eNote 20

Gradients and Tangent Planes

In this eNote we will take a closer look at the geometric analysis and inspection of functions of two variables. In particular we will study the relation between the gradient vectors in the (x, y) plane and the tangent planes in (x, y, z) space. We will look at the parametric representations of – and the normal vectors to – the tangent planes to the graph of a function $f(x, y)$ of two variables via the approximating first-degree polynomial for $f(x, y)$. The partial derivatives in the whole eNote are the basic ingredients. A global fact concerning the partial derivatives is that the gradient can be used for the determination of the function everywhere in a connected region if we know a function value at just a single point. Curves in the (x, y) plane can be elevated to the graph surface of $f(x, y)$, and we will show that all tangents to any curve through a given point on the graph surface are contained within the tangent plane to the surface at the point.

Updated: 10.1.22, DB.
Updated: 2.2.23, shsp.

20.1 The Gradients Determine the Function

For functions of one variable we know that we can find and reconstruct the function $f(x)$ from the derivative $f'(x) = q(x)$ and from the value of $f(x)$ at just a single point x_0. We just have to integrate and find an indefinite integral to $q(x)$. The indefinite integral is then the function $f(x)$ apart from a constant that in the end can be determined from the knowledge of $f(x_0)$. The same applies to functions of two variables.

Even though we only know the partial derivative, i.e. ordinary derivative of the auxiliary functions $f_1(x, y_0)$ and $f_2(x_0, y)$, in the two *coordinate axis directions*, then they are in fact sufficient to reconstruct the function. But in order to integrate and reconstruct the function from the gradient vector field we need the region we consider in the (x, y) plane to be connected:

|||| Definition 20.1 Connected Set

A set M in the plane is said to be ***path-connected*** or just ***connected*** if every point in M can be connected to another point in M via a differentiable parameterized curve $\mathbf{r}(u) = (p(u), q(u))$, where $u \in]\alpha, \beta[$.

|||| Example 20.2 Star-Shaped Sets are Connected

Every star-shaped set in the (x, y) plane is path-connected. We can connect two given points via the straight line segments to the star point, where a 'kink' is typically present. Consider why this is not a problem.

We then have the following theorem, which we will prove constructively:

|||| Theorem 20.3 The Gradient Vector Field Gives the Function Apart from a Constant

Suppose that we know the gradient vector field $\nabla f(x, y)$ of a function $f(x, y)$ everywhere in an open path-connected region in the domain. And suppose we know the function value at an individual point (x_0, y_0). Then we can reconstruct all function values of $f(x, y)$ in all of the region.

|||| Proof

We use the chain rule from eNote 15 on the composite function $h(u) = f(\mathbf{r}(u))$ where $\mathbf{r}(u)$, $u \in]u_0, u_1[$, is an arbitrary differentiable curve from $(x_0, y_0) = \mathbf{r}(u_0)$ to $(x_1, y_1) = \mathbf{r}(u_1)$ and thereby get:

$$h'(u) = \mathbf{r}'(u) \cdot \nabla f(\mathbf{r}(u)) \quad . \tag{20-1}$$

From this it follows that $h(u)$ is an indefinite integral of the function on the right-hand side of the above equation:

$$h(u_1) - h(u_0) = \int_{u_0}^{u_1} \mathbf{r}'(u) \cdot \boldsymbol{\nabla} f(\mathbf{r}(u))\, \mathrm{d}u \quad . \tag{20-2}$$

But since

$$\begin{aligned} h(u_0) &= f(\mathbf{r}(u_0)) = f(x_0, y_0) \quad , \\ h(u_1) &= f(\mathbf{r}(u_1)) = f(x_1, y_1) \quad , \end{aligned} \tag{20-3}$$

we thus get

$$f(x_1, y_1) = f(x_0, y_0) + \int_{u_0}^{u_1} \mathbf{r}'(u) \cdot \boldsymbol{\nabla} f(\mathbf{r}(u))\, \mathrm{d}u \quad , \tag{20-4}$$

And this was precisely what we should do – reconstruct the value of $f(x, y)$ at the point (x_1, y_1) from the value of $f(x, y)$ at the point (x_0, y_0) and the gradient vector field.

■

As a direct consequence of Theorem 20.3 and the proof thereof, in particular Equation (20-4), we have

|||| Theorem 20.4 Zero Gradient Gives Constant Function

If a function $f(x, y)$ has the gradient $\boldsymbol{\nabla} f(x, y) = (0, 0)$ everywhere in an open path-connected region then the function is constant in the whole region.

Remark that it is arbitrary which curve one chooses for the integration path in order to find the function value at the end point. If it is possible one would naturally choose the simplest curve for the purpose. This we will do in the following example:

|||| Example 20.5 Determination of the Function from the Partial Derivatives

We want to determine the function $f(x, y)$ entirely from our knowledge of the function value $f(0, 0) = 0$ and knowledge of the gradient vector field in the (x, y) plane:

$$\boldsymbol{\nabla} f(x, y) = (6\,x + 10\,y^7\,,\; 21\,y^2 + 70\,x\,y^6) \quad . \tag{20-5}$$

Let $\mathbf{r}(u)$ be a parametric form of the *straight line segment* in the (x, y) plane from $(x_0, y_0) = (0, 0)$ to an arbitrary point $(x_1, y_1$ in the (x, y) plane:

$$\mathbf{r}(u) = (0, 0) + u \cdot (x_1, y_1) = (u\,x_1, u\,y_1) \quad , \quad \text{where } u \in]0, 1[\quad , \tag{20-6}$$

such that

$$\mathbf{r}'(u) = (x_1, y_1) \quad ,$$

$$\nabla f(\mathbf{r}(u)) = (6ux_1 + 10u^7 y_1^7 \,,\; 21u^2 y_1^2 + 70u^7 x_1 y_1^6) \quad , \tag{20-7}$$

$$\mathbf{r}'(u) \cdot \nabla f(\mathbf{r}(u)) = 6ux_1^2 + 10u^7 x_1 y_1^7 + 21u^2 y_1^3 + 70u^7 x_1 y_1^7 \quad .$$

