

МОСКОВСКИЙ ФИЗИКО-ТЕХНИЧЕСКИЙ ИНСТИТУТ (НАЦИОНАЛЬНЫЙ
ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ)

Физтех-школа прикладной математики и информатики

Заметки к лекциям по многомерному анализу интегралам и рядам.

Лектор:
Гусев Н. А.

Содержание

МА-28. Евклидова топология в \mathbb{R}^d	3
МА-30. Компакты в \mathbb{R}^d	10
МА-31. Непрерывность функций. Связность и линейная связность	12
МА-32. Теорема Бэра на вещественной прямой	18
МА-33. Определённый интеграл Римана	23
МА-34. Определённый интеграл Римана (продолжение)	30
МА-35. Несобственный интеграл Римана	37
МА-36. Знакопостоянные числовые ряды	44
МА-37. Знакопеременные числовые ряды	49
МА-38. Поточечная и равномерная сходимость функциональных последовательностей	55
МА-39. Исследование функциональных последовательностей и рядов на равномерную сходимость	60
МА-40. Признаки Дирихле и Абеля для функциональных рядов	63
МА-41. Степенные ряды	67
МА-42. Вещественно-аналитические функции	71
МА-43. Ряд Тейлора	75
МА-44. Частные производные, производные по направлениям и дифференцируемость функций многих переменных	82
МА-45. Дифференцирование композиции отображений. Теорема о неявной функции.	88

МА-46. Теорема о системе функций, заданных неявно	93
МА-47. Смешанные частные производные и дифференциалы высших порядков	97
МА-48. (Безусловные) экстремумы функций многих переменных	104
МА-49. Теорема об обратном отображении. Криволинейные координаты	106
МА-50. Условные экстремумы функций многих переменных	108
Примечания	118

МА-28. Евклидова топология в \mathbb{R}^d

Окрестности точек и сходимость последовательностей

Напомним, что ε -окрестностью точки $x = (x(1), x(2), \dots, x(d)) \in \mathbb{R}^d$ называется множество

$$U_\varepsilon(x) := \{y \in \mathbb{R}^d : |y - x| < \varepsilon\}$$

где

$$|y - x| = \sqrt{\sum_{k=1}^d |y(k) - x(k)|^2}$$

В случае $d = 1$ имеем $U_\varepsilon(x) = (x - \varepsilon, x + \varepsilon)$, а при $d > 1$ множество $U_\varepsilon(x)$ представляет собой шар радиуса ε с центром в точке x (при этом сфера радиуса ε с центром в x в этот шар не включается).

Проколотой окрестностью точки $x \in \mathbb{R}^d$ называется множество $\overset{\circ}{U}_\varepsilon(x) := U_\varepsilon(x) \setminus \{x\}$

Дополнением множества $E \subset \mathbb{R}^d$ называется множество $E^c := \mathbb{R}^d \setminus E$.

Сходимость последовательностей точек в \mathbb{R}^d определяется так же, как в одномерном случае. А именно, последовательность $\{x_n\} \subset \mathbb{R}^d$ сходится к $x \in \mathbb{R}^d$ при $n \rightarrow \infty$, если для любого $\varepsilon > 0$ существует $N \in \mathbb{N}$ такое, что для всех $n > N$ выполнено $x_n \in U_\varepsilon(x)$

Утверждение 1: Последовательность $\{x_n\} \subset \mathbb{R}^d$ сходится к $x \in \mathbb{R}^d$ при $n \rightarrow \infty$ тогда и только тогда, когда

$$|x - x_n| \rightarrow 0$$

при $n \rightarrow \infty$

Доказательство:

Утверждение очевидно, так как по определению условие $x_n \in U_\varepsilon(x)$ равносильно тому, что $|x - x_n| < \varepsilon$ покомпонентной сходимости (сходимости координат):

Утверждение 2: Последовательность $\{x_n\} \subset \mathbb{R}^d$ сходится к $x \in \mathbb{R}^d$ при $n \rightarrow \infty$ тогда и только тогда, когда для каждого $k \in \overline{1, d}$ выполнено

$$x_n(k) \rightarrow x(k)$$

при $n \rightarrow \infty$

Доказательство:

Необходимость вытекает из того, что для каждого $k \in \overline{1, d}$ выполнено

$$|x_n(k) - x(k)| \leq |x_n - x|$$

Достаточность следует из того, что

$$|x_n - x| \leq \sqrt{d} \max_{1 \leq k \leq d} |x_n(k) - x(k)|$$

Внутренность и замыкание множества

Определение 3: Пусть $E \subset \mathbb{R}^d$ и $x \in \mathbb{R}^d$.

1. Если существует $\varepsilon > 0$ такое, что $U_\varepsilon(x) \subset E$, то x называется внутренней точкой множества E .
 - Множество всех внутренних точек множества E называется внутренностью E и обозначается $\text{int } E$.
2. Если для любого $\varepsilon > 0$ выполнено $U_\varepsilon(x) \cap E \neq \emptyset$, то x называется точкой прикосновения множества E .
 - Множество всех точек прикосновения множества E называется замыканием множества E и обозначается \bar{E} .

По определению для любого множества E выполнено

$$\text{int } E \subset E \subset \bar{E}$$

Например, для множества $E = [0,1) \cup \{2\}$ выполнено:

$$\text{int } E = (0,1), \quad \bar{E} = [0,1] \cup \{2\}$$

Утверждение 4: Пусть $E \subset \mathbb{R}^d$ и $x \in \mathbb{R}^d$.

1. Точка x является внутренней точкой множества E тогда и только тогда, когда для любой последовательности $\{x_n\}$ точек из \mathbb{R}^d , сходящейся к x , существует $N \in \mathbb{N}$ такое, что для всех $n > N$ выполнено $x_n \in E$.
2. Точка x является точкой прикосновения множества E тогда и только тогда, когда существует последовательность $\{x_n\}$ точек из E , сходящаяся к x при $n \rightarrow \infty$.

Доказательство:

1. Необходимость следует из определения внутренней точки. Обратно, если точка x не является внутренней, то для каждого $n \in \mathbb{N}$ в $1/n$ -окрестности x найдётся точка x_n , не принадлежащая E . Тогда мы получаем последовательность точек из дополнения E , сходящуюся к x , что противоречит предположенному свойству.
2. Достаточность следует из определения точки прикосновения. Обратно, если точка x является точкой прикосновения множества E , то для каждого $n \in \mathbb{N}$ существует точка $x_n \in E$, принадлежащая $1/n$ -окрестности точки x . По построению $x_n \rightarrow x$ при $n \rightarrow \infty$.

Открытые и замкнутые множества

Определение 5: Множество $E \subset \mathbb{R}^d$ называется

- открытым, если $E = \text{int } E$
- замкнутым, если $E = \bar{E}$.

Например, для любых $x \in \mathbb{R}^d$ и $\varepsilon > 0$ множество $U_\varepsilon(x)$ открыто. Действительно, для любой точки $y \in U_\varepsilon(x)$ существует $\delta := \varepsilon - |x - y|$ такое, что $U_\delta(y) \subset U_\varepsilon(x)$. Аналогично, множество $P_\varepsilon(x) := \{y \in \mathbb{R}^d : |y - x| \leq \varepsilon\}$ замкнуто.

