$ but it is not a proper local maximum point because all points on the circle $x^2 + y^2 = 1/2$ are also local maximum points of $f(x,y)$. See also Figure 21.7.

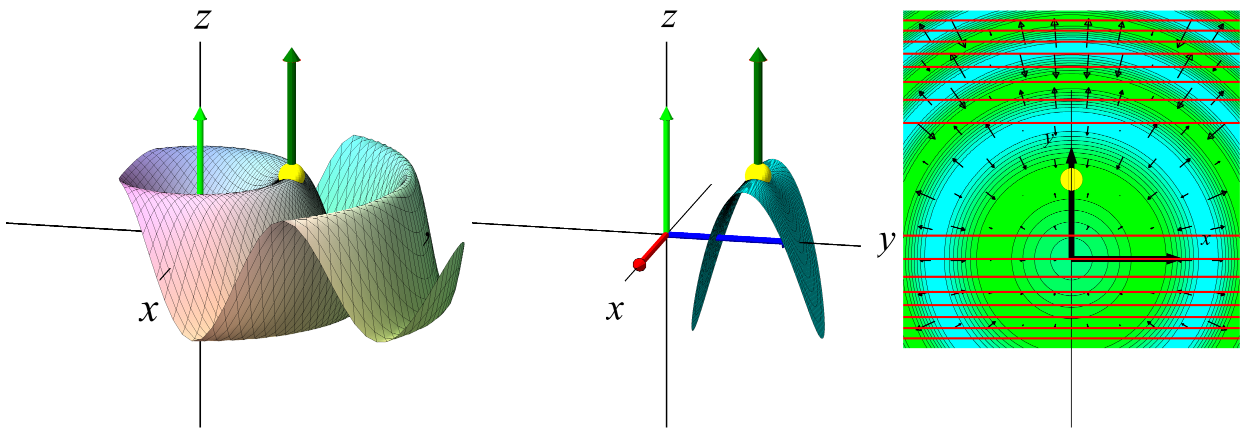

Figure 21.7: To the left: The graph of the function $f(x,y) = \frac{1}{2} \cdot \sin(\pi \cdot x^2 + \pi \cdot y^2)$. In the middle: The approximating quadratic polynomial of the function with the point of development $(0, 1/\sqrt{2})$. To the right: The gradient vector field of the function $f(x,y)$ and its level curves (black) and the level curves of the approximating quadratic polynomial (red) around the stationary point $(x_0, y_0) = (0, 1/\sqrt{2})$.

|||| Example 21.23 Inspection of a Stationary Point

We consider a function $f(x,y)$ that is defined by

$$f(x,y) = \frac{1}{5} \cdot \left(1 - (1-y)^2 - x^2\right) \cdot \left(4 - (2-y)^2 - x^2\right) \quad . \tag{21-58}$$

then $(x_0, y_0) = (0,0)$ is a stationary point of $f(x,y)$ and at this point we have

$$\mathbf{H}f(0,0) = \begin{bmatrix} 0 & 0 \\ 0 & 16/5 \end{bmatrix} \tag{21-59}$$

In the stationary points the Hessian matrix has the eigenvalues 0 and $16/5$. Therefore Lemma 21.17 cannot decide whether we have a local maximum or minimum at the point. It is, though, relatively easy to see that arbitrarily close to $(0,0)$ points (x,y) exist in which the functional values $f(x,y)$ are negative; and at the same time arbitrarily close to $(0,0)$ points (x,y) exist where the functional values $f(x,y)$ are positive. See figure 21.8. Therefore the stationary point $(0,0)$ is not a local minimum point of $f(x,y)$.