From this we then get by using formula (20-4) that

$$\begin{aligned} f(x_1, y_1) &= 0 + \int_0^1 \left(6ux_1^2 + 10u^7 x_1 y_1^7 + 21u^2 y_1^3 + 70u^7 x_1 y_1^7\right) \mathrm{d}u \\ &= x_1^2 \int_0^1 6u \,\mathrm{d}u \;+\; 80x_1 y_1^7 \int_0^1 u^7 \,\mathrm{d}u \;+\; 21y_1^3 \int_0^1 u^2 \,\mathrm{d}u \\ &= 3x_1^2 + 10x_1 y_1^7 + 7y_1^3 \quad . \end{aligned} \tag{20-8}$$

Therefore we get the reconstructed function:

$$f(x, y) = 3\,x^2 + 7\,y^3 + 10\,x\,y^7 \quad . \tag{20-9}$$

For the sake of completeness we can test and investigate whether this function actually is 0 at $(0, 0)$ (it is) and moreover if it has the gradient vector field, given by Equation (20-5) (it has).

|||| Exercise 20.6

A function $f(x, y)$ is known to be differentiable in all of $\mathbb{R}^2$ with $f(1, 0) = 1$ and

$$\nabla f(x, y) = (\cos(y), -x\,\sin(y)) \quad . \tag{20-10}$$

Determine $f(x, y)$ at every (x, y).

20.1.1 Gradient Vector Fields

The gradient vector of a function of two variables $f(x, y)$ contains information about the local increase of the function around every point (x_0, y_0):

|||| Theorem 20.7 The Gradient Points to the Direction with the Largest Increase

Let $f(x,y)$ denote a function of two variables that has a proper gradient vector at a point (x_0,y_0), i.e. $\nabla f(x_0,y_0) \neq \mathbf{0}$. Then it applies that:

- The directional derivative $f'((x_0,y_0);\mathbf{e})$ of $f(x,y)$ at the point (x_0,y_0) is *largest* in the direction $\mathbf{e}$ that is determined by the gradient vector $\nabla f(x_0,y_0)$, i.e. for

$$\mathbf{e} = \mathbf{e}_{max} = \frac{\nabla f(x_0,y_0)}{|\,\nabla f(x_0,y_0)\,|} \tag{20-11}$$

- The directional derivative $f'((x_0,y_0);\mathbf{e})$ of $f(x,y)$ is *smallest* in the opposite direction ($\mathbf{e}_{min}$) of that ($\mathbf{e}_{max}$) determined by the gradient vector at the point:

$$\mathbf{e} = \mathbf{e}_{min} = -\frac{\nabla f(x_0,y_0)}{|\,\nabla f(x_0,y_0)\,|} \tag{20-12}$$

- The directional derivative is 0 in both of the two (contour) directions that are perpendicular to the gradient vector.

|||| Proof

We let $\mathbf{g}$ denote the unit vector that states the *direction* of the gradient :

$$\mathbf{g} = \frac{\nabla f(x_0,y_0)}{|\nabla f(x_0,y_0)|} \quad . \tag{20-13}$$

Every (other) unit direction vector $\mathbf{e}$ in the plane can then 'be resolved' relative to the direction of $\mathbf{g}$ in the following way

$$\mathbf{e} = \cos(\varphi)\cdot\mathbf{g} + \sin(\varphi)\cdot\mathbf{g}^{\perp} \quad , \tag{20-14}$$

where φ is the angle between the two unit vectors $\mathbf{e}$ and $\mathbf{g}$ and where $\mathbf{g}^{\perp}$ denotes the orthogonal vector to $\mathbf{g}$. Then the directional derivative of $f(x,y)$ at the point (x_0,y_0) is:

$$\begin{aligned} f'((x_0,y_0);\mathbf{e}) &= \nabla f(x_0,y_0) \bullet \mathbf{e} \\ &= (|\nabla f(x_0,y_0)|\cdot\mathbf{g}) \bullet \mathbf{e} \\ &= (|\nabla f(x_0,y_0)|\cdot\mathbf{g}) \bullet \left(\cos(\varphi)\cdot\mathbf{g} + \sin(\varphi)\cdot\mathbf{g}^{\perp}\right) \\ &= |\nabla f(x_0,y_0)|\cdot\cos(\varphi)\cdot\left(\mathbf{g}\bullet\mathbf{g} + \mathbf{g}\bullet\mathbf{g}^{\perp}\right) \\ &= |\nabla f(x_0,y_0)|\cdot\cos(\varphi)\cdot(1+0) \\ &= |\nabla f(x_0,y_0)|\cdot\cos(\varphi) \quad . \end{aligned} \tag{20-15}$$

Since $\cos(\varphi)$ is largest (with the value 1) when $\varphi = 0$ (i.e. when **e** and **g** point in the *same* direction) and since $\cos(\varphi)$ is smallest (with the value -1) when $\varphi = \pi$ (i.e. when **e** and **g** point in *opposite* directions), and since $\cos(\varphi)$ is 0 when $\varphi = \pi/2$ (i.e. when **e** and **g** are orthogonal), then the three statements in the theorem follow. The last statement in the theorem we already know from eNote 15 where we saw that the directional derivative is 0 in the directions that are perpendicular to the gradient vector.

■

If we are positioned at the point (x_0, y_0) in the (x, y) plane and wish to move to neighboring points with *higher function values* $f(x, y)$ then it is most efficient to move in the direction in which the gradient of $f(x, y)$ points from that point, that is, in the direction of the unit directional vector $\nabla f(x_0, y_0)/|\nabla f(x_0, y_0)|$.

If we want to move to neighboring points with *lower functional values*, then it is most efficient to go in the opposite direction of the gradient at (x_0, y_0).

If we wish to go to neighboring points with the *same functional value* $f(x_0, y_0)$ then we only need to go along the contour $\mathcal{K}_{f(x_0,y_0)}$ for $f(x, y)$ that goes through (x_0, y_0) (!) – this corresponds (locally) to going along one of the two (contour) directions that are determined by a vector that is *perpendicular* to the gradient vector at the point.

|||| Example 20.8 Geometric Analysis of a Function of Two Variables

A function is given by

$$f(x, y) = 3 + 2\,\mathrm{e}^{-x^2-2y^2} \quad . \tag{20-16}$$

The partial derivatives of $f(x, y)$ at (x, y) are then

$$\begin{aligned} f'_x(x, y) &= -4x\mathrm{e}^{-x^2-2y^2} \\ f'_y(x, y) &= -8y\mathrm{e}^{-x^2-2y^2} \quad . \end{aligned} \tag{20-17}$$

The gradient vector field is therefore

$$\nabla f(x, y) = \left(-4x\mathrm{e}^{-x^2-2y^2},\ -8y\mathrm{e}^{-x^2-2y^2}\right) \quad . \tag{20-18}$$

This vector field is sketched in Figure 20.1 together with the level curves of $f(x, y)$ and together with a curve (a circle) with the parametric representation

$$\mathbf{r}(u) = \left(\frac{1}{2} + \cos(u)\,,\ \frac{1}{2} + \sin(u)\right) \quad , \quad \text{in which} \quad u \in\,] - \pi, \pi[\quad . \tag{20-19}$$

Note that the gradient vector field everywhere is perpendicular to the level curves and that they all point towards the centre where the function values are largest – see the graph of the function in Figure 20.2.

