Производные функции многих переменных

В пространстве R^n имеются различные возможности определения производных для функции многих перерееменных. Рассмотрим некоторые из них.

Самый простой подход к введению производных для функции многих переменных заключается в следующем.

Пусть имеется функция, зависящая от нескольких переменных $f(x) = f\Big(\, x_1,\, x_2,\, \dots\, ,\, x_n\, \Big)$. Выберем в области ее определения некоторую точку $x_{(0)}$ и зафиксируем (то есть, превратим в константы) все компоненты (координаты) этой точки, кроме одной, имеющей номер k.

В этом случае $f(x)$ можно рассматривать как функцию только от одной переменной, от $x_{(0)k}$. Для такой функции у нас имеется понятие *производной функции в точке*, определяемое формулой

$$f'_{x_{(0)k}} = \lim_{t\to 0} \frac{f\Big(\, x_{(0)1},\, \dots\, ,\, x_{(0)k} + t,\, \dots,\, x_{(0)n}\, \Big) - f\Big(\, x_{(0)1},\, \dots\, ,\, x_{(0)n}\, \Big)}{t}\ . \tag{3.01}$$

Эту производную принято называть *частной производной от функции* $f(x)$ *по переменной* x_k и обозначать как $\dfrac{\partial f}{\partial x_k}$.

Заметим, что по аналогии со случаем функции одной переменной для функции многих переменных мы можем ввести понятие *производной функции*, то есть рассматривать новую функцию *многих переменных* вида

$$\frac{\partial f}{\partial x_k}\Big(x_1, x_2, \ldots , x_n \Big) . \tag{3.02}$$

Несложно убедиться, что для функции, являющейся частной производной, справедливы все правила и теоремы, которые описывают свойства производной функции, зависящей только от одной переменной.

Для их использования достаточно считать все компоненты точки $x_{(0)}$, с номерами не равными k некоторыми константами.

Производные высших порядков для функции многих переменных

Поскольку частная производная (3.02) сама является функций многих переменных, то естественно также допустить существование для нее частных производных. При этом переменная, по которой берется новая производная может быть не x_k, а другой, скажем, x_j.

Такие производные называются *смешанными* частными производными и для них принято использовать обозначение вида $f''_{x_k x_j}$ или

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

То есть производные последовательно берутся в порядке, обратном тому, в котором они перечислены в знаменателе записи смешанной производной.

Если $k = j$, то используется обозначение $\dfrac{\partial^2 f}{\partial x_k^2}$.

Оказывается справедливой

Теорема 3.1 **Если частные производные $\dfrac{\partial^2 f}{\partial x_k \partial x_j}$ и $\dfrac{\partial^2 f}{\partial x_j \partial x_k}$ непрерывны в некоторой точке, то они** *равны* **в этой точке.**

Поскольку основные специфические свойства функций многих переменных проявляются уже при $n = 2$, то для их иллюстрации будем ограничиваться только этим случаем.

Задача 3.1 *Пусть* $f(x, y) = \sqrt{x^2 + y^2}$. *Требуется найти все частные производные второго порядка.*

Решение. Областью определения данной функции в R^2 является вся координатная плоскость Oxy.
Для вычисления производных, воспользуемся правилами дифференцирования функции одной переменной.
Последовательно полагая y и x константами, получаем

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

Для вторых производных находим

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

Полученные формулы справедливы не во всей области определения функции $f(x, y)$, поскольку применение *правил* дифференцирования предполагает, что *все* производные, входящие в запись этих правил, существуют.
В нашем примере это верно лишь при $x^2 + y^2 \neq 0$.

Здесь обратим внимание на следующую важную деталь: из того факта, что в начале координат не определены функции в формулах (3.03), вообще говоря, не следует, что функция $f(x,y)$ в этой точке не имеет частных производных.

Отсутствие частных производных в точке $(0;0)$ следует не из неприменимости в ней правил дифференцирования, а из формулы (3.01).

Действительно, при $x_{(0)} = 0$ и $y_{(0)} = 0$ имеем

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

Решение получено.

а вот этот предел не существует.

