Московский государственный технический университет
имени Н.Э. Баумана

Факультет «Фундаментальные науки»
Кафедра «Математическое моделирование»

А.Н. Канатников, А.П. Крищенко

ФУНКЦИИ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ

Конспект лекций

Учебное пособие по дисциплине
«Линейная алгебра и функции нескольких переменных»
для студентов всех специальностей

Москва

Лекция 16

ДИФФЕРЕНЦИРУЕМОСТЬ ВЕКТОРНЫХ ФУНКЦИЙ НЕСКОЛЬКИХ ПЕРЕМЕННЫХ

Матрица Якоби ВФНП, якобиан (при $n = m$). Дифференцируемость ВФНП, ее дифференциал. Производная сложной ВФНП в матричной форме. Теорема о неявной функции в общем случае. Теорема об обратной функции.

16.1. Частные производные

Пусть векторная функция нескольких переменных $f\colon \mathbb{R}^n \to \mathbb{R}^m$ определена в δ-окрестности $\mathrm{U}(a,\delta)$ точки $a \in \mathbb{R}^n$. Обозначим через Δx_i такое приращение независимого переменного x_i в точке a, при котором точка $a = (a_1,\ \ldots,\ a_{i-1},\ a_i{+}\Delta x_i,\ a_{i+1},\ \ldots,\ a_n)$ принадлежит $\mathrm{U}(a,\delta)$. Для этого достаточно, чтобы выполнялось неравенство $|\Delta x_i| < \delta$. Тогда определена разность значений функции f, соответствующая приращению Δx_i:

$$\Delta_i f(a, \Delta x_i) = f(a_1, \ldots, a_{i-1}, a_i + \Delta x_i, a_{i+1}, \ldots, a_n) - f(a_1, \ldots, a_n).$$

Эту разность называют ***частным приращением функции нескольких переменных*** f в точке a по независимому переменному x_i. Частное приращение обозначают также через $\Delta_i f(a)$ или $\Delta_{x_i} f(a)$.

Определение 16.1. Если для функции нескольких переменных $f\colon \mathbb{R}^n \to \mathbb{R}^m$, определенной в окрестности точки a, существует предел

$$\lim_{\Delta x_i \to 0} \frac{\Delta_i f(a)}{\Delta x_i} \tag{16.1}$$

отношения частного приращения функции по переменному x_i к приращению Δx_i этого же переменного при $\Delta x_i \to 0$, то этот предел называют ***частной производной векторной функции нескольких переменных f в точке a*** по переменному x_i и обозначают f'_{x_i}.

Теорема 16.1. Для того чтобы *векторная функция* $f\colon \mathrm{U}(a,\delta) \subset \mathbb{R}^n \to \mathbb{R}^m$ имела частную производную в точке a по переменному x_i, необходимо и достаточно, чтобы все ее *координатные функции* имели частную производную в точке a по тому же переменному x_i.

◄ Пусть Δx_i — приращение независимого переменного x_i в точке a. Тогда соответствующее приращение функции f в точке a можно записать в виде

$$\Delta_i f(a) = \begin{pmatrix} f_1(a_1, \ldots, a_{i-1}, a_i + \Delta x_i, a_{i+1}, \ldots, a_n) \\ \vdots \\ f_m(a_1, \ldots, a_{i-1}, a_i + \Delta x_i, a_{i+1}, \ldots, a_n) \end{pmatrix} - \begin{pmatrix} f_1(a) \\ \vdots \\ f_m(a) \end{pmatrix} =$$

$$= \begin{pmatrix} f_1(a_1, \ldots, a_{i-1}, a_i + \Delta x_i, a_{i+1}, \ldots, a_n) - f_1(a) \\ \vdots \\ f_m(a_1, \ldots, a_{i-1}, a_i + \Delta x_i, a_{i+1}, \ldots, a_n) - f_m(a) \end{pmatrix} = \begin{pmatrix} \Delta_i f_1(a) \\ \vdots \\ \Delta_i f_m(a) \end{pmatrix}.$$

Согласно определению 16.1 частной производной, имеем

$$\frac{\partial f(a)}{\partial x_i} = \lim_{\Delta x_i \to 0} \frac{\Delta_i f(a)}{\Delta x_i} = \lim_{\Delta x_i \to 0} \begin{pmatrix} \dfrac{\Delta_i f_1(a)}{\Delta x_i} \\ \vdots \\ \dfrac{\Delta_i f_m(a)}{\Delta x_i} \end{pmatrix}.$$

Используя теорему 15.1, получаем

$$\frac{\partial f(a)}{\partial x_i} = \begin{pmatrix} \lim\limits_{\Delta x_i \to 0} \dfrac{\Delta_i f_1(a)}{\Delta x_i} \\ \vdots \\ \lim\limits_{\Delta x_i \to 0} \dfrac{\Delta_i f_m(a)}{\Delta x_i} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial f_1(a)}{\partial x_i} \\ \vdots \\ \dfrac{\partial f_m(a)}{\partial x_i} \end{pmatrix},$$

что и требовалось доказать. ►

При доказательстве теоремы установлена формула, согласно которой частная производная векторной функции $f(x)$ равна векторной функции, координатными функциями которой являются соответствующие частные производные координатных функций для $f(x)$. Следовательно, вычисление частных производных векторной функции сводится к вычислению соответствующих частных производных ее координатных функций, которые являются функциями скалярными.

