3 Applications of partial derivatives

3.1 Gradient

The function of two variables $z = f(x, y)$ associates to any point $P(x, y)$ in the domain of that function D one value of the dependent variable z or a scalar. To any point in the domain of the function there is related a scalar. Hence, the function of two variables creates a *scalar field* in the plane.

The function of two variables $w = f(x, y, z)$ associates to any point $P(x, y, z)$ in its domain V a scalar, i.e creates a scalar field in the domain V. Examples used in physics include the temperature distribution throughout space, the pressure distribution in a fluid or in a gas. Scalar fields are contrasted with other physical quantities such as vector fields, which associate a vector to every point of a region.

Definition 1.

$$\operatorname{grad} z = \left(\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}\right) \tag{3.1}$$

is called the *gradient* of the scalar field $z = f(x, y)$.

Definition 2. The vector

$$\operatorname{grad} w = \left(\frac{\partial w}{\partial x}, \frac{\partial w}{\partial y}, \frac{\partial w}{\partial z}\right) \tag{3.2}$$

is called the gradient of the scalar field $w = f(x, y, z)$.

In the first case there is defined a vector field in the plane and in the second case a vector field in the space. These are called the *gradient field*.

If $\overrightarrow{s^{\circ}} = (\cos\alpha, \cos\beta)$ denotes the unit vector in the direction of the vector $\overrightarrow{s}$, the formula (2.20) can be written as the scalar product of the gradient and the unit vector $\overrightarrow{s^{\circ}}$

$$\frac{\partial z}{\partial \overrightarrow{s}} = \operatorname{grad} z \cdot \overrightarrow{s^{\circ}}$$

Since $\overrightarrow{s^{\circ}} = \dfrac{\overrightarrow{s}}{\Delta s}$, then

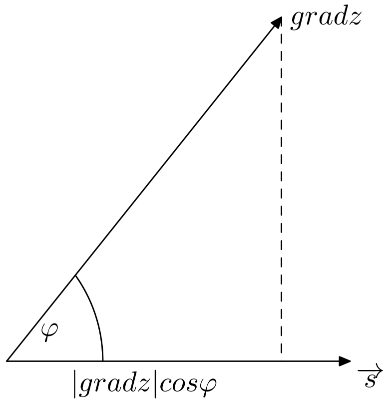
$$\frac{\partial z}{\partial \overrightarrow{s}} = \operatorname{grad} z \cdot \frac{\overrightarrow{s}}{\Delta s} = |gradz| \frac{\operatorname{grad} z \cdot \overrightarrow{s}}{|\operatorname{grad} z| \Delta s}$$

where $|\operatorname{grad} z|$ is the length of the gradient vector. Denoting by φ the angle between the gradient and the vector $\overrightarrow{s}$ we obtain

$$\cos\varphi = \frac{\operatorname{grad} z \cdot \overrightarrow{s}}{|\operatorname{grad} z| \Delta s}$$

and

$$\frac{\partial z}{\partial \overrightarrow{s}} = |\operatorname{grad} z| \cos\varphi. \tag{3.3}$$



Now we formulate this result as a theorem.

Theorem 1. The directional derivative of the function $z = f(x, y)$ equals to the projection of the gradient vector onto the direction of vector $\overrightarrow{s}$.

Two important conclusions of this theorem.

Conclusion 1. The directional derivative in direction perpendicular to the gradient equals to zero.

This conclusion is obvious because in our case $\varphi = \frac{\pi}{2}$ and $\frac{\partial z}{\partial \overrightarrow{s}} = 0$.

Conclusion 2. The directional derivative has the greatest value in the direction of the gradient and equals to the length of the gradient.

It's enough to recall that the cosine function obtains its greatest value 1 if $\varphi = 0$. Thus, the direction of fastest change for a function is given by the gradient vector at that point.

Example 1. Find the greatest rate of growth of the function $z = x^2 + y^2$ at the point $P(1; 1)$.

The directional derivative gives the instantaneous rate of change at the given point. The greatest instantaneous rate of change equals to the length of the gradient. We find the gradient vector at the point P

$$\operatorname{grad} z = (2x, 2y)\Big|_P = (2; 2)$$

and its length $|\operatorname{grad} z| = 2\sqrt{2}$.

