Derivatives of functions of several variables

There are different types of derivatives for functions of several variables in R^n. One of them is defined as follows.

Let $f(x) = f\Big(x_1, x_2, \ldots, x_n\Big)$ is a function depending on several variables. In it we can replace all its arguments with constants, except for one with number k.

In this case, $f(x)$ can be considered as a function of a single variable x_k. For such function there is the concept of derivative, i.e. new particular function with values

$$\lim_{t\to 0} \frac{f\Big(x_{(0)1}, \ldots, x_{(0)k}+t, \ldots, x_{(0)n}\Big) - f\Big(x_{(0)1}, \ldots, x_{(0)k}, \ldots, x_{(0)n}\Big)}{t}. \tag{3.01}$$

Here point $\Big\{x_{(0)1}, \ldots, x_{(0)k}, \ldots, x_{(0)n}\Big\}$ belongs to the domain of definition of function $f(x)$.

It is known that the value of the limit (if it exists) is unique. Therefore, formula (3.01) can be used as a *definition* of a new function of n variables.

This new function is called *partial derivative of the function* $f(x)$ *with respect to variable* x_k This name is justified by the fact: subject to $n = 1$ formula (3.01) is the definition of the derivative for the function of *single* variable

$$\frac{df}{dx_k} = f'(x_k)\,. \tag{3.02}$$

Let us note again: partial derivative of the function $f(x)$ with respect to variable x_k is a new function of n variables. Therefore, to denote partial derivative it is used not (3.02), but

$$\frac{\partial f}{\partial x_k} = f'_{x_k}\Big(\,x_1,\,x_2,\,\ldots\,,\,x_n\,\Big)\,.$$

It is easy to verify that for the partial derivative are valid all the rules and theorems, that are true for derivative of a function depending on only one variable.

Higher order derivatives for a function of several variables

Since the partial derivative is itself a function of several variables, then it is natural to also assume the existence of partial derivatives for it.

In this case, the variable with respect to which the new derivative is taken may not be x_k, but another, say x_j.

Such derivatives are called mixed partial derivatives and it is customary to use the notation of the form $f''_{x_k x_j}$ or

$$\frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_k}\right) = \frac{\partial^2 f}{\partial x_k \partial x_j}\Big(x_1,\, x_2,\, \ldots\,,\, x_n \Big)\,.$$

If $k = j$, then the notation is used $\dfrac{\partial^2 f}{\partial x_k^2}$.

Turns out to be fair

`Theorem 3.1` **If partial derivatives $\dfrac{\partial^2 f}{\partial x_k \partial x_j}$ and $\dfrac{\partial^2 f}{\partial x_j \partial x_k}$ are continuous at some point, then they are *equal* at the point.**

Since the main specific properties of functions of several variables appear already at $n = 2$, then to illustrate them we will limit ourselves only to this case.

Problem 3.1 *Let* $f(x, y) = \sqrt{x^2 + y^2}$. *We need to find all second order partial derivatives.*

Solution. The domain of definition of this function in R^2 is the entire coordinate plane Oxy.
To calculate derivatives, we will use the rules for differentiating a function of one variable.
Consistently assuming y and x to be constants, we get

$$\frac{\partial f}{\partial x} = \frac{x}{\sqrt{x^2 + y^2}} \qquad \text{and} \qquad \frac{\partial f}{\partial y} = \frac{y}{\sqrt{x^2 + y^2}}.$$

For the second derivatives we find

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2} &= \frac{y^2}{\sqrt{\left(x^2+y^2\right)^3}}\,, \\ \frac{\partial^2 f}{\partial x \partial y} &= \frac{xy}{\sqrt{\left(x^2+y^2\right)^3}} = \frac{\partial^2 f}{\partial y \partial x}\,, \\ \frac{\partial^2 f}{\partial x^2} &= \frac{y^2}{\sqrt{\left(x^2+y^2\right)^3}}\,. \end{aligned} \tag{3.03}$$

The resulting formulas are not valid throughout the entire domain of definition functions $f(x, y)$, since the application of *rules* of differentiation assumes that *all* derivatives, included in the record of these rules exist.
In our example this is true only for $x^2 + y^2 \neq 0$.