Если функция $f\colon \mathbb{R}^n \to \mathbb{R}^m$ в точке $a \in \mathbb{R}^n$ имеет частные производные по всем независимым переменным x_1, x_2, ..., x_n, то из этих производных (а точнее, из частных производных координатных функций $f_1(x)$, $f_2(x)$, ..., $f_m(x)$ векторной функции $f(x)$) можно составить матрицу $\left(\dfrac{\partial f_i(a)}{\partial x_j}\right)$ типа $m \times n$, где i соответствует номеру строки матрицы, а j — номеру столбца. Эту матрицу называют ***матрицей Якоби*** функции f в точке a и обозначают

$$f'(x) = \frac{\partial f(a)}{\partial x} = \begin{pmatrix} \dfrac{\partial f_1(a)}{\partial x_1} & \dfrac{\partial f_1(a)}{\partial x_2} & \dots & \dfrac{\partial f_1(a)}{\partial x_n} \\ \dfrac{\partial f_2(a)}{\partial x_1} & \dfrac{\partial f_2(a)}{\partial x_2} & \dots & \dfrac{\partial f_2(a)}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \dfrac{\partial f_m(x)}{\partial x_1} & \dfrac{\partial f_m(x)}{\partial x_2} & \dots & \dfrac{\partial f_m(x)}{\partial x_n} \end{pmatrix}. \tag{16.2}$$

Часто используют запись матрицы Якоби в виде блочной матрицы-строки

$$f'(x) = \begin{pmatrix} \dfrac{\partial f(x)}{\partial x_1} & \dfrac{\partial f(x)}{\partial x_2} & \dots & \dfrac{\partial f(x)}{\partial x_n} \end{pmatrix} \tag{16.3}$$

или блочной матрицы-столбца

$$f'(x) = \begin{pmatrix} \dfrac{\partial f_1(x)}{\partial x} \\ \vdots \\ \dfrac{\partial f_m(x)}{\partial x} \end{pmatrix}. \tag{16.4}$$

В последнем случае каждый блок представляет собой матрицу Якоби соответствующей координатной функции.

Пример 16.1. Для векторной функции

$$f(x,y) = \begin{pmatrix} xe^{-y} & x^2y^3 & y \end{pmatrix}^{\mathrm{T}}$$

двух переменных x и y найдем все частные производные и запишем матрицу Якоби.

Данная функция имеет три координатные функции $u(x,y) = xe^{-y}$, $v(x,y) = x^2y^3$ и $w(x,y) = y$. Вычисляем частные производные этих функций:

$$\begin{aligned} &u'_x(x,y) = e^{-y}, \qquad && u'_y(x,y) = -xe^{-y},\\ &v'_x(x,y) = 2xy^3, && v'_y(x,y) = 3x^2y^2,\\ &w'_x(x,y) = 0, && w'_y(x,y) = 1. \end{aligned}$$

Составляем из вычисленных частных производных матрицу Якоби:

$$f'(x,y) = \begin{pmatrix} u'_x(x,y) & u'_y(x,y)\\ v'_x(x,y) & v'_y(x,y)\\ w'_x(x,y) & w'_y(x,y) \end{pmatrix} = \begin{pmatrix} e^{-y} & -xe^{-y}\\ 2xy^3 & 3x^2y^2\\ 0 & 1 \end{pmatrix}.$$

16.2. Дифференцируемые векторные функции

Пусть векторная функция нескольких переменных $f\colon \mathbb{R}^n \to \mathbb{R}^m$ определена в некоторой *окрестности точки* $x \in \mathbb{R}^n$ и $\Delta x = (\Delta x_1 \ \ldots \ \Delta x_n)^{\mathrm{T}}$ — такой вектор приращений независимых переменных, что точка $x + \Delta x$ тоже принадлежит этой окрестности. В этом случае определено ***полное приращение функции*** f

$$\Delta f(x) = f(x+\Delta x) - f(x),$$

соответствующее приращению Δx переменных в точке x. Полное приращение функции $f(x) = (f_1(x) \ \ldots \ f_m(x))^{\mathrm{T}}$ в точке x можно выразить через полные приращения координатных функций $f_1(x)$, $f_2(x)$, $\ldots$, $f_m(x)$:

$$\Delta f(x) = f(x+\Delta x) - f(x) = \begin{pmatrix} f_1(x+\Delta x)\\ \vdots\\ f_m(x+\Delta x) \end{pmatrix} - \begin{pmatrix} f_1(x)\\ \vdots\\ f_m(x) \end{pmatrix} =$$

$$= \begin{pmatrix} f_1(x+\Delta x) - f_1(x)\\ \vdots\\ f_m(x+\Delta x) - f_m(x) \end{pmatrix} = \begin{pmatrix} \Delta f_1(x)\\ \vdots\\ \Delta f_m(x) \end{pmatrix}.$$

Кроме того, напомним, что $|\Delta x| = \sqrt{(\Delta x_1)^2 + \ldots + (\Delta x_n)^2}$.

Определение 16.2. Функцию $f\colon \mathbb{R}^n \to \mathbb{R}^m$, определенную в некоторой окрестности точки x, называют ***дифференцируемой в точке*** x, если ее полное приращение в окрестности этой точки можно представить в виде

$$\Delta f(x) = A\Delta x + \alpha(\Delta x)|\Delta x|, \tag{16.5}$$

где A — матрица типа $m \times n$, элементы которой не зависят от Δx, а функция $\alpha(\Delta x)$ является *бесконечно малой* при $\Delta x \to 0$.

Функцию f называют ***дифференцируемой в области*** $X \subset \mathbb{R}^n$, если она дифференцируема в каждой точке этой области.

При $m = 1$ функция f скалярная, и в равенстве (16.5) матрица A является строкой длины n, т.е. $A = (a_1 \ a_2 \ \ldots \ a_n)$, а функция $\alpha(\Delta x)$ — это бесконечно малая при $\Delta x \to 0$ скалярная функция. Поэтому в данном случае равенство (16.5) сводится к равенству (9.2).

Следующая теорема сводит исследование дифференцируемости *векторной функции* к скалярному случаю.

Теорема 16.2. Векторная функция $f\colon \mathbb{R}^n \to \mathbb{R}^m$ дифференцируема в точке x тогда и только тогда, когда в этой точке дифференцируемы все ее *координатные функции.*

◀ Доказательство фактически состоит в переходе от матричной формы записи условия (16.5) дифференцируемости векторной функции к его записи в координатной форме. Действительно, в равенстве (9.2) положим $f(x) = \bigl(f_1(x) \ \ \ldots \ \ f_m(x)\bigr)^{\mathrm{т}}$, $\alpha(\Delta x) = \bigl(\alpha_1(\Delta x) \ \ \ldots \ \ \alpha_m(x)\bigr)^{\mathrm{т}}$ и $A = (a_{ij})$. Тогда (9.2) можно записать следующим образом:

$$\begin{pmatrix} \Delta f_1(x) \\ \vdots \\ \Delta f_m(x) \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} \Delta x_1 \\ \vdots \\ \Delta x_n \end{pmatrix} + \begin{pmatrix} \alpha_1(x, \Delta x) \\ \vdots \\ \alpha_m(x, \Delta x) \end{pmatrix} |\Delta x|,$$

или в координатной записи

$$\Delta f_i(x) = a_{i1}\Delta x_1 + \ldots + a_{in}\Delta x_n + \alpha_i(\Delta x)|\Delta x|, \quad i = \overline{1,\, m}, \tag{16.6}$$

где $\alpha_i(\Delta x) \to 0$ при $\Delta x \to 0$.