This result is the same as the result in Example 1 of the previous subsection, where we have found the directional derivative in direction of the vector $\overrightarrow{s_1}$. This is natural because the vector $\overrightarrow{s_1} = (1; 1)$ and the gradient have the same directions.

Theorem 2. The gradient is perpendicular to the tangent of level curve.

Proof. The projection of the level curve of the surface $z = f(x, y)$ onto xy-plane is $f(x, y) = c$. This is an implicit function of one variable and the graph is a curve in xy-plane. The slope of the tangent line of this curve is $\dfrac{dy}{dx} = -\dfrac{f'_x}{f'_y}$. Hence, the direction vector of the tangent line is

$$\overrightarrow{s} = \left(1; -\frac{f'_x}{f'_y}\right) = \frac{1}{f'_y}(f'_y, -f'_x)$$

The scalar product of the gradient vector and the direction vector of the tangent line

$$\operatorname{grad} z \cdot \overrightarrow{s} = f'_x f'_y - f'_y f'_x = 0$$

which means that these two vectors are perpendicular.

Now the Conclusion 1 gives us.

Conclusion 3. The derivative in the direction of the tangent line of the level curve equals to zero.

In Example 1 of the previous subsection the vector $\overrightarrow{s_2}$ has the same direction as the tangent line of the level curve. Thus, by Conclusion 3 it is natural that the derivative in the direction of this vector equals to zero.

Definition 3. A vector field $\overrightarrow{F} = (X(x, y), Y(x, y))$ is called a *conservative vector field* if there exists a scalar field $z = f(x, y)$ such that $\overrightarrow{F} = \operatorname{grad} z$. If $\overrightarrow{F}$ is a conservative vector field then the function $f(x, y)$ is called a *potential function* for $\overrightarrow{F}$.

All this definition is saying is that a vector field is conservative if it is also a gradient vector field for some scalar field.

Example 2. The vector field $\overrightarrow{F} = (2xy; x^2)$ is conservative because there exists the scalar field $z = x^2 y$ such that $\operatorname{grad} z = \overrightarrow{F}$ and $x^2 y$ is the potential function for $\overrightarrow{F}$.

3.2 Divergence and curl

The gradient vector field is just one example of vector fields. More generally, a vector field $\overrightarrow{F} = (X(x, y, z); Y(x, y, z); Z(x, y, z))$ is an assignment of a vector to each point (x, y, z) in a subset of space. Vector fields are often used to model, for example, the strength and direction of some force, such as the magnetic or gravitational force, as it changes from point to point or the speed and direction of a moving fluid throughout space.

Definition 1. The scalar

$$\operatorname{div} \overrightarrow{F} = \frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} + \frac{\partial Z}{\partial z} \tag{3.4}$$

is called the *divergence* of the vector field $\overrightarrow{F}$ at the point $P(x, y, z)$.

Definition 2. The vector

$$\operatorname{curl} \overrightarrow{F} = \left(\frac{\partial Z}{\partial y} - \frac{\partial Y}{\partial z}; \frac{\partial X}{\partial z} - \frac{\partial Z}{\partial x}; \frac{\partial Y}{\partial x} - \frac{\partial X}{\partial y} \right) \tag{3.5}$$

is called the *curl* (or rotor) of the vector field $\overrightarrow{F}$ at the point $P(x, y, z)$.

Example 1. Find the divergence and curl of the vector field $\overrightarrow{F} = \left(xyz; x^2 + z^2; \frac{xy}{z} \right)$.