В силу (1) можно сказать, что множество открыто тогда и только тогда, когда все его точки - внутренние. Аналогично, множество замкнуто тогда и только тогда, когда содержит все свои точки прикосновения. Сразу из определения следует, что внутренность любого множества представляет собой открытое множество, а замыкание любого множества представляет собой замкнутое множество.

$$\bar{E} = (\text{int } E^c)^c$$

Доказательство:

Точка $x \in \mathbb{R}^d$ не является точкой прикосновения множества E тогда и только тогда, когда существует $\varepsilon > 0$ такое, что $U_\varepsilon(x) \cap E = \emptyset$. Последнее равносильно тому, что $U_\varepsilon(x) \subset E^c$, то есть $x \in \text{int } E^c$. Значит $(\bar{E})^c = \text{int } E^c$. Взяв дополнение, получим (2).

Граница множества. Предельные и изолированные точки

Определение 7: Пусть $E \subset \mathbb{R}^d$, $x \in \mathbb{R}^d$.

1. Если для любого $\varepsilon > 0$ выполнено $U_\varepsilon(x) \cap E \neq \emptyset$ и $U_\varepsilon(x) \cap E^c \neq \emptyset$, то x называется граничной точкой множества E .
2. Множество всех граничных точек множества E называется границей E и обозначается ∂E .
3. Если для любого $\varepsilon > 0$ выполнено $\overset{\circ}{U}_\varepsilon(x) \cap E \neq \emptyset$, то x называется предельной точкой множества E .
4. Если $x \in E$ и существует $\varepsilon > 0$ такое, что $\overset{\circ}{U}_\varepsilon(x) \cap E = \emptyset$, то x называется изолированной точкой множества E .

Например, если $d = 1$ и $E = [0, 1) \cup \{2\}$, то $[0, 1]$ является множеством всех предельных точек множества E , а точка 2 является единственной изолированной точкой множества E .

Теорема 8: Для любого $E \subset \mathbb{R}^d$ выполнено

$$\partial E = (\text{int } E \sqcup \text{int } E^c)^c$$

Кроме того,

$$\bar{E} = \text{int } E \sqcup \partial E = E \cup \partial E$$

Доказательство:

В силу (1) имеем $\text{int } E \subset E$ и $\text{int } E^c \subset E^c$, поэтому множества $\text{int } E$ и $\text{int } E^c$ не пересекаются. Точка $x \in \mathbb{R}^d$ не принадлежит $\text{int } E \sqcup \text{int } E^c$ тогда и только тогда, когда в любой её ε -окрестности найдётся хотя бы одна точка из E и хотя бы одна точка из E^c . Это по определению границы и означает, что

$x \in \partial E$. Заметим, что $\partial E \subset \bar{E}$, так как по определению в любой ε -окрестности точки границы E содержится хотя бы одна точка множества E . При этом точка x принадлежит $\bar{E} \setminus \partial E$ тогда и только тогда, когда некоторая её окрестность лежит во множестве E . Значит $\bar{E} \setminus \partial E = \text{int } E$

Наконец, $E \subset \bar{E}$ и $\partial E \subset \bar{E}$, поэтому

$$\bar{E} = \text{int } E \sqcup \partial E \subset E \cup \partial E \subset \bar{E}$$

Отсюда видно, что $\bar{E} = E \cup \partial E$.

Таким образом, для любого множества $E \subset \mathbb{R}^d$ множество \mathbb{R}^d разбивается на 3 попарно непересекающихся подмножества: $\text{int } E$, ∂E и $\text{int } E^c$. Например, если $d = 1$ и $E = (-\infty, 0) \cup (\mathbb{Q} \cap [0, 1])$, то это разбиение имеет вид

$$\mathbb{R} = \underbrace{(-\infty, 0)}_{\text{int } E} \cup \underbrace{(\mathbb{Q} \cap [0, 1])}_{\partial E} \cup \underbrace{(\mathbb{I} \cap [0, 1]) \cup (1, +\infty)}_{\text{int } E^c}$$

Теоретико-множественные свойства открытых и замкнутых множеств

Обсудим свойства открытых и замкнутых множеств, связанные с операциями объединения, пересечения и взятия дополнения. Прежде всего отметим, что понятие замкнутости множества можно было бы эквивалентно определить иначе:

Теорема 9: Множество E замкнуто тогда и только тогда, когда его дополнение открыто.

Доказательство:

В силу (2) равенство $E = \bar{E}$ равносильно тому, что $E = (\text{int } E^c)^c$, то есть $E^c = \text{int } E^c$. Последнее равносильно открытости E^c .

Теорема 10: Объединение любого семейства \mathcal{G} открытых множеств открыто. Если, кроме того, семейство \mathcal{G} конечно, то и пересечение \mathcal{G} открыто.

Доказательство:

Пусть $E := \bigcup \mathcal{G}$. Напомним, по определению объединения семейства множеств точка x принадлежит множеству E тогда и только тогда, когда она принадлежит некоторому элементу G семейства \mathcal{G} . Так как множество G открыто, то существует $\varepsilon > 0$ такое, что $U_\varepsilon(x) \subset G$. Но $G \subset E$, а значит $x \in \text{int } E$. Следовательно, E открыто.

Теперь предположим, что $\mathcal{G} = \{G_1, \dots, G_n\}$ (для некоторого $n \in \mathbb{N}$). Пусть $P = \bigcap G = G_1 \cap \dots \cap G_n$ и $x \in P$. Тогда для каждого $k \in \overline{1, n}$ точка x принадлежит открытому множеству G_k . Значит найдётся $\varepsilon_k > 0$ такое, что $U_{\varepsilon_k}(x) \subset G_k$. Взяв $\varepsilon := \min(\varepsilon_1, \dots, \varepsilon_n)$ получим $U_\varepsilon(x) \subset G_k$ для каждого $k \in \overline{1, n}$. Значит $U_\varepsilon(x) \subset P$. Следовательно P открыто.

Иначе говоря, если для каждого элемента a некоторого множества A дано открытое множество G_a , то множество $\bigcup_{a \in A} G_a$ открыто. Даже если множество индексов A несчётно. Например, если $A = [0, 1]$ и $G_a = (a, a + 1)$ для всех $a \in A$, то $\bigcup_{a \in A} G_a = (0, 2)$ действительно открыто.

Если множество индексов A конечно, то и множество $\bigcap_{a \in A} G_a$ открыто. Но если A бесконечно, то множество $\bigcap_{a \in A} G_a$ не обязано быть открытым. Например, если $A = \mathbb{N}$ и $G_a = (-1/a, 1/a)$ для любого $a \in A$, то $\bigcap_{a \in A} G_a = \{0\}$ не является открытым.

Утверждение 11: Для любого множества E его внутренность является наибольшим по вложению открытым подмножеством множества E . Кроме того, $\text{int } E$ совпадает с объединением всех открытых подмножеств множества E .

Доказательство:

Пусть M - объединение семейства

$$\mathcal{G} := \{G \subset \mathbb{R}^d : G \text{ открыто, } G \subset E\}$$

Данное семейство состоит из открытых множеств, значит по теореме 10 множество M открыто. Кроме того, для любого $G \in \mathcal{G}$ выполнено $G \subset E$, поэтому $M \subset E$. Таким образом, множество M принадлежит семейству \mathcal{G} .