Итак, соотношение (16.5) эквивалентно (16.6), но представление (16.5) по определению означает дифференцируемость векторной функции $f(x)$, а представления (16.6) — дифференцируемость координатных функций $f_i(x)$, $i = \overline{1,\, m}$. ▶

Следствие 16.1. Если функция $f\colon \mathbb{R}^n \to \mathbb{R}^m$ дифференцируема в точке x, то в этой точке существуют частные производные этой функции по всем переменным, определена ее *матрица Якоби* $f'(x)$ и в *окрестности* этой *точки* полное приращение $\Delta f(x)$ можно представить в виде

$$\Delta f(x) = f'(x)\Delta x + \alpha(\Delta x)|\Delta x|, \tag{16.7}$$

где $\alpha(\Delta x) \to 0$ при $\Delta x \to 0$.

◀ При $m = 1$ утверждение следствия вытекает из следствия 9.1. Поэтому остановимся на случае $m > 1$. Согласно теореме 16.2, из дифференцируемости функции $f(x) = \bigl(f_1(x) \ \ \ldots \ \ f_m(x)\bigr)^{\mathrm{т}}$ в точке x следует дифференцируемость в этой точке всех ее *координатных функций* f_i. При этом в представлении (16.6) коэффициенты a_{ij} есть значения частных производных координатной функции f_i в точке x по соответствующим переменным:

$$\Delta f_i(x) = \frac{\partial f_i(x)}{\partial x_1}\Delta x_1 + \ldots + \frac{\partial f_i(x)}{\partial x_n}\Delta x_n + \alpha_i(\Delta x)|\Delta x|. \tag{16.8}$$

Значит, в представлении (16.5) матрица A есть матрица, составленная из значений частных производных координатных функций в точке x, т.е. матрица Якоби $f'(x)$, а векторная функция $\alpha(\Delta x)$ имеет своими координатными функциями функции $\alpha_i(\Delta x)$. Так как все функции $\alpha_i(\Delta x)$ являются бесконечно малыми при $\Delta x \to 0$, то и векторная функция $\alpha(\Delta x)$ является бесконечно малой при $\Delta x \to 0$. ▶

Следствие 16.2. Если векторная функция дифференцируема в некоторой *области*, то во всех точках этой области существуют частные производные ее координатных функций и, следовательно, в области существует ее матрица Якоби.

Как и в случае действительных функций одного действительного переменного, есть еще одно необходимое условие дифференцируемости функции нескольких переменных, связанное с ее непрерывностью.

Теорема 16.3. Если *функция нескольких переменных* дифференцируема в некоторой точке, то она непрерывна в этой точке.

◀ Пусть функция $f(x) = (f_1(x) \ \ldots \ f_m(x))^{\mathrm{T}}$ дифференцируема в точке a. Тогда по теореме 16.2 все ее координатные функции $f_i(x)$ дифференцируемы в точке a, а их полные приращения в точке a можно записать в виде

$$\Delta f_i(a) = \sum_{k=1}^{n} \frac{\partial f_i(a)}{\partial x_k} \Delta x_k + \alpha_i(\Delta x)|\Delta x|,$$

где $\Delta x = (\Delta x_1,\ \Delta x_2,\ \ldots,\ \Delta x_n)$ и $\alpha_i(\Delta x) \to 0$ при $\Delta x \to 0$. Из этого представления следует, что существует предел

$$\lim_{\Delta x \to 0} \Delta f_i(a) = \sum_{k=1}^{n} \frac{\partial f_i(a)}{\partial x_k} \lim_{\Delta x \to 0} \Delta x_k + \lim_{\Delta x \to 0} \bigl(\alpha_i(\Delta x)|\Delta x|\bigr) = 0,$$

означающий, что функции $f_i(x)$, $i = \overline{1,\, m}$, непрерывны в точке a. Действительно, $\Delta x = x - a$, откуда заключаем, что $\Delta x \to 0$ при $x \to a$ и, следовательно, $\Delta f_i(a) \to 0$. Таким образом, $f_i(x) \to f_i(a)$ при $x \to a$, что и означает непрерывность координатной функции $f_i(x)$ в точке a.

Так как все координатные функции $f_i(x)$ непрерывны в точке a, то по теореме 15.3 и векторная функция $f(x)$ непрерывна в точке a. ▶

Следствие 16.3. Если функция нескольких переменных дифференцируема в некоторой области, то она непрерывна в этой области.

Следствие 16.4. Если векторная функция f имеет матрицу Якоби всюду в некоторой окрестности точки a, причем все элементы матрицы Якоби непрерывны в самой точке a, то эта функция дифференцируема в точке a.

◀ Из условий следствия заключаем, что каждая *координатная функция* f_i векторной функции f имеет в окрестности точки a все частные производные, непрерывные в точке a. Значит, все функции f_i дифференцируемы в точке a. Согласно теореме 16.2, векторная функция f дифференцируема в точке a. ▶

Для произвольной области $X \subset \mathbb{R}^n$ через $C^1(X, \mathbb{R}^m)$ обозначают множество всех функций $f\colon X \to \mathbb{R}^m$, имеющих в X непрерывные частные производные по всем переменным. Для дифференцируемости векторной функции f в области X достаточно, чтобы она принадлежала множеству функций $C^1(X, \mathbb{R}^m)$. Функции из множества $C^1(X, \mathbb{R}^m)$ называют ***непрерывно дифференцируемыми*** в области X.