Also note that at exactly two points on the circle the gradient is perpendicular to the circle. Therefore at these points the composite height function $h(u) = f(\mathbf{r}(u))$ has the derivative $h'(u) = 0$. The circle has at these points the same tangent as the respective level curves through the two points. We can determine the two points by determining the two values of u that solve the equation $h'(u) = \boldsymbol{\nabla} f(\mathbf{r}(u)) \cdot \mathbf{r}'(u) = 0$.

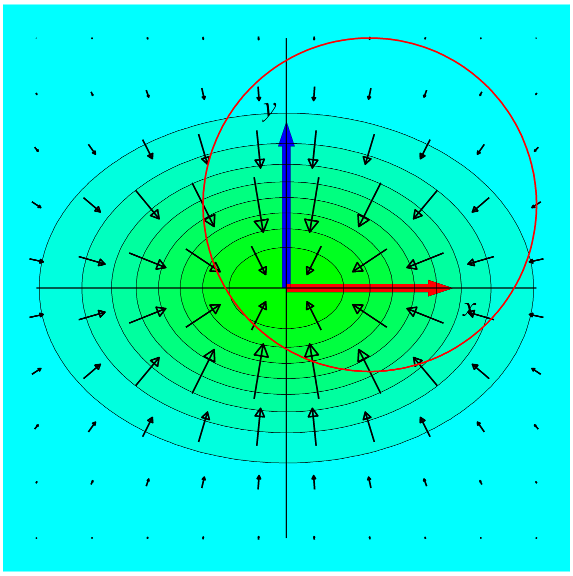

Figure 20.1: Contours and gradient vector field of the function $f(x,y) = 3 + 2\,\mathrm{e}^{-x^2-2y^2}$.

|||| Exercise 20.9

Determine the gradient vector fields of the following functions and sketch the vector fields by drawing a suitable number of the gradient vectors $\boldsymbol{\nabla} f(x_0, y_0)$ with different points (x_0, y_0) in the (x, y) plane:

$$\begin{aligned} f(x,y) &= 3x + y \quad , \\ f(x,y) &= x^2 + y^2 \quad , \\ f(x,y) &= y \cdot \mathrm{e}^x \quad . \end{aligned} \tag{20-20}$$

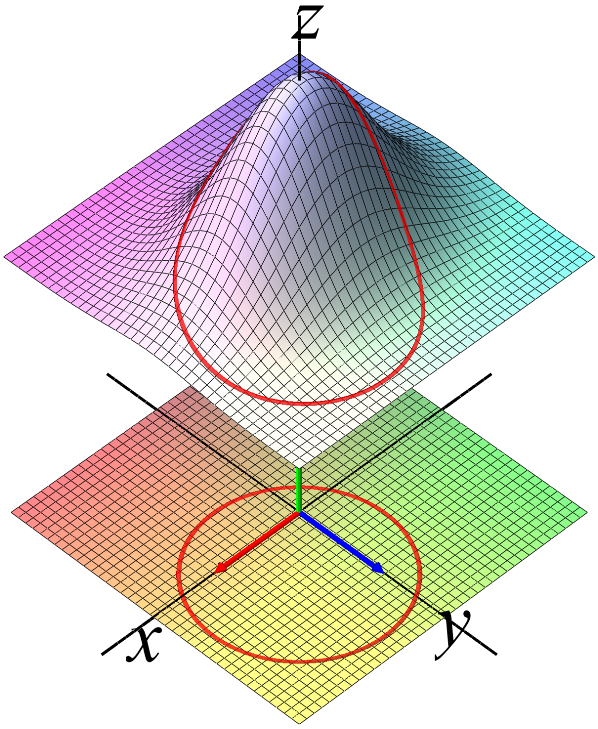

Figure 20.2: The graph in (x, y, z) space of $f(x, y) = 3 + 2\,e^{-x^2-2y^2}$.

If we move around the circle in the (x, y) plane in Figure 20.1 (e.g. in the positive direction) and at the same time keep an eye on which level curves we cross on our way we will be able to decide whether we move towards increasing or decreasing function values – on the one hand by keeping an eye on the change in color and on the other hand by seeing whether the projection of the gradient vector onto the direction of movement (the directional derivative) is positive or negative; if the projection is positive we are heading towards larger function values, if the projection is negative we are heading towards smaller function values. The same thing is observable from the graph of the function, where at each position of the elevated curve it is clear whether we are on our way up or down on the graph surface.

20.2 Elevated Curves on Graph Surfaces

Motivated by the graph surface with the elevated circle in Figure 20.2 and motivated by the analysis in Example 20.8 we will now entirely in general define and study elevated curves:

Definition 20.10 Elevated Curves

We let $f(x,y)$ denote a function of two variables and let $\mathbf{r}(u)$ denote a differentiable parametrized curve in the (x,y) plane:

$$\mathbf{r}(u) = (p(u), q(u)) \quad , \quad \text{where } u \in \,]\alpha, \beta[\quad , \tag{20-21}$$

and where $p(u)$ and $q(u)$ denote two differentiable functions of one variable u.

The ***elevated curve*** that corresponds to the plane curve $\mathbf{r}(u)$ is defined to be the following curve in the (x,y,z) space:

$$\widetilde{\mathcal{C}}_{\mathbf{r}}(f) \quad : \quad \widetilde{\mathbf{r}}(u) = (p(u), q(u), f(p(u), q(u))) \quad , \quad \text{where } u \in \,]\alpha, \beta[\quad . \tag{20-22}$$

The elevated curve $\widetilde{\mathbf{r}}(u)$ is a spatial curve that lies on the graph surface $\mathcal{G}(f)$ of the function $f(x,y)$.

The projection (vertical, along the z-axis) of $\widetilde{\mathbf{r}}(u)$ on the (x,y) plane is of course the plane curve $\mathbf{r}(u)$.