In this example $X = xyz$, $Y = x^2 + z^2$, $Z = \frac{xy}{z}$, thus, $\frac{\partial X}{\partial x} = yz$, $\frac{\partial Y}{\partial y} = 0$ and $\frac{\partial Z}{\partial z} = -\frac{xy}{z^2}$. Hence, the divergence

$$\operatorname{div} \overrightarrow{F} = yz - \frac{xy}{z^2}$$

The components of the curl vector

$$\frac{\partial Z}{\partial y} - \frac{\partial Y}{\partial z} = \frac{x}{z} - 2z$$

$$\frac{\partial X}{\partial z} - \frac{\partial Z}{\partial x} = xy - \frac{y}{z}$$

and

$$\frac{\partial Y}{\partial x} - \frac{\partial X}{\partial y} = 2x - xz$$

Consequently,

$$\operatorname{curl} \overrightarrow{F} = \left(\frac{x}{z} - 2z; xy - \frac{y}{z}; 2x - xz \right)$$

If the vector field represents the velocity of a moving flow in space, then the divergence of a vector field $\overrightarrow{F}$ at point $P(x, y, z)$ represents a measure of the rate at which the flow diverges (spreads away) from P. That is, $\operatorname{div} \overrightarrow{F}\big|_P$ is the limit of the flow per unit volume out of the infinitesimal sphere centered at P. The curl represents the rotation of a flow, i.e. $\operatorname{curl} \overrightarrow{F}\big|_P$ measures the extent to which the vector field $\overrightarrow{F}$ rotates around P.

Suppose that $\overrightarrow{F}$ is the velocity field in a flowing fluid. Then the curl $\overrightarrow{F}$ represents the tendency of particles at the point (x, y, z) to rotate about the axis that points in direction of curl $\overrightarrow{F}$. The length of curl vector represents the velocity of that rotation.

If $\operatorname{curl} \overrightarrow{F} = \overrightarrow{0}$, the vector field $\overrightarrow{F}$ is called *irrotational*.

In field theory there is used a formal vector.

Definition 3. The vector

$$\nabla = \left(\frac{\partial}{\partial x}; \frac{\partial}{\partial y}; \frac{\partial}{\partial z}\right)$$

is called *Hamilton nabla vector* or *Hamilton nabla operator*.

The coordinates of this vector are not numbers but some operators. The first coordinate means that we find the partial derivative with respect to x for some function etc.

If we treat this vector as an usual vector, we can write for the scalar field $w = f(x, y, z)$

$$\nabla w = \left(\frac{\partial}{\partial x}; \frac{\partial}{\partial y}; \frac{\partial}{\partial z}\right) w = \left(\frac{\partial w}{\partial x}; \frac{\partial w}{\partial y}; \frac{\partial w}{\partial z}\right) = \operatorname{grad} w$$

Here we have the formal scalar multiplication of ∇ and w. The order of factors is important. The quantities on which ∇ acts must appear to the right of ∇.

The scalar product of ∇ and the vector field $\overrightarrow{F} = (X(x, y, z); Y(x, y, z); Z(x, y, z))$ is

$$\nabla \cdot \overrightarrow{F} = \left(\frac{\partial}{\partial x}; \frac{\partial}{\partial y}; \frac{\partial}{\partial z}\right) \cdot (X; Y; Z) = \frac{\partial X}{\partial x} + \frac{\partial Y}{\partial y} + \frac{\partial Z}{\partial z} = \operatorname{div} \overrightarrow{F}$$

The vector product of ∇ and the vector field $\overrightarrow{F} = (X(x, y, z); Y(x, y, z); Z(x, y, z))$ is

$$\nabla \times \overrightarrow{F} = \left(\frac{\partial}{\partial x}; \frac{\partial}{\partial y}; \frac{\partial}{\partial z}\right) \times (X; Y; Z) = \left(\frac{\partial Z}{\partial y} - \frac{\partial Y}{\partial z}; \frac{\partial X}{\partial z} - \frac{\partial Z}{\partial x}; \frac{\partial Y}{\partial x} - \frac{\partial X}{\partial y}\right) = \operatorname{curl} \overrightarrow{F}$$

Hence, using the nabla operator, we can write

$$\operatorname{grad} w = \nabla w$$

$$\operatorname{div} \overrightarrow{F} = \nabla \cdot \overrightarrow{F}$$

$$\operatorname{curl} \overrightarrow{F} = \nabla \times \overrightarrow{F}$$

Definition 4. The scalar product of nabla vector by itself $\nabla^2 = \nabla \cdot \nabla$ is called Laplacian operator and denoted

$$\triangle = \nabla^2$$

The scalar product of nabla vector by itself is not a real quantity

$$\triangle = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

but applying this operator to some function, we obtain at every point of the space a scalar.