По определению объединения для любого $G \in \mathcal{G}$ выполнено $G \subset M$. Значит M является наибольшим по вложению открытым подмножеством множества E . (Отметим, что другого такого открытого множества M' не существует, поскольку в силу минимальности по вложению мы имели бы $M \subset M'$ и $M' \subset M$, откуда $M = M'$.)

Все точки каждого $G \in \mathcal{G}$ являются внутренними точками E (так как G открыто и $G \subset E$), а значит $G \subset \text{int } E$. Тогда по определению объединения $M \subset \text{int } E$. С другой стороны, ясно что $\text{int } E$ является одним из элементов семейства \mathcal{G} . Тогда по определению объединения $\text{int } E \subset M$. Из полученных вложений заключаем, что $M = \text{int } E$. Следствие 12. Для любого множества E его замыкание является наименьшим по вложению замкнутым множеством, подмножеством которого является E . Кроме того, \bar{E} совпадает с пересечением всех замкнутых множеств, включающих в себя E .

Доказательство:

В силу теоремы 9 замкнутое множество F содержит E тогда и только тогда, когда открытое множество F^c лежит в E^c . Иначе говоря, множество F принадлежит семейству

$$\mathcal{F} := \{F \subset \mathbb{R}^d : F \text{ замкнуто, } E \subset F\}$$

тогда и только тогда, когда его дополнение принадлежит семейству

$$\mathcal{G} := \{G \subset \mathbb{R}^d : G \text{ открыто, } G \subset E^c\}$$

Тогда

$$\bigcap_{F \in \mathcal{F}} F = \bigcap_{F^c \in \mathcal{G}} (F^c)^c = \bigcap_{G \in \mathcal{G}} G^c = \left(\bigcup_{G \in \mathcal{G}} G \right)^c = (\text{int } E^c)^c = \bar{E}.$$

Отсюда по определению пересечения имеем $\bar{E} \subset F$ для любого $F \in \mathcal{F}$. С другой стороны, \bar{E} замкнуто и $E \subset \bar{E}$, значит $\bar{E} \in \mathcal{F}$. Таким образом, \bar{E} является наименьшим по вложению элементом семейства \mathcal{F} .

Отделимость замкнутых множеств

Открытое множество $G \subset \mathbb{R}^d$ называется (открытой) окрестностью точки $x \in \mathbb{R}^d$, если $x \in G$. Аналогично, открытое множество $G \subset \mathbb{R}^d$ называется (открытой) окрестностью множества $E \subset \mathbb{R}^d$, если $E \subset G$.

Легко видеть, что у любых двух различных точек $x, y \in \mathbb{R}^d$ имеются непересекающиеся окрестности. Действительно, если $|x - y| = \varepsilon > 0$, то $G_1 = U_{\varepsilon/2}(x)$ и $G_2 = U_{\varepsilon/2}(y)$ открыты и не пересекаются, и при этом $x \in G_1$ и $y \in G_2$.

Расстоянием от точки x до непустого множества $E \subset \mathbb{R}^d$ называется число

$$\rho(x, E) := \inf_{y \in E} |x - y|$$

Аналогично, расстоянием между множествами $E, F \subset \mathbb{R}^d$ называется число

$$\rho(E, F) := \inf_{x \in E, y \in F} |x - y|.$$

Если $\rho(E, F) > 0$, то непересекающиеся окрестности множеств E и F можно построить аналогично тому, как это делалось для двух точек. Оказывается, что непересекающиеся окрестности можно построить для любых двух непересекающихся замкнутых множеств (даже если расстояние между ними равно нулю):

Утверждение 13: Если замкнутые множества F_1 и F_2 не пересекаются, то существуют непересекающиеся открытые множества G_1 и G_2 такие, что $F_1 \subset G_1$ и $F_2 \subset G_2$.

Доказательство:

Действительно, так как $F_1 \subset F_2^c$, то для каждого $x \in F_1$ существует $\varepsilon = \varepsilon(x) > 0$ такое, что $U_{\varepsilon(x)}(x) \subset F_2^c$. Аналогично, для каждого $y \in F_2$ существует $\delta = \delta(y) > 0$ такое, что $U_{\delta(y)}(y) \subset F_1^c$. Возьмём $G_1 := \bigcup_{x \in F_1} U_{\varepsilon(x)/2}(x)$ и $G_2 := \bigcup_{y \in F_2} U_{\delta(y)/2}(y)$. Ясно, что G_1 и G_2 открыты. Если существует $z \in G_1 \cap G_2$, то существуют $x \in F_1$ и $y \in F_2$ такие, что $z \in U_{\varepsilon(x)/2}(x) \cap U_{\delta(y)/2}(y)$. Тогда

$$|x - y| \leq |x - z| + |z - y| < (\varepsilon(x) + \delta(y))/2 < \max(\varepsilon(x), \delta(y)).$$

Значит либо $x \in U_{\delta(y)}(y)$, либо $y \in U_{\varepsilon(x)}(x)$, что противоречит определению ε и δ .

Структура открытых подмножеств вещественной прямой

Ясно, что любое объединение интервалов представляет собой открытое множество. Верно и обратное: любое открытое множество $E \subset \mathbb{R}$ можно представить в виде объединения интервалов. Действительно, по определению открытого множества для любого $x \in E$ существует $\varepsilon = \varepsilon(x) > 0$ такое, что $U_{\varepsilon(x)}(x) \subset E$. Тогда

$$E = \bigcup_{x \in E} \{x\} \subset \bigcup_{x \in E} U_{\varepsilon(x)}(x) \subset E$$

значит $E = \bigcup_{x \in E} U_{\varepsilon(x)}(x)$. Таким образом, E представлено в виде объединения некоторого семейства открытых интервалов, в общем случае несчётного. Оказывается, что всегда можно ограничиться счётным семейством, а сами интервалы выбирать попарно непересекающимися:

Теорема 14: Любое открытое множество $E \subset \mathbb{R}$ можно представить в виде объединения не более чем счётного семейства попарно непересекающихся интервалов.

Доказательство:

Для каждой точки $x \in E$ выберем максимальный по вложению интервал I_x такой, что $x \in I_x \subset E$. Для этого рассмотрим семейство $\mathcal{F}_x := \{(a, b) : a, b \in \mathbb{R}, x \in (a, b) \subset E\}$ и возьмём $I_x := \bigcup \mathcal{F}_x$. Множество I_x открыто (как объединение открытых множеств). Кроме того, I_x является непустым

промежутком, так как для любого $\xi \in I_x$ отрезок $[x, \xi]$ (или $[\xi, x]$ если $\xi < x$) лежит в некотором интервале из \mathcal{F}_x и, следовательно, лежит в I_x . Тогда $I_x \in \mathcal{F}_x$. Так как для любого $J \in \mathcal{F}_x$ выполнено $J \subset I_x$, то I_x является наибольшим по вложению элементом \mathcal{F}_x .

Заметим, что если $x, y \in E$, то либо $I_x \cap I_y = \emptyset$, либо $I_x = I_y$. Действительно, если $I_x \cap I_y \neq \emptyset$, то $I := I_x \cup I_y$ является интервалом, лежит в E и содержит точки x и y . Последнее возможно лишь в случае, когда $I_x = I_x \cup I_y = I_y$ (иначе получаем противоречие максимальнойности интервалов I_x и I_y).

Рассмотрим семейство $\mathcal{G} := \{I_x : x \in E\}$. Так как \mathcal{G} является дизъюнктивным семейством непустых интервалов, то \mathcal{G} не более чем счётно.