The elevated curves $\widetilde{\mathbf{r}}(u)$ have the tangent vectors and tangent lines in (x,y,z) space – and they must depend partly on the curve $\mathbf{r}(u)$ in the (x,y) plane and partly on the function $f(x,y)$. The following expression can be found directly from the chain rule for the function $f(x,y)$ along the curve $\mathbf{r}(u)$ – i.e. the chain rule for the composite function $h(u) = f(\mathbf{r}(u)) = f(p(u), q(u))$:

|||| Theorem 20.11 Tangents to Elevated Curves

The u-derivative of the elevated curve $\widetilde{\mathbf{r}}(u)$ is given by the u-derivative of the three coordinate functions:

$$\begin{aligned}\widetilde{\mathbf{r}}'(u) &= \left(p'(u), q'(u), \frac{d}{du}f(p(u), q(u))\right)\\ &= (p'(u), q'(u), h'(u))\\ &= (p'(u), q'(u), \nabla f(p(u), q(u)) \cdot (p'(u), q'(u))) \quad .\end{aligned} \tag{20-23}$$

The *tangent* to the elevated curve on the graph surface of $f(x, y)$ is therefore determined by the following expression at a given curve point $\widetilde{\mathbf{r}}(u_0) = (p(u_0), q(u_0), f(p(u_0), q(u_0)) = (x_0, y_0, f(x_0, y_0))$:

$$\begin{aligned}\widetilde{L}_{u_0} \ : \ \widetilde{\mathbf{T}}(t) &= \widetilde{\mathbf{r}}(u_0) + t \cdot \widetilde{\mathbf{r}}'(u_0)\\ &= (x_0, y_0, f(x_0, y_0)) +\\ &\quad t \cdot (p'(u_0), q'(u_0), \nabla f(x_0, y_0) \cdot (p'(u_0), q'(u_0))) \ ,\end{aligned} \tag{20-24}$$

where the parameter t runs through all of $\mathbb{R}$.

20.2.1 Elevated Coordinate Curves

The coordinate curves in the (x, y) plane are the particularly simple curves (straight lines) that are parallel to the coordinate axes. Through the point (x_0, y_0) we have the following two:

$$\begin{aligned}\mathbf{r}_1(u) &= (u, y_0) \quad \text{where } u \in \mathbb{R} \quad ,\\ \mathbf{r}_2(u) &= (x_0, u) \quad \text{where } u \in \mathbb{R} \quad .\end{aligned} \tag{20-25}$$

Therefore they have also particularly simple elevations to the graph surface of a given function $f(x, y)$:

$$\begin{aligned}\widetilde{\mathbf{r}}_1(u) &= (u, y_0, f(u, y_0)) \quad , \quad u \in \mathbb{R} \quad ,\\ \widetilde{\mathbf{r}}_2(u) &= (x_0, u, f(x_0, u)) \quad , \quad u \in \mathbb{R} \quad ,\end{aligned} \tag{20-26}$$

with particularly simple u-derivatives for every u:

$$\begin{aligned}\widetilde{\mathbf{r}}_1'(u) &= (1, 0, f_x'(u, y_0)) \quad , \quad u \in \mathbb{R} \quad ,\\ \widetilde{\mathbf{r}}_2'(u) &= (0, 1, f_y'(x_0, u)) \quad , \quad u \in \mathbb{R} \quad .\end{aligned} \tag{20-27}$$

The tangents from the point $(x_0, y_0, f(x_0, y_0))$ to the elevated coordinate curves are therefore also fairly simple:

$$\begin{aligned} \widetilde{L}_1 \quad &: \quad (x_0, y_0, f(x_0, y_0)) + t \cdot (1, 0, f'_x(x_0, y_0)) \quad , \quad t \in \mathbb{R} \\ \widetilde{L}_2 \quad &: \quad (x_0, y_0, f(x_0, y_0)) + t \cdot (0, 1, f'_y(x_0, y_0)) \quad , \quad t \in \mathbb{R} \end{aligned} \tag{20-28}$$

|||| Exercise 20.12

Determine (for every point (x_0, y_0) in the (x, y) plane) the tangents to the elevated coordinate curves through the graph surface point $(x_0, y_0, f(x_0, y_0))$ of the functions:

$$\begin{aligned} f(x, y) &= 3x + y \quad , \\ f(x, y) &= x^2 + y^2 \quad , \\ f(x, y) &= y \cdot e^x \quad . \end{aligned} \tag{20-29}$$

In the Figure 20.3 the elevated circle and the graph surface of the function $f(x, y)$ from Example 20.8 are seen. The tangents of the elevated curve are drawn from a couple of chosen points. Note that the spatial tangents to the elevated curve project down on the corresponding tangents to the circle in the (x, y) plane – as it also follows from equation (20-24).

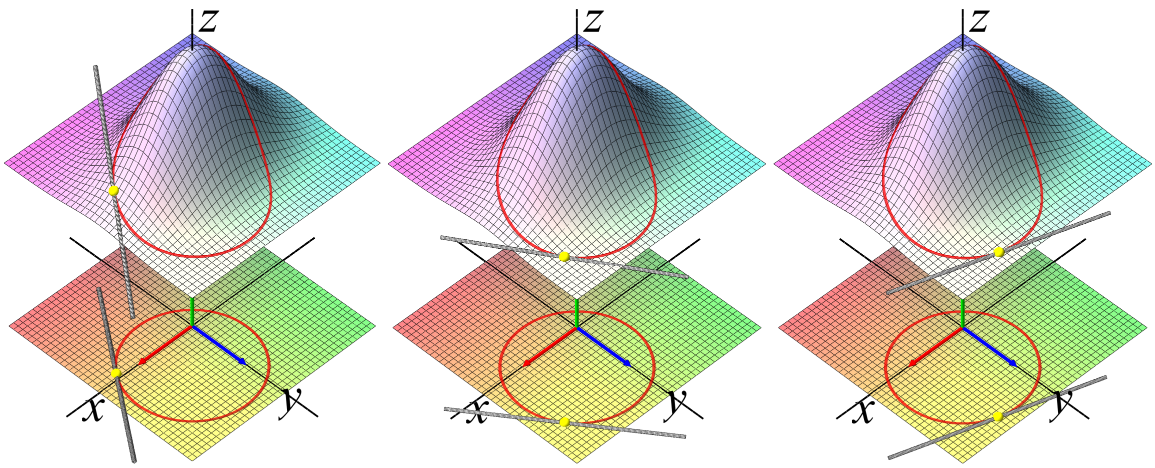

Figure 20.3: The graph of $f(x, y) = 3 + 2\,e^{-x^2-2y^2}$ with chosen tangents to an elevated circle.