Example 2. Find the Laplacian operator for the function $w = e^x \sin(yz)$.

First we find the first-order partial derivatives

$$\frac{\partial w}{\partial x} = e^x \sin(yz)$$

$$\frac{\partial w}{\partial y} = ze^x \cos(yz)$$

$$\frac{\partial w}{\partial z} = ye^x \cos(yz)$$

and next

$$\begin{aligned}\triangle w &= \frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} + \frac{\partial^2 w}{\partial z^2} \\ &= e^x \sin(yz) - z^2 e^x \sin(yz) - y^2 e^x \sin(yz) \\ &= e^x \sin(yz)(1 - z^2 - y^2) = w(1 - z^2 - y^2)\end{aligned}$$

Finally we prove some equalities that hold for the scalar field $w = f(x, y, z)$ and vector field $\overrightarrow{F} = (X; Y; Z)$.

Corollary 1. $\operatorname{div} \operatorname{grad} w = \triangle w$

Proof We write

$$\operatorname{div} \operatorname{grad} w = \nabla \cdot \operatorname{grad} w = \left(\frac{\partial}{\partial x}; \frac{\partial}{\partial y}; \frac{\partial}{\partial z}\right) \cdot \left(\frac{\partial w}{\partial x}; \frac{\partial w}{\partial y}; \frac{\partial w}{\partial z}\right) = \frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} + \frac{\partial^2 w}{\partial z^2}$$

3.3 Local extrema of function of two variables

The theory of maxima and minima for the functions of two variables is similar to the theory for one variable.

Definition 1. It is said that the function of two variables $f(x, y)$ has a local maximum at the point $P_1(x_1, y_1)$, if there exists a neighborhood of this point $U_\varepsilon(x_1, y_1)$ such that for any $P(x, y) \in U_\varepsilon(x_1, y_1)$

$$f(x, y) < f(x_1, y_1)$$

Definition 2. It is said that the function of two variables $f(x, y)$ has a local minimum at the point $P_2(x_2, y_2)$, if there exists a neighborhood of this point $U_\varepsilon(x_2, y_2)$ such that for any $P(x, y) \in U_\varepsilon(x_2, y_2)$

$$f(x, y) > f(x_2, y_2)$$

Local extremum is either a local maximum or a local minimum.

Example 1. By Definition 2 the function $z = x^2 + y^2$ has the local minimum at the point $P_0(0;0)$ because $f(0;0) = 0$ and for any point $P(x, y)$ different of P_0 there holds $f(x, y) = x^2 + y^2 > 0$.

Example 2. The function $z = x^2 - y^2$ has no local extremum at the point $P_0(0;0)$. We have $f(0;0) = 0$ and any neighborhood $U_\varepsilon(0;0)$ contains the points of x-axis and y-axis. At the points on x-axis $y = 0$ and $z = x^2 > 0$, at the points of y-axis $x = 0$ and $z = -y^2 < 0$.

If the function of two variables has local extremum at the point $P_0(x_0, y_0)$ then the intersection curve of surface (the graph of the function of two variables) and the plain $y = y_0$ has local extremum at x_0. Hence, the function of one variable $z = f(x, y_0)$ has local extremum at x_0. It follows that at the point P_0 either $\dfrac{\partial z}{\partial x} = 0$ or does not exist.

As well, the intersection curve of surface and the plain $x = x_0$ has local extremum at y_0. The function of one variable $z = f(x_0, y)$ has local extremum at y_0. Then at the point P_0 either $\dfrac{\partial z}{\partial y} = 0$ or does not exist.

Definition 3. The points, where $\dfrac{\partial z}{\partial x} = 0$ or does not exist and $\dfrac{\partial z}{\partial y} = 0$ or does not exist, are called the *critical points* of the function of two variables.

Now we can formulate the theorem.

Theorem 1. (Necessary condition for existence of local extremum). If the function $z = f(x, y)$ has local extremum at the point P_0, then P_0 is the critical point of this function.