На векторные функции нескольких переменных можно распространить *правило дифференцирования сложной функции*, установленное для скалярных функций. При этом формула дифференцирования с учетом использования матричной записи упрощается.

Теорема 16.4. Если функция нескольких переменных $f\colon \mathbb{R}^n \to \mathbb{R}^m$ дифференцируема в точке a, а функция нескольких переменных $g\colon \mathbb{R}^m \to \mathbb{R}^k$ дифференцируема в точке $b = f(a)$, $b \in \mathbb{R}^m$, то в некоторой окрестности точки a определена *сложная функция* $g \circ f\colon \mathbb{R}^n \to \mathbb{R}^k$, причем эта функция дифференцируема в точке a и выполнено равенство

$$(g \circ f)'(a) = g'(b) f'(a). \tag{16.9}$$

◀ Пусть функция g определена в *окрестности* $\mathrm{U}(b, \sigma)$ *точки* b. Так как функция f дифференцируема в точке a, она определена в некоторой *окрестности* этой *точки* и является *непрерывной функцией в точке* a. Значит, согласно определению непрерывности, существует такая окрестность $\mathrm{U}(a, \delta)$ из *области определения* функции f, для которой $f(\mathrm{U}(a, \delta)) \subset \mathrm{U}(b, \sigma)$. В окрестности $\mathrm{U}(a, \delta)$ определена сложная функция $g \circ f$.

Пусть x — произвольная точка в $\mathrm{U}(a, \delta)$ и $y = f(x)$, $z = g(y)$. Обозначим $\Delta x = x - a$, $\Delta y = y - b$, $\Delta z = z - c$, где $c = g(b)$. В силу дифференцируемости функции f в точке a имеем представление

$$\Delta y = f'(a) \Delta x + \alpha(\Delta x)|\Delta x|, \tag{16.10}$$

где $\alpha(\Delta x) \to 0$ при $\Delta x \to 0$. В силу дифференцируемости g в точке b имеем аналогичное представление

$$\Delta z = g'(b)\Delta y + \beta(\Delta y)|\Delta y|, \tag{16.11}$$

где $\beta(\Delta y) \to 0$ при $\Delta y \to 0$. Подставив (16.10) в (16.11), получим

$$\Delta(g\circ f)(a) = \Delta z = g'(b)\big(f'(a)\Delta x + \alpha(\Delta x)|\Delta x|\big) + \beta\big(\Delta y\big)|\Delta y| = g'(b)f'(a)\Delta x + \gamma(\Delta x)|\Delta x|, \tag{16.12}$$

где

$$\gamma(\Delta x) = g'(b)\alpha(\Delta x) + \beta(\Delta f(a))\big|f'(a)\nu(\Delta x) + \alpha(\Delta x)\big|,$$

и $\nu(\Delta x) = \dfrac{\Delta x}{|\Delta x|}$.

Функция $\beta(\Delta y)$ *бесконечно малая* при $\Delta y \to 0$, причем на представление (16.11) не влияет значение этой функции при $\Delta y = 0$. Поэтому можно считать, что $\beta(0) = 0$ и что функция $\beta(\Delta y)$ непрерывна при $\Delta y = 0$. Но тогда функция $\beta(\Delta f(a))$ непрерывна при $\Delta x = 0$ как *композиция* непрерывных *функций* и, следовательно, является бесконечно малой при $\Delta x \to 0$. Функция $\nu(\Delta x)$ является ограниченной: $|\nu(\Delta x)| = 1$. Следовательно, функция

$$\beta(\Delta f(a))\big|f'(a)\nu(\Delta x) + \alpha(\Delta x)\big|$$

является бесконечно малой, так как представляет собой произведение бесконечно малой функции $\beta(\Delta f(a))$ на ограниченную функцию $\big|f'(a)\nu(\Delta x) + \alpha(\Delta x)\big|$. В результате заключаем, что $\gamma(\Delta x) \to 0$ при $\Delta x \to 0$. Согласно определению 16.2, представление (16.12) означает, что функция $g\circ f$ дифференцируема в точке a. При этом произведение $g'(b)\,f'(a)$ двух матриц Якоби является, согласно (16.12), матрицей Якоби сложной функции $g\circ f$, т.е. имеет место равенство (16.9). ►

Композицию $(g\circ f)(x) = g(f(x))$ двух функций $f\colon \mathbb{R}^n \to \mathbb{R}^m$ и $g\colon \mathbb{R}^m \to \mathbb{R}^k$ часто задают в виде $z = g(y)$, $y = f(x)$, вводя ***промежуточные переменные*** $y = (y_1,\ y_2,\ \ldots,\ y_m) \in R^m$. Используя координатные функции $f_1, f_2, \ldots, f_m$ и $g_1, g_2, \ldots, g_k$ функций нескольких переменных f и g, равенство (16.9) матриц Якоби можно записать в координатной форме

$$\frac{\partial z_i}{\partial x_j} = \sum_{s=1}^{m} \frac{\partial g_i}{\partial y_s}\frac{\partial f_s}{\partial x_j} = \frac{\partial g_i}{\partial y_1}\frac{\partial f_1}{\partial x_j} + \ldots + \frac{\partial g_i}{\partial y_m}\frac{\partial f_m}{\partial x_j}, \quad i = \overline{1,\ k},\ j = \overline{1,\ n}. \tag{16.13}$$

Пусть $n = m = k = 2$, т.е. $f\colon \mathbb{R}^2 \to \mathbb{R}^2$, $f = (f_1\ \ f_2)^{\mathrm{T}}$, $g\colon \mathbb{R}^2 \to \mathbb{R}^2$, $g = (g_1\ \ g_2)^{\mathrm{T}}$. Тогда для сложной функции $z = g\big(f(x_1, x_2)\big)$ равенство (16.9) в матричной форме при обозначениях $z = g(y)$, $y = f(x_1, x_2)$, $y = (y_1,\ y_2)$ имеет следующий вид:

$$\begin{pmatrix} (z_1)'_{x_1} & (z_1)'_{x_2} \\ (z_2)'_{x_1} & (z_2)'_{x_2} \end{pmatrix} = \begin{pmatrix} (g_1)'_{y_1} & (g_1)'_{y_2} \\ (g_2)'_{y_1} & (g_2)'_{y_2} \end{pmatrix} \begin{pmatrix} (f_1)'_{x_1} & (f_1)'_{x_2} \\ (f_2)'_{x_1} & (f_2)'_{x_2} \end{pmatrix}.$$