Here we pay attention to the following important detail: from the fact that the functions in formulas (3.03) are not defined at the origin, generally speaking, it does not follow that that the function $f(x,y)$ at this point has no partial derivatives.

The absence of partial derivatives at the point $(0;0)$ follows from formula (3.01). Indeed, for $x_{(0)} = 0$ and $y_{(0)} = 0$ we have

$$\lim_{t\to 0} \frac{\sqrt{(x_{(0)}+t)^2+(y_{(0)})^2} - \sqrt{(x_{(0)})^2+(y_{(0)})^2}}{t} = \lim_{t\to 0} \frac{|\,t\,|}{t},$$

`Solution found.`

but this limit does not exist.

Gradient and directional derivative of functions of several variables

Due to partial derivatives it is possible to more accurately describe and analyze the properties of functions of several variables.

One of the tools for such analysis is the *gradient* of function of several variables.

Let a function of several variables $f(x) = f\Big(x_1, x_2, \dots, x_n \Big)$ has at some point $x = \left\| x_1, x_2, \dots, x_n \right\|$ of its domain all partial derivatives first order

$$\frac{\partial f}{\partial x_k}\Big(x_1, x_2, \dots, x_n \Big) \qquad \forall k = \overline{1, n}\,.$$

Definition 3.1	*The gradient* at a point x of a function of several variables $f(x)$ is a vector in R^n, whose components (coordinates) are the values of the partial derivatives of this function calculated at point x.

The gradient in R^n is usually denoted either in expanded, component-wise form as

$$\left\| \frac{\partial f}{\partial x_1} \; \frac{\partial f}{\partial x_2} \; \cdots \; \frac{\partial f}{\partial x_n} \right\| ,$$

or symbolically (unexpanded) as $\text{grad} f(x)$ or ∇f.

Let us find out the geometric meaning of the gradient for $n = 2$, first for linear function $f(x, y) = Ax + By$. Recall that the equation

$$Ax + By = C \tag{3.04}$$

on the coordinate plane Oxy defines a straight line, which is *level line* for the function $f(x, y) = Ax + By$.

For for this straight line vector $\|A\,B\|$ with additional restriction (in an orthonormal basis) $A^2 + B^2 \neq 0$ is a normal (orthogonal) vector.

The equalities $\dfrac{\partial f}{\partial x} = A$ and $\dfrac{\partial f}{\partial y} = B$ are obvious. It follows that the gradient of a linear function is collinear to the normal vector of the line (3.04).

Returning to the nonlinear case, remember that the equation of the tangent to the level line $f(x, y) = C$ for the function $f(x, y)$ at the point $\| \, x_0 \, y_0 \, \|$ looks like

$$A(\, x_0 \, y_0 \,)(x - x_0) + B(\, x_0 \, y_0 \,)(y - y_0) = 0 \,,$$

where $A(\, x_0 \, y_0 \,) = \dfrac{\partial f}{\partial x}(\, x_0 \, y_0 \,)$ and $B(\, x_0 \, y_0 \,) = \dfrac{\partial f}{\partial y}(\, x_0 \, y_0 \,)$.

Whence we conclude that the gradient of a function of two variables is *the normal tangent vector* to the level line of this function, calculated at this point.

Graphically this statement is illustrated in Fig. 1.