This theorem says that the function of two variables has a local extremum only at the critical point of this function. But the condition given in this theorem is not sufficient for the function to have a local extremum. For instance the point $O(0;0)$ is the critical point of the function $z = x^2 - y^2$ because the partial derivatives $\dfrac{\partial z}{\partial x} = 2x$ and $\dfrac{\partial z}{\partial y} = 2y$ both equal to zero at this point, but as we know by Example 2, this function has no local maximum and local minimum at $O(0;0)$.

Because of this theorem we know that if we have all the critical points of a function then we also have every possible local extremum for the function. The fact tells us that all local extrema must be at the critical points so we know that if the function does have local extrema then they must be in the set of all the critical points. However, it will be completely possible that at least at one of the critical points the function hasn't a local extremum.

So the question is how to determine whether the function of two variables has a local extremum at the critical point or not and if it has, is at that point a local maximum or a local minimum.

In the following we consider only the critical points where both partial derivatives equal to zero, i.e. the system of equations

$$\begin{cases} \dfrac{\partial z}{\partial x} = 0 \\ \dfrac{\partial z}{\partial y} = 0 \end{cases} \tag{3.6}$$

The solutions of this system of equations are called the *stationary points* of the function $z = f(x, y)$. Every stationary point is also a critical point of the function of two variables but not vice versa. There exist the critical points that are not the stationary points. For instance, for the function $z = \sqrt{x^2 + y^2}$ the partial derivatives

$$\frac{\partial z}{\partial x} = \frac{x}{\sqrt{x^2 + y^2}}$$

and

$$\frac{\partial z}{\partial y} = \frac{y}{\sqrt{x^2 + y^2}}$$

are never simultaneously zero, however they both don't exist at $O(0; 0)$. Therefore, $O(0; 0)$ is a critical point and a possible extremum. The graph of $z = \sqrt{x^2 + y^2}$ is a cone opening upwards with vertex at the origin. Therefore, at $O(0; 0)$ this function has a local minimum at $O(0; 0)$.

We find the sufficient conditions for existence of the local extremum at the stationary points. Let P_0 be a stationary point of the function $z = f(x, y)$. Evaluate the second order partial derivatives at P_0 and denote

$$A = \left.\frac{\partial^2 z}{\partial x^2}\right|_{P_0} \quad B = \left.\frac{\partial^2 z}{\partial x \partial y}\right|_{P_0} \quad \text{and} \quad C = \left.\frac{\partial^2 z}{\partial y^2}\right|_{P_0}$$

Theorem 2 (sufficient conditions for existence of a local extremum). Let P_0 be a stationary point of the function $z = f(x, y)$.

1. If $AC - B^2 > 0$ and $A < 0$ then the function $z = f(x, y)$ has a local maximum at P_0.

2. If $AC - B^2 > 0$ and $A > 0$ then the function $z = f(x, y)$ has a local minimum at P_0.

3. If $AC - B^2 < 0$ then the function $z = f(x, y)$ has no local extremum at P_0.

Definition 4. If $AC - B^2 < 0$ then the stationary point P_0 is called the *saddle point* of the function $z = f(x, y)$.

We obtain the stationary point $P_0(0; 0)$ of the function $z = x^2 + y^2$ as the solution of the system of equations (3.6)

$$\begin{cases} 2x = 0 \\ 2y = 0 \end{cases}$$

We find

$$A = \frac{\partial^2 z}{\partial x^2} = 2$$

$$B = \frac{\partial^2 z}{\partial x \partial y} = 0$$

and

$$C = \frac{\partial^2 z}{\partial y^2} = 2$$

Hence, $AC - B^2 = 4 > 0$ and $A > 0$. Consequently, by Theorem 2 the function $z = x^2 + y^2$ has at stationary point $P_0(0; 0)$ a local minimum.