Перемножая матрицы в правой части этого равенства, получаем

$$\frac{\partial z_1}{\partial x_1} = \frac{\partial g_1}{\partial y_1}\frac{\partial f_1}{\partial x_1} + \frac{\partial g_1}{\partial y_2}\frac{\partial f_2}{\partial x_1}, \qquad \frac{\partial z_1}{\partial x_2} = \frac{\partial g_1}{\partial y_1}\frac{\partial f_1}{\partial x_2} + \frac{\partial g_1}{\partial y_2}\frac{\partial f_2}{\partial x_2},$$

$$\frac{\partial z_2}{\partial x_1} = \frac{\partial g_2}{\partial y_1}\frac{\partial f_1}{\partial x_1} + \frac{\partial g_2}{\partial y_2}\frac{\partial f_2}{\partial x_1}, \qquad \frac{\partial z_2}{\partial x_2} = \frac{\partial g_2}{\partial y_1}\frac{\partial f_1}{\partial x_2} + \frac{\partial g_2}{\partial y_2}\frac{\partial f_2}{\partial x_2},$$

где частные производные $\dfrac{\partial f_j}{\partial x_i}$, $\dfrac{\partial z_k}{\partial x_i}$ вычисляются в точке $(x_1,\ x_2)$, а частные производные $\dfrac{\partial g_k}{\partial y_j}$ — в точке $(y_1(x_1, x_2),\ \ y_2(x_1, x_2))$.

16.3. Дифференциал векторной функции

Пусть *векторная функция нескольких переменных* $f\colon \mathbb{R}^n \to \mathbb{R}^m$ определена в *окрестности точки* $x = (x_1, \ldots, x_n)$ и *дифференцируема* в этой точке. Тогда, согласно следствию 16.1, *полное приращение* этой *функции* в точке x в зависимости от приращения $\Delta x = (\Delta x_1 \ \ldots \ \Delta x_n)^{\mathrm{T}}$ независимых переменных можно представить в виде

$$\Delta f(x) = f'(x)\Delta x + \alpha(\Delta x)|\Delta x|,$$

где $f'(x)$ — *матрица Якоби* функции $f(x)$, а функция $\alpha(\Delta x)$ является *бесконечно малой функцией* при $\Delta x \to 0$. Как и в случае скалярных функций, можно ввести следующее понятие.

Определение 16.3. Линейную относительно Δx часть $f'(x)\Delta x$ полного приращения функции $f(x)$, дифференцируемой в точке x, называют (***полным***) ***дифференциалом функции*** f и обозначают через $df(x)$.

Итак, дифференциал функции нескольких переменных, дифференцируемой в точке x, вычисляется по той же формуле, что и в случае функции одного переменного: $df(x) = f'(x)\Delta x$. Правда, в многомерном случае $f'(x)$ обозначает не производную функции, а ее матрицу Якоби. Отметим, что формула (10.1) для дифференциала скалярной функции нескольких переменных также может быть записана в матричной форме, если в качестве матрицы Якоби ввести матрицу-строку частных производных скалярной функции.

Дифференциалы независимых переменных x_i, $i = \overline{1, n}$, как и в случае скалярных функций, по определению равны приращениям этих переменных: $dx_i = \Delta x_i$. С учетом этого дифференциал функции f можно записать в виде

$$df(x) = f'(x)\, dx, \quad dx = (dx_1 \ dx_2 \ \ldots \ dx_n)^{\mathrm{T}}. \tag{16.14}$$

Для полного приращения дифференцируемой функции нескольких переменных имеем равенство

$$\Delta f(x) = df(x) + \alpha(\Delta x)|\Delta x| = f'(x)\, dx + \alpha(\Delta x)|\Delta x|, \tag{16.15}$$

где $\alpha(\Delta x) \to 0$ при $\Delta x \to 0$.

Представив матрицу Якоби $f'(x)$ как набор столбцов: $f'(x) = \bigl(f'_{x_1} \ f'_{x_2} \ \ldots f'_{x_n}\bigr)$, равенство (16.14) можно записать следующим образом:

$$df(x) = f'_{x_1}(x)\, dx_1 + f'_{x_2}(x)\, dx_2 + \ldots + f'_{x_n}(x)\, dx_n.$$

Слагаемые $f'_{x_i}\, dx_i$ в правой части равенства называют ***частными дифференциалами функции*** $f(x)$ в точке x. Каждое слагаемое $f'_{x_i}\, dx_i$ представляет собой линейную часть *частного приращения* $\Delta_i f(x)$ *функции* $f(x)$ в данной точке.

Дифференциал функции нескольких переменных, как и функции одного действительного переменного, имеет свойство, которое называют ***инвариантностью формы записи дифференциала***.

Пусть функция $f\colon \mathbb{R}^n \to \mathbb{R}^m$ дифференцируема в точке $a \in \mathbb{R}^n$, а функция $g\colon \mathbb{R}^m \to \mathbb{R}^k$ дифференцируема в точке $b = f(a)$. Согласно теореме 10.1, композиция $g \circ f$ двух функций дифференцируема в точке a. Введем набор промежуточных переменных y и запишем сложную функцию в виде $z = g(y)$, $y = f(x)$. Тогда в соответствии с (16.9)

$$dz = (g \circ f)'(a)\, dx = g'(b)f'(a)\, dx = g'(b)\, dy$$

Мы видим, что дифференциал dz сложной функции $z = g(f(x))$ выражается через дифференциал dy промежуточных переменных так же, как и в случае, когда эти переменные являются независимыми.

Замечание 16.1. Для дифференциала векторной функции нескольких переменных сохраняются свойства дифференциала скалярной функции нескольких переменных. Например, для дифференцируемых функций $f, g\colon \mathbb{R}^n \to \mathbb{R}^m$ и произвольного действительного числа c верны равенства $d(cf) = cdf$, $d(f \pm g) = df \pm dg$.