Наконец, для любого $x \in E$ выполнено $I_x \subset E$, поэтому $\bigcup \mathcal{G} \subset E$. С другой стороны, $E = \bigcup_{x \in E} \{x\} \subset \bigcup_{x \in E} I_x = \bigcup \mathcal{G}$. Значит $E = \bigcup \mathcal{G}$, что и требовалось доказать.

Следствие 15. Множества \emptyset и \mathbb{R} являются единственными множествами, которые открыты и замкнуты одновременно.

Доказательство:

Предположим, что $E \subset \mathbb{R}$ непусто, открыто и замкнуто одновременно. Тогда E^c открыто. Возьмём точку $x \in E$ и выберем для неё максимальный по вложению интервал I_x такой, что $x \in I_x \subset E$. Обозначим $\alpha := \inf I_x, \beta := \sup I_x$. Если $\beta \in \mathbb{R}$, то $\beta \in E^c$. Но тогда существует $\delta > 0$ такое, что $(\beta - \delta, \beta + \delta) \subset E^c$. Это противоречит тому, что $(\alpha, \beta) \subset E$. Таким образом, $\beta = +\infty$ и из аналогичных соображений $\alpha = -\infty$.

МА-30. Компакты в \mathbb{R}^d

Множество $E \subset \mathbb{R}^d$ называется ограниченным, если оно лежит в некотором шаре. Это эквивалентно существованию константы $C > 0$ такой, что $|x| \leq C$ для всех $x \in E$.

Определение 1: Множество $K \subset \mathbb{R}^d$ называется компактом, если оно ограничено и замкнуто.

Теорема 2: (свойство Больцано-Вейерштрасса). Множество $K \subset \mathbb{R}^d$ является компактом тогда и только тогда, когда из любой последовательности точек x_n множества K можно выделить подпоследовательность, сходящуюся к некоторой точке множества K .

Доказательство:

Необходимость. Пусть $x_n \in K$ для всех $n \in \mathbb{N}$. Так как K ограничено, то числовая последовательность $\{x_n(1)\}_{n \in \mathbb{N}}$ также ограничена. По теореме Больцано-Вейерштрасса из неё можно выделить сходящуюся подпоследовательность $x_{n_k}(1)$. Аналогично, из ограниченной числовой последовательности $x_{n_k}(2)$ можно выделить сходящуюся подпоследовательность $x_{n_{k_\ell}}(2)$. Переходя к следующим компонентам и выбирая сходящуюся подпоследовательность получим строго возрастающую последовательность индексов $\{m_k\}_k \subset \mathbb{N}$ такую, что для каждого $j \in \overline{1, d}$ последовательность $\{x_{m_k}(j)\}$ сходится при $k \rightarrow \infty$. Следовательно, последовательность $\{x_{m_k}\}$ сходится к некоторой точке $x \in \mathbb{R}^d$ при $k \rightarrow \infty$. Так как K замкнуто, то $x \in K$.

Достаточность. Пусть K - компакт. Если K неограничено, то можно выбрать последовательность точек $x_n \in K$ так, что $|x_n| \rightarrow \infty$ при $n \rightarrow \infty$. Из такой последовательности, очевидно, нельзя выбрать подпоследовательность, сходящуюся к некоторой точке $x \in \mathbb{R}^d$. Противоречие. Значит K ограничено.

Пусть последовательность точек $\{x_n\}_{n \in \mathbb{N}}$ лежит в K и сходится к $x \in \mathbb{R}^d$ при $n \rightarrow \infty$. По определению компакта мы можем выделить подпоследовательность $\{x_{n_k}\}_{k \in \mathbb{N}}$, сходящуюся к некоторой точке множества K . Но $x_{n_k} \rightarrow x$ при $k \rightarrow \infty$. Поэтому $x \in K$. Значит K замкнуто.

Определение 3: Семейство \mathcal{F} называется покрытием множества E , если для каждого $x \in E$ существует множество $A \in \mathcal{F}$ такое, что $x \in A$. Семейство $\mathcal{G} \subset \mathcal{F}$ называется подпокрытием множества E , если \mathcal{G} также является покрытием множества E .

Покрытие \mathcal{F} множества $E \subset \mathbb{R}^d$ называется открытым, если \mathcal{F} состоит из открытых множеств.

Например, семейство $\mathcal{F} = \{(a-1, a+1) : a \in \mathbb{I}\}$ является (несчётным) открытым покрытием отрезка $E = [-1, 1]$. Семейство $\mathcal{G} = \left\{ \left(-\frac{1}{\sqrt{2}} - 1, -\frac{1}{\sqrt{2}} + 1 \right), \left(\frac{1}{\sqrt{2}} - 1, \frac{1}{\sqrt{2}} + 1 \right) \right\} \subset \mathcal{F}$ является (конечным) подпокрытием отрезка E .

Теорема 4: (лемма Гейне-Бореля). Множество $K \subset \mathbb{R}^d$ является компактом тогда и только тогда, когда из любого открытого покрытия множества K можно выделить конечное подпокрытие.

Доказательство:

Достаточность. Рассмотрим открытое покрытие $\mathcal{F} = \{ \{x \in \mathbb{R}^d : |x| < r\} : r > 0 \}$ множества K . Так как из \mathcal{F} можно выделить конечное подпокрытие, то K ограничено.

Пусть последовательность точек $\{x_n\}_{n \in \mathbb{N}} \subset K$ сходится к точке $x \in \mathbb{R}^d$. Если $x \notin K$, то семейство $\mathcal{F} = \{ \{y \in \mathbb{R}^d : |x - y| > r\} : r > 0 \}$ покрывает K и из него можно выделить конечное подпокрытие. Тогда для некоторого $r > 0$ выполнено $K \subset \{y \in \mathbb{R}^d : |x - y| > r\}$. Значит для всех $n \in \mathbb{N}$ выполнено $|x - x_n| > r$, что противоречит сходимости $x_n \rightarrow x$ при $n \rightarrow \infty$.

Необходимость. Пусть K - компакт и $\mathcal{F} = \{U_a\}_{a \in A}$ - его открытое покрытие, из которого нель-

зя выделить конечное подпокрытие. Так как K ограничено, то оно лежит в некотором d -мерном параллелепипеде P .

Разделив каждую сторону P пополам, получим 2^d параллелепипедов P_i . Если для каждого i множество $K \cap P_i$ покрывается конечным набором элементов \mathcal{F} , то и всё множество K можно покрыть конечным набором элементов \mathcal{F} , что противоречит сделанному предположению. Значит найдётся параллелепипед Q_1 такой, что $K \cap Q_1$ не покрывается никаким конечным набором элементов \mathcal{F} (в частности, $K \cap Q_1 \neq \emptyset$).

Теперь в качестве P можно взять Q_1 и повторить сделанный шаг. Продолжая данный процесс, в итоге мы получим последовательность стягивающихся вложенных параллелепипедов $\{Q_j\}_{j \in \mathbb{N}}$ таких, что $K \cap Q_j$ не покрывается никаким конечным набором элементов \mathcal{F} . По теореме Кантора (строго говоря, её мы применяем к проекциям рассматриваемых параллелепипедов) существует единственная точка $x \in \bigcap_{j=1}^{\infty} Q_j$. Так как $K \cap Q_j \neq \emptyset$, то $x \in \bar{K}$. В силу замкнутости K заключаем, что $x \in K$.