The tangent plane to the graph surface at a given point $(x_0, y_0, f(x_0, y_0))$ is in itself the graph surface of the approximating first-degree polynomial for $f(x, y)$ with the development point (x_0, y_0) and is therefore independent of which curve we choose to elevate to the graph surface through the point $(x_0, y_0, f(x_0, y_0))$! Nevertheless in the figures it looks as though the *tangents to the curves* lie entirely in their respective tangent planes.

This inspection therefore gives rise to the following supposition which we will prove in Theorem 20.13 below: Every tangent to every elevated curve through a given point $(x_0, y_0, f(x_0, y_0))$ lies within the tangent surface to $f(x, y)$ at the point. And vice versa: Every straight line in the tangent plane that goes through the point $(x_0, y_0, f(x_0, y_0))$ is the tangent to some elevated curve through the point.

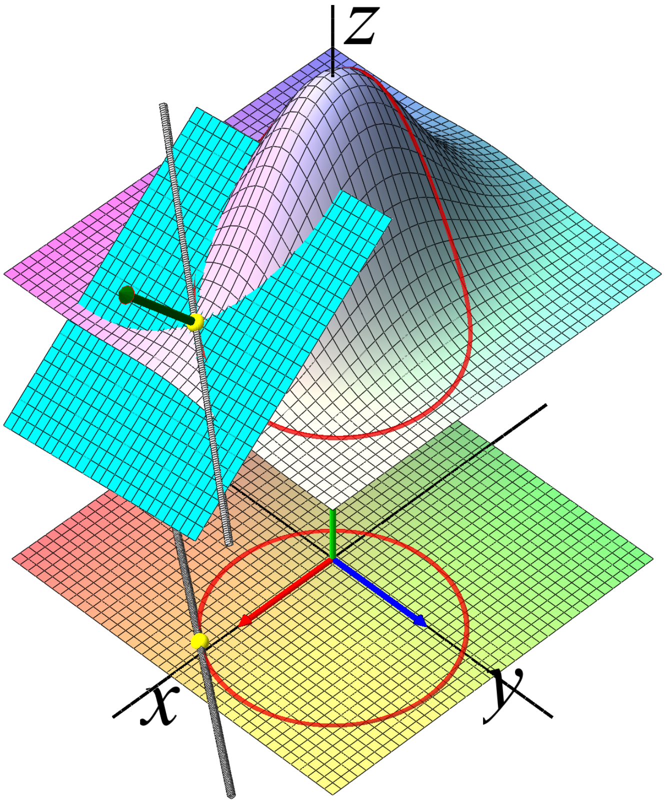

Figure 20.4: The graph of $f(x, y) = 3 + 2\,e^{-x^2-2y^2}$ and the tangent plane through a chosen point on the elevated circle, see Figure 20.2.

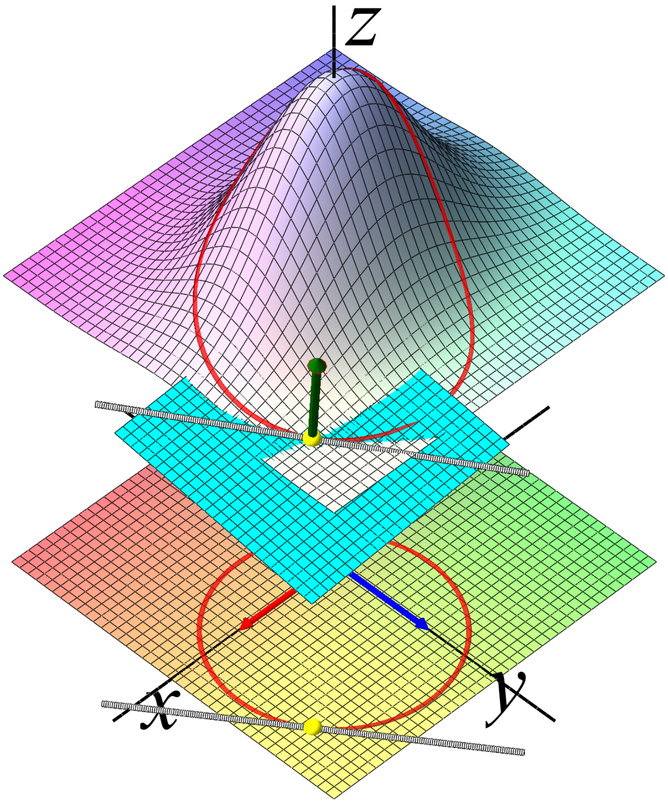

Figure 20.5: The graph of $f(x,y) = 3 + 2\,e^{-x^2-2y^2}$ and the tangent plane through another chosen point on the elevated circle, see Figure 20.2.

|||| Theorem 20.13 The Tangents Lie in the Tangent Plane

Let $f(x,y)$ be a differentiable function of two variables and let $\widetilde{\mathbf{r}}(u)$ denote a elevated curve through the point $\widetilde{\mathbf{r}}(u_0) = (x_0, y_0, f(x_0,y_0))$ on the graph surface $\mathcal{G}(f)$. Then the tangent to the elevated curve is contained in the tangent plane to $\mathcal{G}(f)$ of $f(x,y)$.