16.4. Теорема о неявной функции (общий случай)

Ранее (см. **11.1**) было введено общее понятие *неявной функции* (*неявно заданной функции*) как решения системы уравнений

$$\begin{cases} f_1(x_1, x_2, \ldots, x_{n+m}) = 0, \\ f_2(x_1, x_2, \ldots, x_{n+m}) = 0, \\ \ldots\ldots\ldots\ldots\ldots \\ f_m(x_1, x_2, \ldots, x_{n+m}) = 0, \end{cases} \tag{16.16}$$

где функции $f_1(x)$, $\ldots$, $f_m(x)$ определены в некоторой *области* $X \subset \mathbb{R}^{n+m}$. Был рассмотрен частный случай системы (16.16) при $m = 1$, т.е. случай одного уравнения, при котором использовался аппарат скалярных функций нескольких переменных. Остановимся на общем случае системы (16.16), используя аппарат *векторных функций нескольких переменных*.

При изучении системы (16.16) удобно использовать векторные способы записи. Подобную систему будем записывать в виде $F(x, y) = 0$, где $x \in \mathbb{R}^n$ объединяет переменные, значения которых задаются произвольно (свободные переменные), $y \in \mathbb{R}^m$ объединяет переменные, относительно которых решается система (зависимые переменные), а $F\colon \mathbb{R}^{n+m} \to \mathbb{R}^m$ — некоторая, вообще говоря, *векторная функция нескольких переменных*.

Поставим вопрос: при каких условиях система $F(x, y) = 0$ разрешима относительно переменных $y = (y_1,\ y_2,\ \ \ldots,\ y_m)$ в окрестности данной точки $(a,\ b) \in \mathbb{R}^{n+m}$? Через $F'_x(x, y) = \dfrac{\partial F(x,y)}{\partial x}$ и $F'_y(x, y) = \dfrac{\partial F(x,y)}{\partial y}$ будем обозначать соответственно матрицы Якоби функции F по части переменных x и по части переменных y, т.е.

$$F'_x(x,y) = \begin{pmatrix} \dfrac{\partial f_1(x,y)}{\partial x_1} & \dfrac{\partial f_1(x,y)}{\partial x_2} & \ldots & \dfrac{\partial f_1(x,y)}{\partial x_n} \\ \dfrac{\partial f_2(x,y)}{\partial x_1} & \dfrac{\partial f_2(x,y)}{\partial x_2} & \ldots & \dfrac{\partial f_2(x,y)}{\partial x_n} \\ \ldots & \ldots & \ldots & \ldots \\ \dfrac{\partial f_m(x,y)}{\partial x_1} & \dfrac{\partial f_m(x,y)}{\partial x_2} & \ldots & \dfrac{\partial f_m(x,y)}{\partial x_n} \end{pmatrix}$$

и

$$F'_y(x,y) = \begin{pmatrix} \dfrac{\partial f_1(x,y)}{\partial y_1} & \dfrac{\partial f_1(x,y)}{\partial y_2} & \ldots & \dfrac{\partial f_1(x,y)}{\partial y_m} \\ \dfrac{\partial f_2(x,y)}{\partial y_1} & \dfrac{\partial f_2(x,y)}{\partial y_2} & \ldots & \dfrac{\partial f_2(x,y)}{\partial y_m} \\ \ldots & \ldots & \ldots & \ldots \\ \dfrac{\partial f_m(x,y)}{\partial y_1} & \dfrac{\partial f_m(x,y)}{\partial y_2} & \ldots & \dfrac{\partial f_m(x,y)}{\partial y_m} \end{pmatrix}.$$

Отметим, что матрица $F'_y(x, y)$ является квадратной порядка m, а матрица Якоби $F'(x, y)$ по всей совокупности переменных может быть записана как блочная матрица $\bigl(F'_x(x, y)\ \ F'_y(x, y)\bigr)$. Отметим также, что определитель квадратной матрицы Якоби (по части переменных или по всем переменным — неважно) называют ***якобианом***.

Теорема 16.5 (*теорема о неявной функции (общий случай)*). Пусть система m уравнений $F(x, y) = 0$, $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, удовлетворяет следующим трем условиям:

1) координаты точки $(a,\ b) \in \mathbb{R}^{n+m}$ удовлетворяют системе уравнений, т.е. $F(a, b) = 0$;

2) функция $F(x, y)$ определена и непрерывно дифференцируема в некоторой окрестности V точки $(a,\ b)$, т.е. $F \in C^1(V)$;

3) матрица Якоби функции $F(x, y)$ в точке $(a,\ b)$ по части переменных y невырождена, т.е. $\det F'_y(a, b) \neq 0$.

Тогда в $\mathbb{R}^{n+m}$ найдется окрестность U точки $(a,\ b)$, определяемая неравенствами $|x - a| < \delta_x$, $|y - b| < \delta_y$, в которой система уравнений $F(x, y) = 0$ разрешима относительно группы переменных y и тем самым задает функцию $y = \varphi(x)$, $x \in U_x = \{x \in \mathbb{R}^n\colon |x - a| < \delta_x\}$. При этом функция $y = \varphi(x)$ непрерывно дифференцируема в *области* U_x, $\varphi(a) = b$, а ее матрица Якоби $\varphi'(x)$ может быть вычислена по формуле

$$\varphi'(x) = -\Big(\big(F'_y(x,y)\big)^{-1} F'_x(x,y)\Big)\Big|_{y=\varphi(x)}. \quad \# \tag{16.17}$$

Замечание 16.2. Как и в скалярном случае, теорема 16.5 не только дает формулу вычисления матрицы частных производных (матрицы Якоби) неявной функции, но и устанавливает достаточные условия существования неявной функции в окрестности заданной точки.