По определению покрытия существует $U \in \mathcal{F}$ такое, что $x \in U$. Но тогда при достаточно больших j выполнено $Q_j \subset U$. (Действительно, так как $x \in U$, то существует $\varepsilon > 0$ такое, что $U_\varepsilon(x) \subset U$. Тогда для того, чтобы $Q_j \subset U$ достаточно, чтобы каждая из сторон Q_j была меньше ε/\sqrt{d} .)

В силу теорем 2 и 4 следующие 3 свойства множества $K \subset \mathbb{R}^d$ эквивалентны:

1. K ограничено и замкнуто.
2. из любой последовательности элементов множества K можно выделить подпоследовательность, сходящуюся к некоторому элементу множества K .
3. из любого открытого покрытия множества K можно выделить конечное подпокрытие.

Забегая вперёд отметим, что эти три свойства эквивалентны только в случае евклидовой топологии. В общем случае за определение компактности берётся третье свойство, а второе свойство называется секвенциальной компактностью.

Множество Кантора

Простейшим примером компакта в \mathbb{R} является отрезок $[0,1]$. Другой примера компакта - множество вида $\{\frac{1}{n} : n \in \mathbb{N}\} \cup \{0\}$. Может ли компакт быть несчётным и при этом не содержать ни одного отрезка положительной длины? Оказывается, что да. Для построения такого компакта C начнём со множества $C_0 := [0,1]$. Исключив из C_0 интервал $(\frac{1}{3}, \frac{2}{3})$ получим множество $C_1 := [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$. В каждом из двух отрезков, составляющих C_1 , исключим интервал с центром в середине и длиной $1/9$. Продолжая по аналогии на n -ом шаге получим множество C_n , состоящее из 2^n попарно непересекающихся отрезков длины 3^{-n} .

Определение 5: Множество $C := \bigcap_{n=1}^{\infty} C_n$ называется (тонким) множеством Кантора.

Утверждение 6: Множество Кантора C замкнуто, равномощно \mathbb{R} и не включает в себя ни одного отрезка ненулевой длины. ненулевой длины, следует из того, что в C не входят числа вида $m/2^n$, где $m, n \in \mathbb{N}$, а в любом отрезке ненулевой длины содержится хотя бы одно число такого вида.

МА-31. Непрерывность функций. Связность и линейная СВЯЗНОСТЬ

Пределы по направлениям и по совокупности переменных.

Открытое множество $U \subset \mathbb{R}^n$ называется (открытой) окрестностью точки $x \in \mathbb{R}^n$, если $x \in U$.

Рассмотрим отображение $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$. По аналогии с одномерным случаем точка $y \in \mathbb{R}^m$ называется пределом f в точке $x \in \mathbb{R}^n$ (по совокупности переменных) если для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что $f(\overset{\circ}{U}_\delta(x)) \subset U_\varepsilon(y)$. Обозначение: $y = \lim_{z \rightarrow x} f(z)$.

Если отображение f определено не на всём \mathbb{R}^n , а лишь на некотором подмножестве $E \subset \mathbb{R}^n$, для которого x является предельной точкой, то условие $f(\overset{\circ}{U}_\delta(x)) \subset U_\varepsilon(y)$ необходимо заменить на $f(E \cap \overset{\circ}{U}_\delta(x)) \subset U_\varepsilon(y)$; в этом случае говорят о пределе по множеству E . Чтобы это подчеркнуть, иногда используют обозначение $\lim_{E \ni z \rightarrow x} f(z)$.

Если $a \in \mathbb{R}^n$ - ненулевой вектор, то предел $\lim_{t \rightarrow 0} f(x+at)$, если он существует, называется пределом f по направлению a . При $t \rightarrow \pm 0$ говорят об односторонних пределах по направлению.

Утверждение 1. Если отображение f имеет в точке $x \in \mathbb{R}^n$ предел $y \in \mathbb{R}^m$ (по совокупности переменных), то для любого ненулевого вектора $a \in \mathbb{R}^n$ отображение f имеет предел по направлению a , равный y .

Доказательство:

По определению предела для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что $f(\overset{\circ}{U}_\delta(x)) \subset U_\varepsilon(y)$. Если $|t| < \delta/|a|$, то $x+at \in U_\delta(x)$. Тогда $f(x+at) \in U_\varepsilon(y)$.

Обратное в общем случае не имеет места. Например, функция $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ вида

$$f(x, y) = \begin{cases} 1, & y = x^2 \\ 0, & \text{иначе} \end{cases}$$

не имеет предела в точке $0 = (0,0)$, но для любого ненулевого вектора $a \in \mathbb{R}^n$ функция f имеет предел по направлению a , равный нулю. Утверждение 2. Пусть отображение $f = f(x) \in \mathbb{R}^m$ определено в проколотой окрестности точки $x \in \mathbb{R}^n$. Пусть $y \in \mathbb{R}^m$. Для того, чтобы предел f в точке x (по совокупности переменных) был равен y необходимо и достаточно, чтобы для любой дифференцируемой в нуле вектор-функции $g = g(t) \in \mathbb{R}^n$ такой, что $g(0) = x$ и $g'(0) \neq 0$, существовал $\lim_{t \rightarrow 0} f(g(t)) = y$

Доказательство:

Не уменьшая общности пусть $x = 0$ и $y = 0$.

Необходимость.

Пусть $f(z) \rightarrow 0$ при $z \rightarrow 0$. Тогда для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что при $0 < |z| < \delta$ выполнено $|f(z)| < \varepsilon$. Для любой дифференцируемой в нуле (и, следовательно, непрерывной в нуле) вектор-функции $g = g(t) \in \mathbb{R}^n$ такой, что $g(0) = 0$ и $g'(0) = \nu \in \mathbb{R}^n \setminus \{0\}$ выполнено $g(t) = \nu t + o(t)$ при $t \rightarrow 0$. Поэтому существует $\rho > 0$ такое, что при $0 < |t| < \rho$ выполнено $0 < |g(t)| < \delta$, следовательно $|f(g(t))| < \varepsilon$.

Достаточность.

Пусть $f(z) \rightarrow 0$ при $z \rightarrow 0$. Тогда существуют $\varepsilon > 0$ и сходящаяся к нулю последовательность Гейне $\{z_h\}_{h \in \mathbb{N}}$ такая, что $|f(z_h)| \geq \varepsilon$ для всех $h \in \mathbb{N}$.

По теореме Больцано-Вейерштрасса мы можем выделить подпоследовательность $\{z_{n_k}\}$ такую, что $z_{n_k}/|z_{n_k}| \rightarrow \nu$ при $k \rightarrow \infty$ для некоторого единичного вектора $\nu \in \mathbb{R}^n$. Поэтому не уменьшая общности будем считать, что $z_h/|z_h| \rightarrow \nu$ при $h \rightarrow \infty$.

Кроме того, не уменьшая общности будем считать, что $|z_{h+1}| < |z_h|$ для всех $h \in \mathbb{N}$. (Этого также можно добиться перейдя к соответствующей подпоследовательности.)

Теперь рассмотрим вектор-функцию

$$g(t) = \begin{cases} z_h, & t = |z_h| \\ \theta z_{h+1} + (1 - \theta)z_h, & t = \theta |z_{h+1}| + (1 - \theta)|z_h|, \theta \in (0,1) \end{cases}$$

продолженную при $t < 0$ нечётным образом и равную 0 при $t = 0$. (Заметим, что данная вектор-функция получается линейной интерполяцией пар $(|z_h|, z_h)$.)