In other words: Every point (x,y,z) that lies on the tangent $\widetilde{L}_{u_0}$ also satisfies the equation for the tangent plane:

$$z = P_{1,(x_0,y_0)}(x,y) = f(x_0,y_0) + f'_x(x_0,y_0)\cdot(x-x_0) + f'_y(x_0,y_0)\cdot(y-y_0) \quad . \quad (20\text{-}30)$$

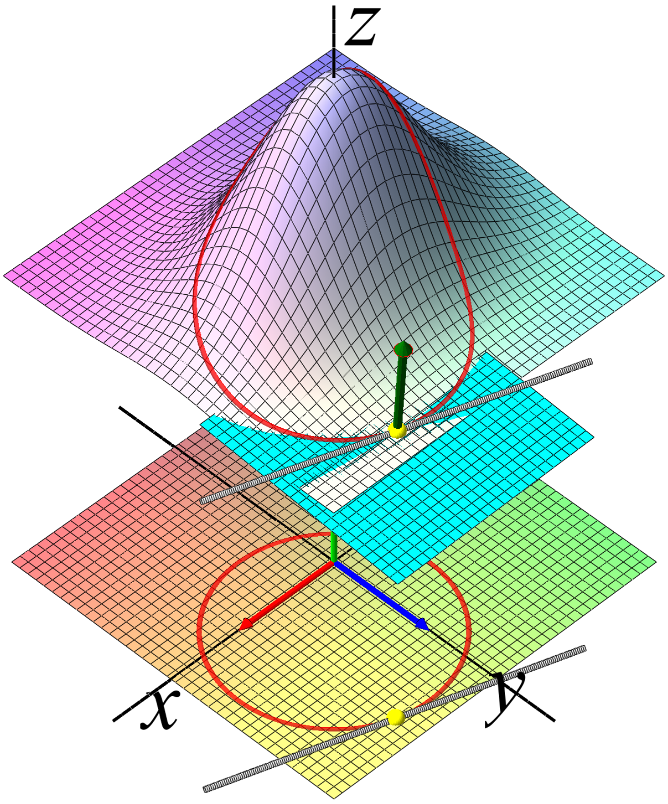

Figure 20.6: The graph of $f(x,y) = 3 + 2\,e^{-x^2-2y^2}$ and the tangent plane through a third chosen point on the elevated circle, see Figure 20.2.

|||| Proof

We only have to realize that the tangent vector $\widetilde{\mathbf{r}}'(u_0)$ is perpendicular to a normal vector to the tangent plane. Such a normal vector is constructed in the next section (see below):

$$\mathbf{N}_{(x_0,y_0)}(f) = (-f'_x(x_0,y_0), -f'_y(x_0,y_0), 1) \quad , \tag{20-31}$$

and since

$$\begin{aligned}\widetilde{\mathbf{r}}'(u_0) &= (p'(u_0), q'(u_0), \boldsymbol{\nabla} f(p(u_0), q(u_0)) \cdot (p'(u_0), q'(u_0))) \\ &= \Big(p'(u_0), q'(u_0), \Big(f'_x(x_0,y_0), f'_y(x_0,y_0)\Big) \cdot (p'(u_0), q'(u_0))\Big) \quad ,\end{aligned} \tag{20-32}$$

we get the orthogonality by computing the scalar product:

$$\begin{aligned}\widetilde{\mathbf{r}}'(u_0) \cdot \mathbf{N}_{(x_0,y_0)}(f) &= -f'_x(x_0,y_0) \cdot p'(u_0) - f'_y(x_0,y_0) \cdot q'(u_0) \\ &\quad + 1 \cdot \Big(f'_x(x_0,y_0), f'_y(x_0,y_0)\Big) \cdot (p'(u_0), q'(u_0)) \\ &= 0 \quad .\end{aligned} \tag{20-33}$$

■

20.3 The Gradient Determines the Normal Vector

A plane in (x, y, z) space with the equation

$$a \cdot x + b \cdot y + c \cdot z + d = 0 \tag{20-34}$$

has a normal vector $\mathbf{N} = (a, b, c)$ that is found directly from the coefficients to x, y, and z in the equation. Now we can write the equation for the tangent plane to the graph surface of $f(x, y)$ at the point $(x_0, y_0, f(x_0, y_0))$ exactly in that form, thus:

$$\begin{aligned}
&z = P_{1,(x_0,y_0)}(x, y) \quad , \\
&z = f(x_0, y_0) + f'_x(x_0, y_0) \cdot (x - x_0) + f'_y(x_0, y_0) \cdot (y - y_0) \quad , \\
&\text{that is equivalent to:} \\
&-f'_x(x_0, y_0) \cdot (x - x_0) - f'_y(x_0, y_0) \cdot (y - y_0) + z - f(x_0, y_0) = 0 \quad , \\
&\text{and thus:} \\
&-f'_x(x_0, y_0) \cdot x - f'_y(x_0, y_0) \cdot y + z + d = 0 \quad ,
\end{aligned} \tag{20-35}$$

where $d = x_0 \cdot f'_x(x_0, y_0) + y_0 \cdot f'_y(x_0, y_0) - f(x_0, y_0)$.

The normal vector to the tangent plane is directly read from the last equation in (20-35):

$$\mathbf{N}_{(x_0,y_0)}(f) = (-f'_x(x_0, y_0), -f'_y(x_0, y_0), 1) \quad . \tag{20-36}$$

A normal vector to the tangent plane to the graph surface of $f(x, y)$ can therefore 'be built' from the same ingredients as the gradient to $f(x, y)$ – the partial derivatives. Note the minus signs and the number 1. And note that while the gradient is a vector in the 2D plane, the normal vector is a vector in 3D space.

An alternative derivation of the coordinates of the normal vector can be found in the following way: If we can find two linearly independent vectors in the tangent plane, then their cross product is a normal vector to the plane.

We always know two *linearly independent vectors in the tangent plane* through $(x_0, y_0, f(x_0, y_0))$, viz. the tangent vectors to the elevated coordinate curves through the point, i.e.

$$\begin{aligned} \widetilde{\mathbf{r}}'_1(x_0) &= (1, 0, f'_x(x_0, y_0)) \quad , \\ \widetilde{\mathbf{r}}'_2(y_0) &= (0, 1, f'_y(x_0, y_0)) \quad . \end{aligned} \tag{20-37}$$

A normal vector to the tangent plane is therefore

$$\begin{aligned} \mathbf{N}_{(x_0,y_0)}(f) &= \widetilde{\mathbf{r}}'_1(x_0) \times \widetilde{\mathbf{r}}'_2(y_0) \\ &= (-f'_x(x_0, y_0)\,,\ -f'_y(x_0, y_0), 1) \end{aligned} \tag{20-38}$$

- that is precisely the same normal vector as found above.