Формула (16.17) может быть получена по правилу дифференцирования сложной функции в предположении, что в дополнение к условиям теоремы 16.5 выполнено условие: неявная функция $y = \varphi(x)$, определяемая системой $F(x,y)$, дифференцируема в точке a. Действительно, в некоторой окрестности $U(a)$ точки a выполняется тождество $F(x, \varphi(x)) \equiv 0$. Дифференцируя это тождество в соответствии с правилом дифференцирования сложной функции, находим

$$F'_x(x,y)\Big|_{y=\varphi(x)} + F'_y(x,y)\Big|_{y=\varphi(x)} \varphi'(x) = 0.$$

Из этого матричного уравнения получаем

$$\varphi'(x) = -\Big(F'_y(x,y)\Big|_{y=\varphi(x)}\Big)^{-1} F'_x(x,y)\Big|_{y=\varphi(x)},$$

что равносильно (16.17).

Пример 16.2. Рассмотрим систему уравнений

$$\begin{cases} u^2 - 2xv + y^2 = 1, \\ v^2 - 2yu + x^2 = 0 \end{cases} \tag{16.18}$$

в окрестности точки $(x,\ y,\ u,\ v) = (0,\ 1,\ 0,\ 0)$. Векторная функция

$$F(x,y,u,v) = \begin{pmatrix} u^2 - 2xv + y^2 - 1 \\ v^2 - 2yu + x^2 \end{pmatrix}$$

является дифференцируемой в $\mathbb{R}^4$ и удовлетворяет условию $F(0,1,0,0) = (0\ \ 0)^{\mathrm{T}}$. Вычислим ее матрицу Якоби:

$$F'(x,y,u,v) = \begin{pmatrix} -2v & 2y & 2u & -2x \\ 2x & -2u & -2y & 2v \end{pmatrix}.$$

В точке $(0,\ 1,\ 0,\ 0)$ значение матрицы Якоби равно

$$F'(0,1,0,0) = \begin{pmatrix} 0 & 2 & 0 & 0 \\ 0 & 0 & -2 & 0 \end{pmatrix}.$$

Видно, что матрица Якоби $F'(0,1,0,0)$ имеет единственный ненулевой минор второго порядка, соответствующий переменным y и u:

$$F'_{(y,\ u)}(0,1,0,0) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}.$$

Следовательно, согласно теореме 16.5, в некоторой окрестности точки $(0,\ 1,\ 0,\ 0)$ систему уравнений (16.18) можно разрешить относительно переменных y и u, т.е. система определяет в окрестности точки $x = 0$, $v = 0$ функции $y = y(x,v)$, $u = u(x,v)$, для которых $F(x, y(x,v), u(x,v), v) \equiv (0\ \ 0)^{\mathrm{T}}$. Эти функции, согласно теореме 16.5, дифференцируемы, но можно показать, что на самом деле они дважды непрерывно дифференцируемы.

Найдем первый и второй дифференциалы функций $y(x, v)$ и $u(x, v)$ в точке $(0,\ 0)$. Дифференцируя уравнения системы, после сокращения на 2 находим

$$\begin{cases} u\,du - x\,dv - v\,dx + y\,dy = 0, \\ v\,dv - y\,du - u\,dy + x\,dx = 0. \end{cases} \tag{16.19}$$

Полученная система двух уравнений является линейной относительно дифференциалов переменных x, y, u, v. В точке $(0,\ 1,\ 0,\ 0)$ она приобретает особенно простой вид

$$\begin{cases} dy = 0, \\ du = 0. \end{cases}$$

Таким образом, функции $y(x, v)$ и $u(x, v)$ имеют нулевой дифференциал при $x = v = 0$.

Еще раз дифференцируем систему (16.19), учитывая, что dx и dv — это дифференциалы независимых переменных, а dy и du — это дифференциалы неявно заданных функций. В результате получаем

$$\begin{cases} du^2 + u\,d^2u - dxdv - dvdx + dy^2 + y\,d^2y = 0, \\ dv^2 - dydu - y\,d^2u - dudy - u\,d^2y + dx^2 = 0. \end{cases} \tag{16.20}$$

В точке $(0,\ 1,\ 0,\ 0)$ имеем $du = dy = 0$. Поэтому система (16.20) в этой точке принимает вид

$$\begin{cases} -2dxdv + d^2y = 0, \\ dv^2 - d^2u + dx^2 = 0. \end{cases}$$

В результате получаем вторые дифференциалы функций $y(x, v)$ и $u(x, v)$ в точке $(0,\ 0)$:

$$d^2y = 2dxdv, \quad d^2u = dx^2 + dv^2. \quad \#$$

Рассмотрим вопрос, при каких условиях *функция нескольких переменных* $G\colon \mathbb{R}^n \to \mathbb{R}^n$ имеет *обратную функцию* G^{-1}, а также вопрос о том, *дифференцируема* ли обратная *функция*. Соответствующие условия в *окрестности* фиксированной *точки* можно получить с помощью *теоремы о неявной функции*.

Теорема 16.6 (*теорема об обратной функции*). Пусть функция $G\colon \mathbb{R}^n \to \mathbb{R}^n$ удовлетворяет условиям:

1°. *Функция* $G(x)$ *непрерывно дифференцируема* в некоторой окрестности V точки a, т.е. $G \in C^1(V)$.

2°. *Матрица Якоби* функции $G(x)$ в точке a невырождена, т.е. $\det G'(a) \neq 0$.

Тогда найдется такая окрестность U точки $b = G(a)$, что:

1*. В U определена функция $G^{-1}(y)$, обратная к функции $G(x)$, т.е. $G^{-1}(y) \in V$ при $y \in U$ и $G\bigl(G^{-1}(y)\bigr) = y$, $y \in U$.