По определению предела для любого $\delta > 0$ существует N такое, что при $h \geq N$ выполнено $|z_h/|z_h| - \nu| < \delta$. Тогда для любого t такого, что $|t| < |z_N|$ выбрав $h \geq N$ так, что $t = \theta |z_{h+1}| + (1 - \theta)|z_h|$ (для некоторого $\theta \in [0,1]$) получим

$$|g(t) - \nu t| = |\theta(z_{h+1} - \nu |z_{h+1}|) + (1 - \theta)(z_h - \nu |z_h|)| < \theta \delta |z_{h+1}| + (1 - \theta)\delta |z_h| = \delta |t|$$

В силу произвольности $\delta > 0$ получаем $g(t) = \nu t + o(t)$ при $t \rightarrow 0$. Значит векторфункция g дифференцируема в нуле, $g(0) = 0$ и $g'(0) = \nu \neq 0$, но при этом

$$f(g(t)) \not\rightarrow 0$$

при $t \rightarrow 0$, так как $f(g(|z_h|)) \not\rightarrow 0$ при $h \rightarrow \infty$. Противоречие.

В общем случае существование пределов функции вдоль дифференцируемых кривых, проходящих через данную точку, не влечёт существование предела данной функции по совокупности переменных. Иначе говоря, в утверждении 2 важно не только существование предела вдоль различных дифференцируемых в нуле кривых g , но и независимость этих пределов от выбора g . Действительно, рассмотрим

$$f(x, y) = \frac{x^2 - y^2}{x^2 + y^2}, \quad x^2 + y^2 > 0$$

Если вектор-функция $g = g(t)$ дифференцируема в нуле и $g(0) = 0$ и $g'(0) \neq 0$, то существует числа α, β такие, что $g(t) = (\alpha t + o(t), \beta t + o(t))$ при $t \rightarrow 0$. Тогда

$$f(g(t)) = \frac{\alpha^2 - \beta^2 + o(1)}{\alpha^2 + \beta^2 + o(1)} \rightarrow \frac{\alpha^2 - \beta^2}{\alpha^2 + \beta^2}$$

при $t \rightarrow 0$. Однако предела по совокупности переменных у f в нуле нет, так как $f(x, 0) = 1$ и $f(0, y) = -1$ (при $x, y \neq 0$).

Опишем предел отображения в точке, не привлекая понятия ε -окрестностей:

Утверждение 3. Пусть $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Точка $y \in \mathbb{R}^m$ является пределом функции f в точке $x \in \mathbb{R}^n$ тогда и только тогда, когда для любой окрестности V точки y существует окрестность U точки x такая, что $f(U \setminus \{x\}) \subset V$.

Доказательство:

Необходимость.

Так как V открыто и $y \in V$ то существует $\varepsilon > 0$ такое, что $U_\varepsilon(y) \subset V$. По определению непрерывности существует $\delta > 0$ такое, что $f(U_\delta(x) \setminus \{x\}) \subset U_\varepsilon(y)$. Тогда мы можем взять $U = U_\delta(x)$.

Достаточность.

Пусть $\varepsilon > 0$. Возьмём открытую окрестность U точки x так, что $f(U \setminus \{x\}) \subset V$. Так как $x \in U$ то существует $\delta > 0$ такое, что $U_\delta(x) \subset U$. Значит $f(\overset{\circ}{U}_\delta(x)) \subset f(U \setminus \{x\}) \subset V$

Непрерывность отображений и топология Отображение $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ называется непрерывным в точке $x \in \mathbb{R}^n$, если $\lim_{z \rightarrow x} f(z) = f(x)$. Если f непрерывно в каждой точке $x \in \mathbb{R}^n$, то оно называется непрерывным (на \mathbb{R}^n).

Напомним, что множество $f^{-1}(V) = \{x \in \mathbb{R}^n : f(x) \in V\}$ называется полным прообразом множества $V \subset \mathbb{R}^m$ под действием f .

Теорема 4. (критерий непрерывности отображения). Отображение $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ непрерывно тогда и только тогда, когда для любого открытого множества $V \subset \mathbb{R}^m$ его полный прообраз $f^{-1}(V)$ открыт.

Доказательство:

Так как для любого множества $V \subset \mathbb{R}^m$ выполнено

$$f^{-1}(V^c) = (f^{-1}(V))^c$$

то несложно проверить, что непрерывность отображения $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ эквивалентна тому, что полный прообраз любого замкнутого множества $F \subset \mathbb{R}^m$ под действием f замкнут.

Свойства непрерывных отображений, связанные с компактами

Напомним, что для любого отображения $f : X \rightarrow Y$ и любого множества $A \subset X$

множество $f(A) := \{f(x) : x \in A\}$ называется образом A под действием отображения f

Обобщением теоремы Вейерштрасса является следующий результат:

Теорема 5. Пусть отображение $f : K \rightarrow \mathbb{R}^m$ непрерывно на компакте $K \subset \mathbb{R}^n$. Тогда $f(K)$ — компакт.

Доказательство

В приведённом выше доказательстве мы использовали одно из трёх эквивалентных определений компактности (связанное с леммой Гейне-Бореля). В качестве упражнения рекомендуется провести доказательство с использованием каждого из двух оставшихся определений компактности.

Связность и линейная связность

Определение 6. Множество $E \subset \mathbb{R}^d$ называется линейно связным, если для любых двух точек $x, y \in E$ существует непрерывная функция $\varphi : [0, 1] \rightarrow E$ такая, что $\varphi(0) = x$ и $\varphi(1) = y$

Иначе говоря, множество $E \subset \mathbb{R}^d$ линейно связно тогда и только тогда, когда любые две его точки можно соединить непрерывной кривой, лежащей в E .

Например, множество $E = [0, 1]$ линейно связно (в \mathbb{R}), так как для любых $x, y \in E$ можно взять $\varphi(t) = (1 - t)x + ty$

Утверждение 8.. Если отображение $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ непрерывно, то для любого линейно связного множества $E \subset \mathbb{R}^n$ множество $f(E)$ линейно связно.

Доказательство:

Близким к понятию линейной связности является понятие (топологической) связности:

Определение 9. Множество $E \subset \mathbb{R}^d$ называется связным (или топологически связным), если не существует открытых множеств $U_1, U_2 \subset \mathbb{R}^d$ таких, что $U_1 \cap U_2 = \emptyset, E \subset U_1 \cup U_2$, причём $E \cap U_1 \neq \emptyset$ и $E \cap U_2 \neq \emptyset$.

Иначе говоря, множество $E \subset X$ связно тогда и только тогда, когда его нельзя нетривиальным образом покрыть двумя попарно не пересекающимися открытыми множествами.

На самом деле обычно используется немного другое определение связности, которое для подмножеств \mathbb{R}^d (и других метрических пространств) эквивалентно данному выше.

Утверждение 10. Множество $E \subset \mathbb{R}^d$ связно тогда и только тогда, когда не существует открытых множеств $U_1, U_2 \subset \mathbb{R}^d$ таких, что $U_1 \cap U_2 \cap E \neq \emptyset, E \subset U_1 \cup U_2$, причём $E \cap U_1 \neq \emptyset$ и $E \cap U_2 \neq \emptyset$.