We obtain the stationary point $P_0(0; 0)$ of the function $z = x^2 - y^2$ as the solution of the system of equations (3.6)

$$\begin{cases} 2x = 0 \\ -2y = 0 \end{cases}$$

We find

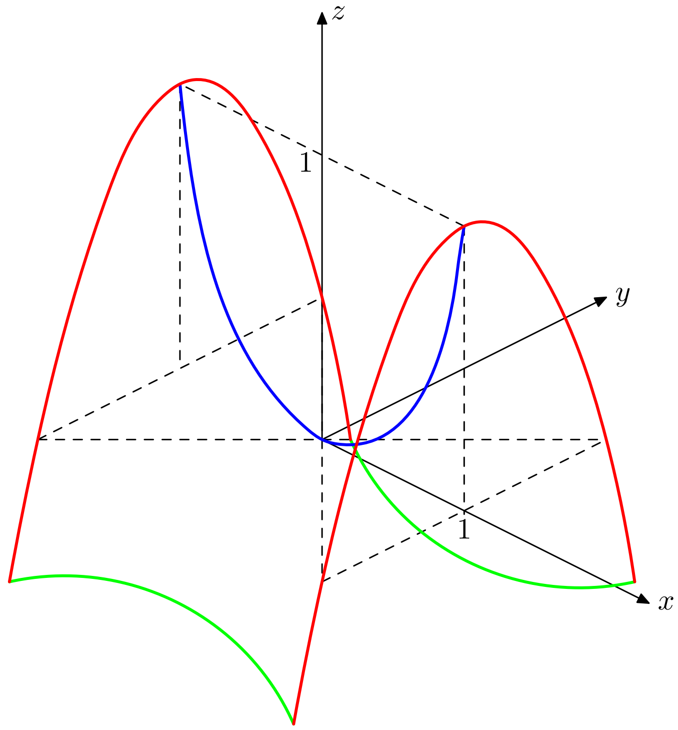
$$A = \frac{\partial^2 z}{\partial x^2} = 2$$

$$B = \frac{\partial^2 z}{\partial x \partial y} = 0$$

and

$$C = \frac{\partial^2 z}{\partial y^2} = -2$$

Thus, $AC - B^2 = -4 < 0$. Consequently, by Theorem 2 the function $z = x^2 - y^2$ has't a local extremum at the stationary point $P_0(0; 0)$. In other words: the point $P_0(0; 0)$ is the saddle point of the function $z = x^2 - y^2$.



Joonis 3.1: Saddle surface

Example 3. Find the local extrema of the function $f(x,y) = 4 + x^3 + y^3 - 3xy$.

The first order partial derivatives are

$$\frac{\partial f}{\partial x} = 3x^2 - 3y \quad \text{and} \quad \frac{\partial f}{\partial y} = 3y^2 - 3x$$

To find the stationary points we solve the system of equations (3.6)

$$\begin{cases} 3x^2 - 3y = 0 \\ 3y^2 - 3x = 0 \end{cases}$$

or

$$\begin{cases} x^2 - y = 0 \\ y^2 - x = 0 \end{cases}$$

The first equation gives $y = x^2$. Substituting this into second equation gives $x^4 - x = 0$ or $x(x^3 - 1) = 0$, whose solutions are $x_1 = 0$ and $x_2 = 1$. Since $y = x^2$, we have two stationary points $P_1(0;0)$ and $P_2(1;1)$. Next we find the second order partial derivatives

$$\frac{\partial^2 f}{\partial x^2} = 6x \quad \frac{\partial^2 f}{\partial x \partial y} = -3 \quad \text{and} \quad \frac{\partial^2 f}{\partial y^2} = 6y$$

Since at the first stationary point $P_1(0;0)$ the values $A = 0$, $B = -3$ and $C = 0$ and

$$AC - B^2 = 0 \cdot 0 - (-3)^2 = -9$$

the point $P_1(0;0)$ is the saddle point of the given function.

At the second stationary point $P_2(1;1)$ the values $A = 6$, $B = -3$ and $C = 6$ and

$$AC - B^2 = 6 \cdot 6 - (-3)^2 = 27 > 0$$

As well $A = 6 > 0$ and by Theorem 2 the given function has a local minimum at the point $P_2(1;1)$ and this local minimum equals to

$$z_{min} = 4 + 1^3 + 1^3 - 3 \cdot 1 \cdot 1 = 3$$

Remark If in Theorem 2 $AC - B^2 = 0$ then anything is possible. More advanced methods are required to classify the stationary point properly.