Градиент и производная по направлению функции многих переменных

С помощью частных производных имеется возможность более точно описывать и анализировать свойства функций многих переменных.

Одним из инструментов такого анализа служит *градиент* функции многих переменных.

Пусть функция многих переменных $f(x) = f\Big(x_1, x_2, \ldots, x_n\Big)$. имеет в некоторой точке $x = \Big\| x_1, x_2, \ldots, x_n \Big\|$ своей области определения it все частные производные первого порядка

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

Определение 3.1	*Градиентом* в точке x функции многих переменных $f(x)$ называется вектор в R^n, компонентами (координатами) которого служат значения частных производных этой функции, вычисленные в данной точке.

Градиент принято обозначать либо в развернутом, покомпонентном виде как

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

либо символически (в неразвернутом виде) как $\operatorname{grad} f(x)$ или ∇f.

Выясним геометрический смысл градиента при $n = 2$, вначале для случая линейной функции $f(x, y) = Ax + By$. Напомним, что уравнение

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

на координатной плоскости Oxy задает прямую, являющуюся *линией уровня* для функции $f(x, y) = Ax + By$.

Для для этой прямой вектор $\left\| A \; B \right\|$ при дополнительном ограничении (в ортонормированном базисе) $A^2 + B^2 \neq 0$ является нормальным (ортогональным) вектором.

С другой стороны, очевидны равенства $\dfrac{\partial f}{\partial x} = A$ и $\dfrac{\partial f}{\partial y} = B$, из которых следует, что градиент линейной функции коллинеарен нормальному вектору прямой (3.04).

Возвращаясь к нелинейному случаю, вспомним, что уравнение касательной к линии уровня $f(x, y) = C$ для функции $f(x, y)$ в точке $\|\, x_0\; y_0 \,\|$ имеет вид

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

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

Откуда заключаем, что градиент функции двух переменных является *нормальным вектором касательной* к линии уровня этой функции, проведенной в данной точке.

Графически это утверждение иллюстрирует рис. 1.

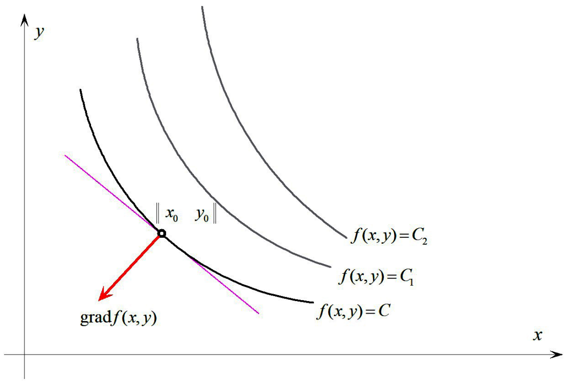

Рис. 1

Другой важной, часто используемой в приложениях количественной характеристикой функции многих переменных $f(x)$ является *производная по направлению.*

Пусть r_0 некоторая точка в R^n , а вектор w, для которого $|w| = 1$, задает в R^n фиксированное направление.

Тогда величина, оценивающая относительное изменение значения функции $f(x)$ при малом смещении из точки r_0 по направлению w, называется *производной* функции $f(x)$ *в точке* r_0 *по направлению* w. Ее значение дает

Определение 3.2	$$\lim_{t\to+0} \frac{f\Big(r_0 + tw\Big) - f\Big(r_0\Big)}{t}\,. \qquad (3.05)$$

Предел (3.05) является функцией очевидно как от r_0, так и от w, то есть, зависящей от $2n$ скалярных переменных. Его принято обозначать $\dfrac{\partial f}{\partial w}(\,x_0,\,w\,)$ или, просто $\dfrac{\partial f}{\partial w}$.