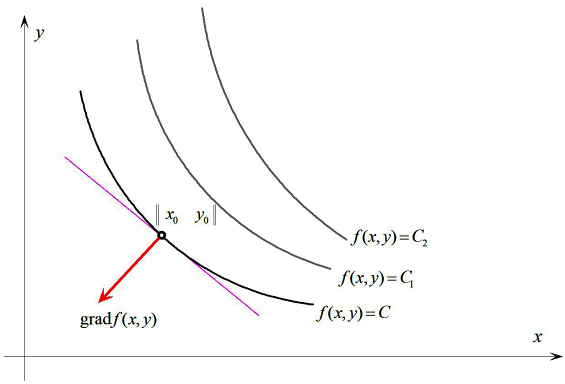

Рис. 1

Another important quantitative characteristic for a function of several variables $f(x)$ is *directional derivative.*

Let r_0 be some point in R^n and w, for which $|w| = 1$, specifies a fixed direction in R^n.

Then the quantity estimating the relative change in the value of the function $f(x)$ for a small displacement from the point r_0 in the direction w is called the *derivative* of the function $f(x)$ *at point* r_0 *in direction* w. Its value gives

Definition 3.2	$$\lim_{t\to+0}\frac{f\left(r_0+tw\right)-f\left(r_0\right)}{t}. \qquad (3.05)$$

The limit (3.05) is obviously a function of both r_0, and w, that is, depending on $2n$ scalar variables. It is usually denoted $\dfrac{\partial f}{\partial w}(x_0, w)$ or simply $\dfrac{\partial f}{\partial w}$.

Differentiability of a function of several variables

As before we will consider this concept in the case of R^2. Let a function $f(x,y)$ be given in the neighborhood of the point $\|x_0\, y_0\|$. Let's give

Definition 3.3	The function $f(x,y)$ is called *differentiable* at the point $\|x_0\, y_0\|$, if there are finite numbers A and B such that $$f(x,y) = f(x_0,y_0) + A\big(x - x_0\big) + B\big(y - y_0\big) + \\ + o\left(\sqrt{\big(x - x_0\big)^2 + \big(y - y_0\big)^2}\right) \qquad (3.06)$$

It is important to note:

unlike the case of a function of one variable, differentiability property and existence of derivatives in R^n *are not equivalent.*

There is

`Theorem 3.2` **For the function $f(x, y)$ to be differentiable at the point $\|x_0\, y_0\|$** *requires the existence* **of partial derivatives $\dfrac{\partial f}{\partial x}$ And $\dfrac{\partial f}{\partial y}$ at this point. In this case $A = \dfrac{\partial f}{\partial x}(x_0, y_0)$ and $B = \dfrac{\partial f}{\partial y}(x_0, y_0)$.**

This condition is not sufficient, which demonstrates

`Example 3.1` Function $f(x, y) = \sqrt{|xy|}$ obviously has partial derivatives at the origin, since on the coordinate axes the function is identically equal to zero.
On the other side,

$$\lim_{\|x\, y\| \to \|0\, 0\|} \frac{\sqrt{|xy|}}{\sqrt{x^2 + y^2}}$$

does not exist by negating the definition of the limit according to Heine.
Indeed, $f(0, 0) = 0$ and when tending to the origin of coordinates along the Ox axis numeric sequence of function values tend to zero. But upon transition to the origin along line $x = y$ this limit is equal to $\frac{1}{\sqrt{2}}$.

At the same time we have

`Theorem 3.3` **If at the point $\|x_0\, y_0\|$ partial derivatives $\dfrac{\partial f}{\partial x}$ and $\dfrac{\partial f}{\partial y}$** *continuous*, **then the function $f(x,y)$ is differentiable at this point.**

This condition, as the following example shows, is not necessary.

`Example 3.2` Consider the function

$$f(x,y) = \begin{cases} 0 & \text{at} \quad x^2+y^2=0, \\ (x^2+y^2)\sin\dfrac{1}{x^2+y^2} & \text{at} \quad x^2+y^2 \neq 0. \end{cases}$$

Its partial derivatives at the origin

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = \lim_{t\to 0} \frac{t^2 \sin\dfrac{1}{t^2} - 0}{t} = 0\,.$$

But they are not continuous functions. For example,

$$\frac{\partial f}{\partial x} = 2x\sin\frac{1}{x^2+y^2} - \frac{2x}{x^2+y^2}\cos\frac{1}{x^2+y^2}$$

is not continuous on $y=0$ at $x\to 0$.