The two tangent vectors $\widetilde{\mathbf{r}}'_1(x_0)$ and $\widetilde{\mathbf{r}}'_2(y_0)$ also give us the ***parametric representation of the tangent plane***:

|||| Theorem 20.14 Parametric Representation of the Tangent Plane

The tangent plane through the point $(x_0, y_0, f(x_0, y_0))$ on the graph surface $\mathcal{G}(f)$ of the function $f(x, y)$ has the parametric representation

$$\begin{aligned} \mathcal{T}_{(x_0,y_0)}(f) \ : \ \mathbf{T}(t_1, t_2) &= (x_0, y_0, f(x_0, y_0)) + t_1 \cdot \widetilde{\mathbf{r}}'_1(x_0) + t_2 \cdot \widetilde{\mathbf{r}}'_2(y_0) \\ &= (x_0, y_0, f(x_0, y_0)) + t_1 \cdot (1, 0, f'_x(x_0, y_0)) + t_2 \cdot (0, 1, f'_y(x_0, y_0)) \ , \end{aligned} \tag{20-39}$$

where $t_1 \in \mathbb{R}$ and $t_2 \in \mathbb{R}$.

|||| Example 20.15 Elevated Coordinate Curves

The unit normal vector in the direction $\mathbf{N}$ to the tangent planes shown in Figures 20.4, 20.5, 20.6 is also shown in the figures. In addition the coordinate curves in the (x, y) plane and the elevated coordinate curves on the graph surfaces both of the function $f(x, y)$ and of the respective approximating first-degree polynomials (the tangent planes) are shown.

|||| Example 20.16 Functions of One Variable

Every function of one variable can be considered as a function of two variables. We may expect that the level curves, the gradient vector fields, the tangent planes, etc. are relatively simple for such functions. We look at some examples that show this:

1. The function $f(x,y) = x^2$ has the following gradient field, tangent planes, and normal vectors:

$$\begin{aligned} & \nabla f(x,y) = (2x, 0) \quad , \\ \mathcal{T}_{(x_0,y_0)}(f) \quad : \quad & \mathbf{T}(t_1,t_2) = (x_0 + t_1\,,\; y_0 + t_2\,,\; x_0^2 + 2x_0 \cdot t_1) \quad , \\ & \mathbf{N}_{(x_0,y_0)}(f) = (-2x_0, 0, 1) \quad . \end{aligned} \tag{20-40}$$

   See Figure 20.7 where the contours, the gradient vector field, and the graph surface are shown for the function.

2. The function $f(x,y) = x^3$ has

$$\begin{aligned} & \nabla f(x,y) = (3x^2, 0) \quad , \\ \mathcal{T}_{(x_0,y_0)}(f) \quad : \quad & \mathbf{T}(t_1,t_2) = (x_0 + t_1\,,\; y_0 + t_2\,,\; x_0^3 + 3x_0^2 \cdot t_1) \quad , \\ & \mathbf{N}_{(x_0,y_0)}(f) = (-3x_0^2, 0, 1) \quad . \end{aligned} \tag{20-41}$$

   See Figure 20.8 where only the contours and the gradient vector field are shown. Compare contours and the gradient vector field of $f(x,y) = x^2$ in Figure 20.7.

3. The function $f(x,y) = \cos(3x)$ has

$$\begin{aligned} & \nabla f(x,y) = (-3\sin(3x), 0) \quad , \\ \mathcal{T}_{(x_0,y_0)}(f) \quad : \quad & \mathbf{T}(t_1,t_2) = (x_0 + t_1\,,\; y_0 - t_1 \cdot 3\sin(x_0) + t_2\,,\; \cos(3x_0)) \quad , \\ & \mathbf{N}_{(x_0,y_0)}(f) = (3\sin(3x_0), 0, 1) \quad . \end{aligned} \tag{20-42}$$

   See Figure 20.9. Compare with Figure 20.10. Inspection of the graph and of the contours of the function in 20.10 points to the idea that by rotating the coordinate system and by thus changing the coordinates in the plane – we can also describe the function as a function of one variable. This is in fact the case since the function shown is $f(x,y) = \cos(3x + 3y)$.

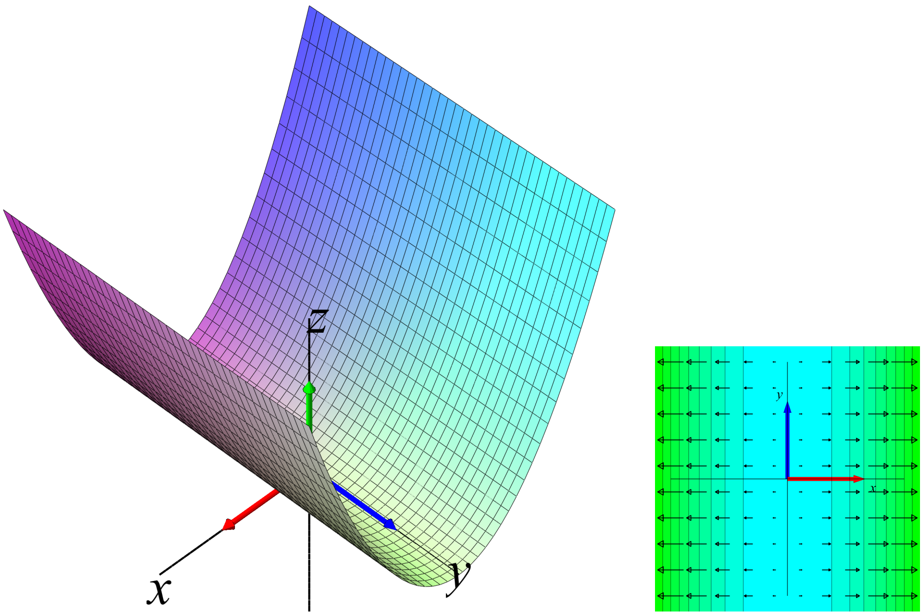

Figure 20.7: The graph of $f(x,y) = x^2$ together with contours and gradient vector fields in the (x,y) plane.

|||| Exercise 20.17

Determine the unit normal vector to the tangent planes through the point $(x_0, y_0, f(x_0, y_0))$ on the graph surface of each of the following functions for every point (x_0, y_0) in the (x,y) plane.

$$\begin{aligned} f(x,y) &= 3x + y \\ f(x,y) &= x^2 + y^2 \\ f(x,y) &= y \cdot \mathrm{e}^x \quad . \end{aligned} \tag{20-43}$$

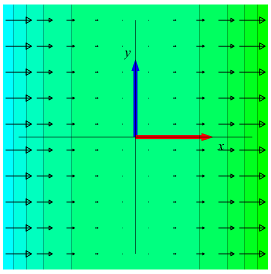

Figure 20.8: Contours and the gradient vector field of $f(x,y) = x^3$.