Дифференцируемость функции многих переменных

Рассмотрим данное понятие в случае R^2. Пусть дана функция $f(x,y)$ и точка $\|x_0\, y_0\|$. Дадим

Определение 3.3	Функция $f(x,y)$ называется *дифференцируемой* в точке $\|x_0\, y_0\|$, если существуют конечные числа A и B такие, что $$f(x,y) = f(x_0,y_0) + A\big(x-x_0\big) + B\big(y-y_0\big) + \\ +o\left(\sqrt{\big(x-x_0\big)^2 + \big(y-y_0\big)^2}\right) \qquad (3.06)$$

Важно отметить: в отличие от случая функции одной переменной, свойство дифференцируемости и существование частных производных в R^n *не равносильны.*

Имеет место

Теорема 3.2 **Для дифференцируемости функции $f(x, y)$ в точке $\|x_0\, y_0\|$** *необходимо существование* **частных производных $\dfrac{\partial f}{\partial x}$ и $\dfrac{\partial f}{\partial y}$ в этой точке. В этом случае $A = \dfrac{\partial f}{\partial x}(x_0, y_0)$ и $B = \dfrac{\partial f}{\partial y}(x_0, y_0)$.**

Это условие не является достаточным, что демонстрирует

Пример 3.1 Функция $f(x, y) = \sqrt{|xy|}$ очевидно имеет в начале координат частные производные, поскольку на координатных осях она тождественно равна нулю.
С другой стороны,

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

не существует по отрицанию определения предела по Гейне, поскольку в данном примере $f(0, 0) = 0$ и при стремлении в начало координат по оси Ox числовая последовательность значений функции стремиться к нулю, а при переходе в начало координат по биссектрисе первого координатного угла (где $x = y$) этот предел равен $\frac{1}{\sqrt{2}}$.

При этом справедлива

Теорема 3.3 **Если в точке $\|x_0\, y_0\|$ частные производные $\dfrac{\partial f}{\partial x}$ и $\dfrac{\partial f}{\partial y}$** *непрерывны*, **то функция $f(x, y)$ дифференцируема в этой точке.**

Это условие достаточное и, как показывает следующий пример, не является необходимым.

Пример 3.2 Рассмотрим функцию

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

Ее частные производные в начале координат

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

но не являются непрерывными функциями поскольку при $x^2 + y^2 \neq 0$, например,

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

С другой стороны,

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

в силу неравенств

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

и теоремы о двух милиционерах. Следовательно, рассматриваемая функция дифференцируема в начале координат.

Частная производная сложной функции многих переменных

Пусть переменные x и y — аргументы дифференцируемой функции $f(x, y)$ — являются дифференцируемыми функциями $x(t)$ и $y(t)$ от новой независимой переменной t в некоторой окрестности точки t_0 . Рассмотрим задачу отыскания производной функции $\Phi(t) = f\big(x(t), y(t)\big)$ по переменной t .

Ее решение сводится к следующим действиям.

Пусть $x_0 = x(t_0)$ и $y_0 = y(t_0))$, а функция $f(x, y)$ дифференцируема в точке $\|x_0\, y_0\|$. Тогда

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

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

Далее, поскольку $x(t)$ и $y(t)$ дифференцируемые функции *одной* переменной в точке t_0 , то будут верны равенства

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

Значит,

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

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

Наконец, разделив обе части равенства (3.07) на Δt и перейдя к пределу при $\Delta t \to 0$, получим искомую формулу

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

Формулу (3.08) нетрудно обобщить на случай, когда, например, функция f зависит от n переменных:

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

или на случай, когда, кроме того, и t является m-мерным вектором

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

В случае, когда функция $f(x)$ имеет частные производные по всем своим аргументам, можно применить теорему о дифференцировании сложной функции для вычисления производеой по направлению.

Поскольку $x_j(t) = x_{0j} + t \cdot w_j \quad \forall j = \overline{1, n}$, то из формулы (3.09) в ортонормированном базисе получаем

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

Теперь можно оценить величину производной по направлению, воспользовавшись неравенством Коши-Буняковского. Получаем

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

поскольку $|w| = 1$.

Откуда следует оценка

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

и вывод о том, что производная функции $f(x)$ в точке x_0 максимальна по направлению w, коллинеарному вектору $\mathrm{grad} f(x_0)$.

Другими словами, исходя из (3.10), можно утверждать, что, имеющая в точке x_0 все частные производные, функция $f(x)$ возрастает быстрее всего по направлению градиента в этой точке, а убывает быстрее всего по противоположному направлению.

Последнее утверждение можно рассматривать также как еще одно геометрическое свойство градиента.