On the other side,

$$\lim_{\|x\ y\| \to \|0\ 0\|} \frac{(x^2+y^2)\sin\dfrac{1}{x^2+y^2}}{\sqrt{x^2+y^2}} = 0$$

due to inequalities

$$0 \ \leq \ \left| \frac{(x^2+y^2)\sin\dfrac{1}{x^2+y^2}}{\sqrt{x^2+y^2}} \right| \ \leq \ \sqrt{x^2+y^2}$$

and theorems about «two policemen». Therefore, the function in question is differentiable at the origin.

Partial derivative of a composit function of several variables

Let the variables x and y be the arguments of the differentiable function $f(x,y)$. Wherein they are differentiable functions of the new independent variable t in some neighborhood of the point t_0. We need to find the derivative of the function $\Phi(t) = f\big(x(t), y(t)\big)$ by variable t.

The solution comes down to the following actions.

Let $x_0 = x(t_0)$ and $y_0 = y(t_0))$, and the function $f(x,y)$ is differentiable at the point $\|x_0\, y_0\|$. Then

$$\Delta f = f(x_0 + \Delta x,\, y_0 + \Delta y) - f(x_0, y_0) =$$

$$= \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y + o(\sqrt{\Delta x^2 + \Delta y^2})$$

Further, since $x(t)$ and $y(t)$ are differentiable functions of *one* variable at point t_0 , then the equalities will be true

$$\Delta x = \frac{dx}{dt}\Delta t \qquad \text{and} \qquad \Delta y = \frac{dy}{dt}\Delta t\,.$$

It gives

$$\Delta f = \frac{\partial f}{\partial x}\frac{dx}{dt}\Delta t + \frac{\partial f}{\partial y}\frac{dy}{dt}\Delta t + o(\Delta t)\,, \tag{3.07}$$

because the $\quad o\left(\sqrt{\Delta x^2 + \Delta y^2}\right) = o\left(\Delta t\sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dx}{dt}\right)^2}\right) = o(\Delta t)\,.$

Finally, dividing both sides of equality (3.07) by Δt and passing to the limit at $\Delta t \to 0$, we obtain the required formula

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}\,. \tag{3.08}$$

Formula (3.08) is easy to generalize to the case when, for example, function f depends on n variables:

$$\frac{df}{dt} = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}\frac{dx_k}{dt} \tag{3.09}$$

or in the case where, in addition, t is an m-dimensional vector, to

$$\frac{\partial f}{\partial t_j} = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}\frac{\partial x_k}{\partial t_j} \qquad \forall j = \overline{1, m}\,.$$

In the case, where the function $f(x)$ has partial derivatives with respect to all its arguments, you can apply the theorem on differentiation of composite functions to calculate the directional derivative.

Because the $x_j(t) = x_0 + t \cdot w_j \quad \forall j = \overline{1,n}$, then from formula (3.09) we obtain

$$\frac{\partial f}{\partial w} = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k} w_k = \big(\text{grad} f, w\big)\,.$$

Now we can estimate the value of the directional derivative, taking advantage of Cauchy-Bunyakovsky inequality. We get

$$\left|\frac{\partial f}{\partial w}\right| = \big|\big(\text{grad} f, w\big)\big| \leq |\,\text{grad} f\,|\,|w| = |\,\text{grad} f\,|\ , \tag{3.10}$$

since $|w| = 1$.

From there follows the assessment

$$-\,|\,\mathrm{grad}f\,| \leq \frac{\partial f}{\partial w} \leq |\,\mathrm{grad}f\,|$$

and the conclusion that the derivative of the function $f(x)$ at the point x_0 is maximum in the direction of w, collinear vector $\mathrm{grad}f(x_0)$.

It can be argued that function $f(x)$ increases fastest in the direction of the gradient and decreases most rapidly in the opposite direction.

The last statement can be considered as well as another geometric property of the gradient.