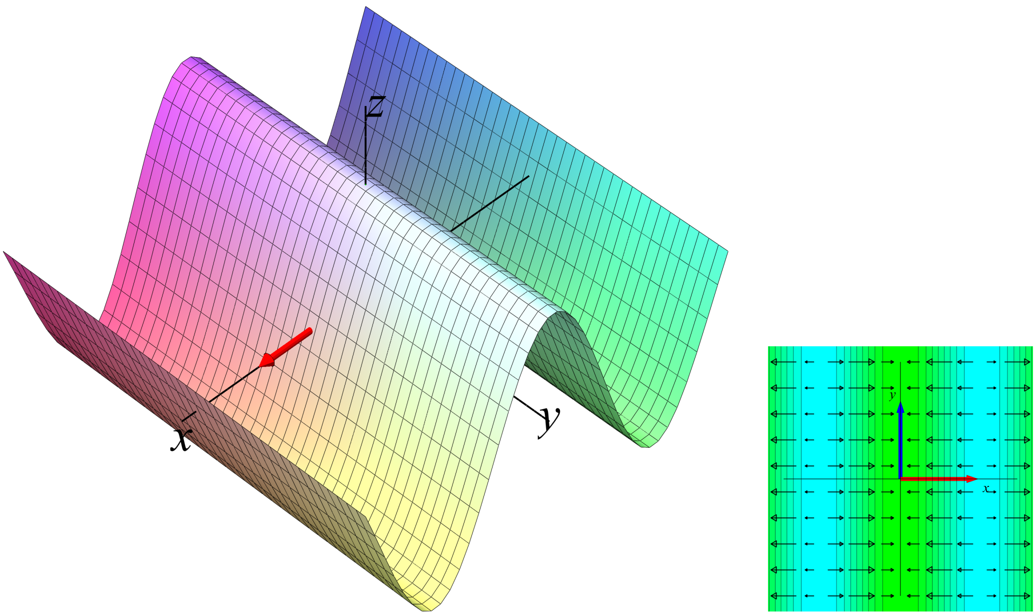

Figure 20.9: The graph, contours and the gradient vector field of $f(x,y) = \cos(3x)$.

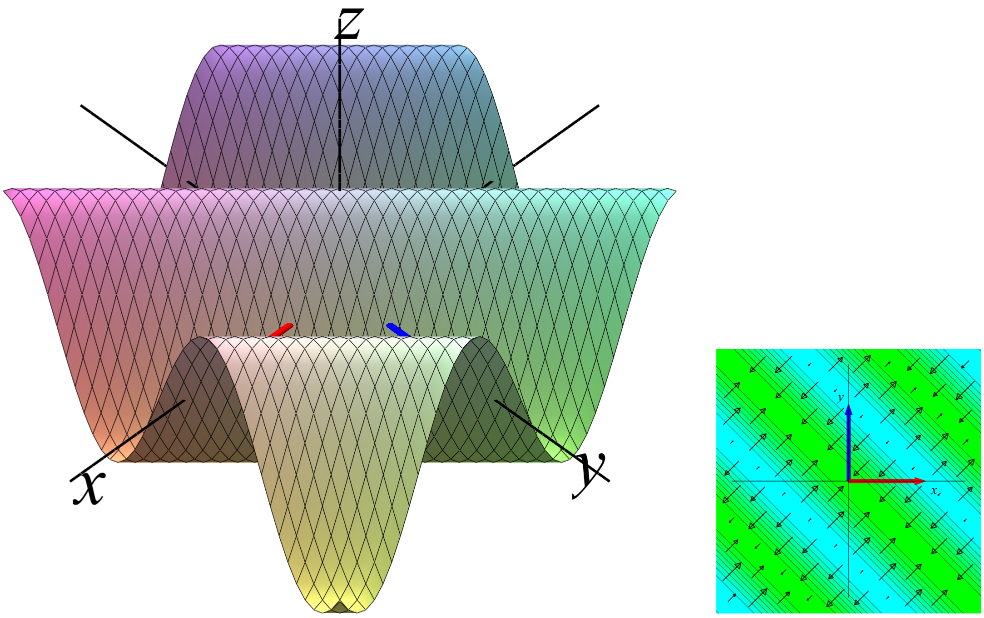

Figure 20.10: The graph, contours, and gradient vector fields of $f(x,y) = \cos(3x+3y)$.

20.4 Summary

The partial derivative of a function of two variables has different geometric appearances that are particularly useful in the description and analysis of the functions. The gradient vector field – that has the partial derivatives as coordinate functions – contains (almost) all information about the function. In this eNote we have worked with the following:

- The gradient at every point points in the direction where the function locally increases the most and the (opposite) direction in which the function decreases the most.
- The length of the gradient of $f(x,y)$ at the point $(x_0, y_0))$ is the value of the largest directional derivative of $f(x,y)$ at the point, and the corresponding negative value is the value of the smallest directional derivative at the point.
- The gradient vector field of $f(x,y)$ is everywhere orthogonal to the contours of $f(x,y)$.
- The elevated curves on the graph surface of a function have tangents that are entirely contained within the tangent plane to the graph surface.
- The tangent plane to the graph surface of the function $f(x,y)$ at a given point $(x_0, y_0, f(x_0, y_0))$ has a normal vector that is 'built' from the partial derivatives:

$$\mathbf{N}_{(x_0,y_0)}(f) = (-f'_x(x_0,y_0)\,,\ -f'_y(x_0,y_0), 1) \quad .$$

- The tangent plane to the graph surface of the function $f(x,y)$ at a given point $(x_0, y_0, f(x_0, y_0))$ is represented both by the equation

$$z = P_{1,(x_0,y_0)}(x,y) = f(x_0,y_0) + f'_x(x_0,y_0)\cdot(x-x_0) + f'_y(x_0,y_0)\cdot(y-y_0)$$

  and by the parametric representation

$$\begin{aligned}\mathcal{T}_{(x_0,y_0)}(f) \ :\ \mathbf{T}(t_1,t_2) &= (x_0, y_0, f(x_0,y_0)) + t_1\cdot\widetilde{\mathbf{r}}'_1(x_0) + t_2\cdot\widetilde{\mathbf{r}}'_2(y_0) \\ &= (x_0, y_0, f(x_0,y_0)) + t_1\cdot(1, 0, f'_x(x_0,y_0)) + t_2\cdot(0, 1, f'_y(x_0,y_0))\ ,\end{aligned}$$

  where $t_1 \in \mathbb{R}$ and $t_2 \in \mathbb{R}$.