Доказательство:

Необходимость очевидна, так как из $U_1 \cap U_2 = \emptyset$ следует, что $U_1 \cap U_2 \cap E = \emptyset$. Для доказательства достаточности построим открытые множества $V_1, V_2 \subset \mathbb{R}^d$ такие, что $V_1 \cap V_2 = \emptyset, E \subset V_1 \cup V_2$, причём $E \cap V_1 \neq \emptyset$ и $E \cap V_2 \neq \emptyset$. В отличие от множеств U_1 и U_2 , данных в условии, множества V_1 и V_2 не должны пересекаться.

Обозначим $E_1 := E \cap U_1$ и $E_2 := E \cap U_2$. Так как $E_1 \cap E_2 = \emptyset$, то $E_2 \subset U_1^c$. Поэтому для каждой точки $x \in E_1$ существует число $\varepsilon(x) > 0$ такое, что $U_{2\varepsilon(x)}(x) \cap E_2 = \emptyset$. Аналогично для каждой точки $x \in E_2$ существует число $\delta(x) > 0$ такое, что $U_{2\delta(x)}(x) \cap E_1 = \emptyset$.

Возьмём $V_1 := \bigcup_{x \in E_1} U_{\varepsilon(x)}(x)$ и $V_2 := \bigcup_{y \in E_2} U_{\delta(y)}(y)$. Если $z \in V_1 \cap V_2$ то существуют $x \in E_1$ и $y \in E_2$ такие, что $z \in U_{\varepsilon(x)}(x) \cap U_{\delta(y)}(y)$. Тогда

$$|x - y| \leq |x - z| + |z - y| < \varepsilon(x) + \delta(y) < 2 \max(\varepsilon(x), \delta(y)).$$

Тогда либо $x \in U_{2\delta(y)}(y)$, либо $y \in U_{2\varepsilon(x)}(x)$. Противоречие. Значит $V_1 \cap V_2 = \emptyset$.

По построению множества V_1 и V_2 открыты. Кроме того, $\emptyset \neq E_1 \subset V_1$ и $\emptyset \neq E_2 \subset V_2$. Утверждение 11. Множество $E = [0, 1]$ связно (в \mathbb{R}).

Доказательство:

Иначе найдутся открытые $U_1, U_2 \subset \mathbb{R}$ такие, что $U_1 \cap U_2 = \emptyset, E \subset U_1 \cup U_2$, причём $E \cap U_1 \neq \emptyset$ и $E \cap U_2 \neq \emptyset$.

Не уменьшая общности пусть $0 \in U_1$. Множество $M = \{x \in E : [0, x] \subset U_1\}$ непусто, так как U_1 открыто. Пусть $x = \sup M$. Так как U_2 открыто и не пересекается с U_1 , то точка x не может принадлежать U_2 . Следовательно, $x \in U_1$. Если $x \neq 1$ то в силу открытости U_1 существует $x' > x$, принадлежащее M , что противоречит определению точной верхней грани.

На самом деле для доказательства связности отрезка $[0, 1]$ можно было бы воспользоваться уже доказанной для него линейной связностью:

Утверждение 12. Если множество $E \subset X$ линейно связно, то оно связно.

Доказательство:

Предположим противное. Тогда существуют открытые множества $U_1, U_2 \subset X$ такие, что $U_1 \cap U_2 = \emptyset, E \subset U_1 \cup U_2$, причём $E \cap U_1 \neq \emptyset$ и $E \cap U_2 \neq \emptyset$. Возьмём $x \in E \cap U_1$, $y \in E \cap U_2$. В силу линейной

связности существует непрерывная функция $f : [0,1] \rightarrow E$ такая, что $f(0) = x$ и $f(1) = y$. В силу непрерывности этой функции множества $f^{-1}(U_1)$ и $f^{-1}(U_2)$ открыты и покрывают отрезок $[0,1]$, при этом не пересекаясь. В силу связности отрезка $[0,1]$ одно из них должно содержать весь отрезок, что противоречит выбору точек x и y .

Обратное в общем случае неверно, как показывает следующий классический пример.

Задача 13 (Синусоида тополога). Привести пример связного множества, не являющегося линейно связным

Решение

Пусть $X = \mathbb{R}^2$. Рассмотрим множество $E = E_1 \cup E_2$, где

$$E_1 = \{0\} \times [-1,1], \quad E_2 = \left\{ \left(x, \sin \frac{1}{x} \right) : x > 0 \right\}$$

Линейную связность множеств E_1 и E_2 легко проверить. Например, если $x > 0$ и $y > 0$ и

$$f(t) = \left((1-t)x + ty, \sin \frac{1}{(1-t)x + ty} \right)$$

то $f : [0,1] \rightarrow E$ непрерывна, $f(0) = (x, \sin \frac{1}{x})$ и $f(1) = (y, \sin \frac{1}{y})$. Допустим, что существуют открытые $U_1, U_2 \subset \mathbb{R}^2$ такие, что $U_1 \cap U_2 = \emptyset$, $E \subset U_1 \cup U_2$, причём $E \cap U_1 \neq \emptyset$ и $E \cap U_2 \neq \emptyset$. Так как E_1 и E_2 связны, то каждое из множеств U_1 и U_2 либо не пересекает E_i , либо содержит его целиком (где $i \in \{1,2\}$). Не уменьшая общности пусть $E_1 \subset U_1$ и $E_2 \subset U_2$.

Так как U_1 открыто, то некоторый круг $B_r(0)$ с центром в $(0,0)$ и радиусом $r > 0$ целиком лежит в U_1 . С другой стороны, точки $(\frac{1}{\pi k}, 0)$ принадлежат $B_r(0)$ при достаточно больших $k \in \mathbb{N}$ и в то же время всегда принадлежат $E_2 \subset U_2$, что даёт противоречие. Следовательно, E связно.

Если бы E было линейно связным, то нашлась бы непрерывная функция $f : [0,1] \rightarrow E$ такая, что $f(0) = (0,0)$ и $f(1) = (1, \sin 1)$. Обозначим $f(t) = (x(t), y(t))$, $t \in [0,1]$. Пусть $t_0 = \sup\{t \in [0,1] : x(t) = 0\}$. В силу непрерывности функции x имеем $x(t_0) = 0$, а также $t_0 < 1$. Тогда в силу непрерывности f для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что

$$f([t_0, t_0 + \delta]) \subset B_\varepsilon(f(t_0)).$$

По определению t_0 имеем $y := x(t_0 + \delta) > 0$. Тогда по теореме о промежуточных значениях непрерывной функции имеем $(0, y) \subset x([t_0, t_0 + \delta])$. Следовательно,

$$\left\{ \left(z, \sin \frac{1}{z} \right) : z \in (0, y) \right\} \subset f([t_0, t_0 + \delta]).$$

Так как для любого $y > 0$ выполнено $\sin \frac{1}{(0,y)} = [-1,1]$, то из (1) и (2) следует, что диаметр шара $B_\varepsilon(f(t_0))$ не меньше 2, что противоречит произвольности ε .

Тем не менее, для открытых подмножеств \mathbb{R}^n связность равносильна линейной связности:

Задача 14. Доказать, что всякое связное открытое множество $E \subset \mathbb{R}^n$ линейно связно.