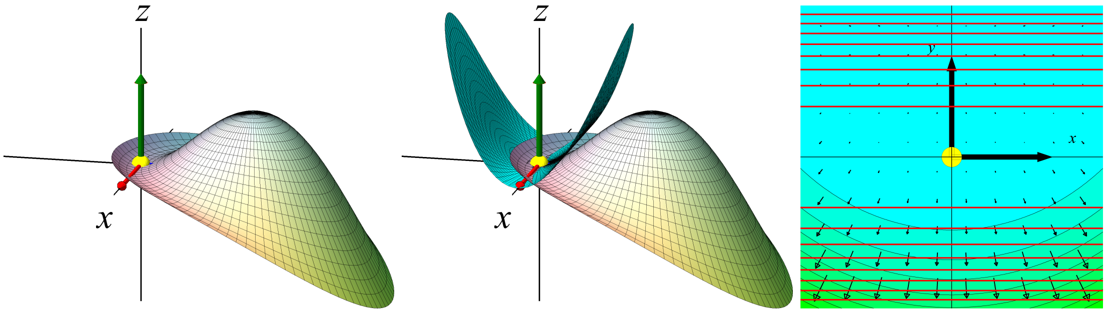

Figure 21.8: To the left: The graph of the function $f(x,y) = \frac{1}{5} \cdot (1-(1-y)^2 - x^2) \cdot (4-(2-y)^2 - x^2)$. Centre: The graph of $f(x,y)$ together with the graph of the approximating quadratic polynomial of the function with the point of development $(0,0)$. To the right: The gradient vector field of the function $f(x,y)$ and its level curves (black) and the level curves of the approximating quadratic polynomial (red) around the stationary point $(x_0, y_0) = (0,0)$.

The function that is inspected in Example 21.23 has the following surprising property: Let $\mathbf{r}(u)$ be the parametric representation of an arbitrary line through $\mathbf{r}(0) = (0,0)$. Then the composite (height) function $h(u) = f(\mathbf{r}(u))$ has a proper local minimum at $u = 0$. A complete investigation of a function of two variables about a stationary point cannot be obtained by only investigating the restrictions of the function to the straight lines through the point!

21.5 Summary

This eNote is about the analysis of functions of two variables.

- We use the compact notation for the gradient vector field and the Hessianmatrix given by the derivatives of $f(x,y)$ until and including second order:

$$\nabla f(x_0,y_0) = (f'_x(x_0,y_0), f'_y(x_0,y_0)) \quad , \tag{21-60}$$

$$\mathbf{H}f(x_0,y_0) = \begin{bmatrix} f''_{xx}(x_0,y_0) & f''_{xy}(x_0,y_0) \\ f''_{yx}(x_0,y_0) & f''_{yy}(x_0,y_0) \end{bmatrix} \quad . \tag{21-61}$$

- The approximating polynomial of the second order of $f(x,y)$ is then given by:

$$\begin{aligned} P_{2,(x_0,y_0)}(x,y) &= f(x_0,y_0) + \nabla f(x_0,y_0)\cdot(x-x_0, y-y_0) \\ &+ \frac{1}{2}\cdot\begin{bmatrix} x-x_0 & y-y_0 \end{bmatrix}\cdot \mathbf{H}f(x_0,y_0)\cdot\begin{bmatrix} x-x_0 \\ y-y_0 \end{bmatrix} \quad . \end{aligned} \tag{21-62}$$

- Taylor's limit formula for functions of two variables can thus be written

$$\begin{aligned} f(x,y) = f(x_0,y_0) &+ \nabla f(x_0,y_0)\cdot(x-x_0, y-y_0) \\ &+ \frac{1}{2}\cdot\begin{bmatrix} x-x_0 & y-y_0 \end{bmatrix}\cdot \mathbf{H}f(x_0,y_0)\cdot\begin{bmatrix} x-x_0 \\ y-y_0 \end{bmatrix} \\ &+ \rho^2_{(x_0,y_0)}(x,y)\cdot\varepsilon_f(x-x_0,y-y_0) \quad . \end{aligned} \tag{21-63}$$

- Stationary points (x_0,y_0) of functions $f(x,y)$ of two variables are characterized by $\nabla f(x_0,y_0) = (0,0)$.

- The eigenvalues of the Hessian matrix at a stationary point can assist us in the decision whether the point is a proper local maximum or minimum point of a smooth function $f(x,y)$. If both eigenvalues are positive, then the point is a proper local minimum point. If they both are negative then the point is a proper local maximum point. If the eigenvalues both are different from 0 and are of opposite sign then the point is neither a maximum nor a minimum point at the stationary point.