2*. Функция $G^{-1}(y)$ непрерывно дифференцируема в U (в частности, непрерывна в U), а ее матрица Якоби связана с матрицей Якоби функции $G(x)$ равенством

$$\bigl(G^{-1}\bigr)'(y) = \bigl(G'(x)\bigr)^{-1}\Bigr|_{x=G^{-1}(y)}. \tag{16.21}$$

◄ Рассмотрим функцию $F\colon \mathbb{R}^{2n} \to \mathbb{R}^n$, определяемую равенством $F(x, y) = G(x) - y$. Эта функция непрерывно дифференцируема в окрестности точки $(a,\ b) \in \mathbb{R}^{2n}$, а множество решений системы n уравнений $F(x, y) = 0$ представляет собой график функции $G(x)$, т.е. множество точек $(x,\ y)$, удовлетворяющих условию $y = G(x)$. В частности, $F(a, b) = 0$. Так как $\det G'(a) \neq 0$, то матрица Якоби $F'_x(a, b) = G'(a)$ невырождена. Таким образом, для функции $F(x,y)$ в окрестности точки $(a,\ b)$ выполнены условия *теоремы* 16.5 *о неявной функции.* Это значит, что система уравнений $F(x,y)=0$ в некоторой окрестности W вида $W = \{(x,\ y)\in\mathbb{R}^{2n}\colon\ |x-a|<\delta_x,\ \ |y-b|<\delta_y\}$ разрешима относительно переменных x, т.е. существует такая функция $\varphi(y)$, определенная в окрестности $|y-b|<\delta_y$ точки b, что

$$F(\varphi(y),y)\equiv 0, \tag{16.22}$$

причем функция $\varphi(y)$ непрерывно дифференцируема, а ее матрица Якоби равна

$$\varphi'(y) = -\bigl(F'_x(\varphi(y),y)\bigr)^{-1}F'_y(\varphi(y),y). \tag{16.23}$$

Так как $F(x,y)=G(x)-y$, тождество (16.22) означает, что $G\bigl(\varphi(y)\bigr)\equiv y$, т.е. функция $\varphi(y)$ является обратной к функции $G(x)$. Кроме того, матрица $F'_y(x,y)$ совпадает с матрицей $-E$, противоположной единичной матрице E. Поэтому равенство (16.23) сводится к равенству $\varphi'(y)=\bigl(G'(\varphi(y))\bigr)^{-1}$, равносильному (16.21) ►

Пример 16.3. а. Рассмотрим отображение $G\colon \mathbb{R}^2\to\mathbb{R}^2$, заданное уравнениями $z_1=x_1+e^{x_2}$, $z_2=e^{x_1}-x_2$. Это отображение непрерывно дифференцируемо всюду в $\mathbb{R}^2$. Его матрица Якоби в произвольной точке $(x_1,\ x_2)\in\mathbb{R}^2$ имеет вид

$$J(x_1,x_2)=\begin{pmatrix} 1 & e^{x_2}\\ e^{x_1} & -1\end{pmatrix},$$

а определитель матрицы Якоби

$$\det J(x_1,x_2) = -1-e^{x_1+x_2}$$

не обращается в нуль ни в одной точке в $\mathbb{R}^2$. Согласно теореме об обратной функции, в любой точке $b\in\mathbb{R}^2$, $b=G(a)$, существует окрестность, в которой определено обратное отображение G^{-1}, причем $G^{-1}(b)=a$.

б. Для отображения $G\colon \mathbb{R}^2\to\mathbb{R}^2$, заданного уравнениями

$$z_1=x_1+x_2^2,\qquad z_2=2x_1, \tag{16.24}$$

найдем те точки множества в области значений отображения, в окрестности которых определено обратное отображение G^{-1}. Для это воспользуемся теоремой об обратной функции. Отображение G непрерывно дифференцируемо в $\mathbb{R}^2$, а его матрица Якоби имеет вид

$$J(x_1,x_2)=\begin{pmatrix} 1 & 2x_2\\ 2 & 0\end{pmatrix}.$$

Вычисляем определитель матрицы Якоби: $\det J(x_1,x_2)=-4x_2$. Отсюда заключаем, что матрица Якоби невырождена во всех точках $(x_1,\ x_2)$, для которых $x_2\neq 0$. Таким образом, во всех точках $(x_1,\ x_2)$, удовлетворяющих условию $x_2\neq 0$, можно применить теорему об обратной функции. Точки $(x_1,\ x_2)$, в которых матрица Якоби вырождена, удовлетворяют условию $x_2=0$ и в совокупности составляют прямую — координатную ось $O\mathrm{x}_1$. Найдем ее образ при отображении G. Для этого в уравнения (16.24) отображения G подставим $x_2=0$. В результате находим образ координатной оси $O\mathrm{x}_1$:

$$\{(z_1,\ z_2)\colon\ z_1=x_1,\ z_2=2x_1\},$$

или $z_2=2z_1$.

Итак, обратное отображение G^{-1} определено в окрестности любой точки $(z_1,\ z_2)$, принадлежащей области значений отображения G и не лежащей на прямой $z_2=2z_1$. Теорема об обратной функции не позволяет ответить на вопрос, существует ли обратное отображение G^{-1} в окрестности какой-либо точки прямой $z_2 = 2z_1$. Для ответа на этот вопрос нужно использовать другие методы. В данном случае уравнения (16.24) можно разрешить относительно переменных x_1 и x_2 и тем самым получить аналитическое представление функции G^{-1}:

$$\begin{cases} x_1 = \dfrac{z_2}{2}, \\ x_2 = \pm\sqrt{z_1 - \dfrac{z_2}{2}}. \end{cases}$$

Это представление показывает, что областью значений отображения G является полуплоскость $z_1 \geqslant z_2/2$. Каждая *внутренняя точка* $z = (z_1,\, z_2)$ этой полуплоскости является образом при отображении G двух точек $a = (a_1,\, a_2)$ и $\widetilde{a} = (a_1,\, -a_2)$, отличающихся знаком второй координаты. В окрестности каждой точки $z = (z_1,\, z_2)$, $z_1 > z_2/2$, существуют два обратных отображения, первое удовлетворяет условию $G^{-1}(z) = a$, а второе — условию $G^{-1}(z) = \widetilde{a}$. Оба отображения определены в области $z_1 > z_2/2$.

Любая точка $z^0 = (z_1^0,\, z_2^0)$ на прямой $z_1 = z_2/2$ не имеет окрестности, в которой определено обратное отображение G^{-1}, и тому есть две причины. Во-первых, такие точки не являются внутренними точками области значений отображения G. Во-вторых, каждая такая точка z^0 является образом единственной точки x^0 в области определения отображения, но при этом в любой окрестности точки x^0 можно выбрать такие две точки, в которых отображение G принимает одинаковые значения.

ОГЛАВЛЕНИЕ