Указание Зафиксировав точку $x \in E$ нужно проверить, является ли множество

$$V_x = \{y \in E \mid \text{существует непрерывная } f : [0,1] \rightarrow E \text{ такая, что } f(0) = x \text{ и } f(1) = y\}$$

открытым.

МА-32. Теорема Бэра на вещественной прямой

Начать знакомство с теоремой Бэра проще всего со случая вещественной прямой (далее мы рассмотрим другие метрические пространства). Напомним, что множество $E \subset \mathbb{R}$ всюду плотно в \mathbb{R} , если в любом непустом открытом интервале содержится хотя бы одна точка множества E . Это равносильно тому, что $\bar{E} = \mathbb{R}$.

Теорема 11. (Бэр, для открытых множеств). Пусть для каждого $n \in \mathbb{N}$ множество $G_n \subset \mathbb{R}$ открыто и всюду плотно в \mathbb{R} . Тогда $\bigcap_{n=1}^{\infty} G_n$ всюду плотно в \mathbb{R} .

Доказательство:

Рассмотрим произвольный открытый интервал I_0 в \mathbb{R} . По условию в нём найдётся хотя бы одна точка x_1 , принадлежащая множеству G_1 . Вместе с ней в G_1 войдёт и некоторый открытый интервал I_1 с центром в x_1 . Этот интервал можно выбрать так, что $\bar{I}_1 \subset I_0 \cap G_1$ и длина I_1 не превосходит 1.

Продолжая этот процесс мы получим последовательность открытых интервалов $\{I_k\}_{k \in \mathbb{N}}$ таких, что для каждого $k \in \mathbb{N}$ выполнено $\bar{I}_k \subset I_{k-1} \cap G_k$ и длина интервала I_k не превосходит $1/k$. По теореме Кантора существует единственная точка $x \in \bigcap_{k=1}^{\infty} \bar{I}_k$. Так как для каждого $k \in \mathbb{N}$ по построению $x \in \bar{I}_k \subset G_k \cap I_0$, то $x \in \bigcap_{k=1}^{\infty} G_k \cap I_0$.

На практике бывает удобнее пользоваться контрапозицией доказанного утверждения:

Теорема 2. (Бэр, для замкнутых множеств). Если $\mathbb{R} = \bigcup_{n=1}^{\infty} F_n$, где для каждого $n \in \mathbb{N}$ множество $F_n \subset \mathbb{R}$ замкнуто, то хотя бы одно из множеств F_n включает в себя некоторый непустой открытый интервал.

Доказательство:

В противном случае открытые множества $G_n := \mathbb{R} \setminus F_n$ всюду плотны в \mathbb{R} и, следовательно, их пересечение непусто по теореме Бэра. Переходя к дополнению получаем

$$\mathbb{R} \neq \mathbb{R} \setminus \left(\bigcap_{n=1}^{\infty} G_n \right) = \bigcup_{n=1}^{\infty} F_n = \mathbb{R}$$

Противоречие.

Данной версии теоремы Бэра достаточно для многих приложений. Например, с её помощью можно доказать, что множество всех иррациональных чисел нельзя представить в виде не более чем счётного объединения замкнутых множеств.

Метрические пространства, их полнота

Определение 3. Пусть X - множество. Функция $\rho : X \times X \rightarrow \mathbb{R}$ называется метрикой на X , если для любых $x, y \in X$

1. $\rho(x, y) = 0$ тогда и только тогда, когда $x = y$
2. $\rho(x, y) = \rho(y, x)$ (симметричность);
3. для любой точки $z \in X$ выполнено $\rho(x, y) \leq \rho(x, z) + \rho(z, y)$ (неравенство треугольника).

Если на множестве X введена метрика ρ , то пара (X, ρ) называется метрическим пространством. (На практике само множество X также называют метрическим пространством, если из контекста ясно, какая метрика на нём рассматривается.)

Из определения сразу следует, что для всех $x, y \in X$ выполнено $\rho(x, y) \geq 0$. Действительно, $0 = \rho(x, x) \leq \rho(x, y) + \rho(y, x) = 2\rho(x, y)$.

Любое множество X можно ввести дискретную метрику:

$$\rho(x, y) := \begin{cases} 1, & x \neq y \\ 0, & x = y \end{cases}$$

(несложно проверить, что она действительно удовлетворяет аксиомам метрики.)

В первую очередь в качестве метрических пространств мы будем рассматривать различные подмножества $E \subset \mathbb{R}^d$ с евклидовой метрикой

$$\rho(x, y) = |x - y|.$$

То, что данная функция действительно является метрикой, следует из известных свойств длины вектора.

Определение 4. Последовательность $\{x_n\}_{n \in \mathbb{N}} \subset X$ элементов метрического пространства (X, ρ) называется фундаментальной, если для любого $\varepsilon > 0$ существует $N \in \mathbb{N}$ такое, что для всех $m, n > N$ выполнено

$$\rho(x_n, x_m) < \varepsilon.$$

Утверждение 5. Если последовательность $\{x_n\}$ элементов метрического пространства сходится, то она фундаментальна.

Определение 6. Метрическое пространство (X, ρ) называется полным, если в нём любая фундаментальная последовательность сходится.

Теорема 7. Множество $E \subset \mathbb{R}^d$ является полным метрическим пространством (относительно евклидовой метрики) тогда и только тогда, когда E замкнуто.

Доказательство:

Необходимость.

Пусть E полно и последовательность элементов E сходится к некоторой точке $x \in \mathbb{R}^d$. В силу полноты E она сходится к некоторой точке $y \in E$, а в силу единственности предела $x = y$. Значит $x \in E$. Следовательно, E замкнуто.

Достаточность.

Пусть E замкнуто и последовательность $\{x_n\}_{n \in \mathbb{N}} \subset E$ фундаментальна. Тогда она сходится к некоторому элементу $x \in \mathbb{R}^d$. Значит в силу замкнутости E имеем $x \in E$.

В произвольном метрическом пространстве (X, ρ) , подобно пространству \mathbb{R}^d , открытым шаром радиуса $r \geq 0$ с центром в $x \in X$ называется множество $U_r(x) = \{y \in X : \rho(x, y) < r\}$. Множество $\{y \in X : \rho(x, y) \leq r\}$ называется соответственно замкнутым шаром радиуса r с центром в x . Открытые и замкнутые множества (а также внутренность, замыкание и граница) в метрических пространствах определяются точно так же, как в \mathbb{R}^d . Несложно проверить, что замкнутый шар действительно

является замкнутым множеством. Однако не всегда замкнутый шар совпадает с замыканием соответствующего открытого шара.

Теорема 7. практически дословно обобщается на метрические пространства: если $X \subset Y$ и метрическое пространство (Y, ρ) полно, то для полноты (X, ρ) необходимо и достаточно, чтобы X было замкнуто в (Y, ρ) .

Теорема 8. (теорема Кантора для метрических пространств). Метрическое пространство (X, ρ) полно тогда и только тогда, когда в нём любая стягивающаяся последовательность вложенных замкнутых шаров имеет единственную общую точку. рассматриваемая последовательность является стягивающейся, то для каждого $\varepsilon > 0$ существует натуральное число $N > 0$ такое, что при $k \geq N$ диаметр шара B_k будет меньше чем $\varepsilon/2$. Так как при $k \geq N$ выполнено $x_k \in B_N$, то для всех $m, n \geq N$ выполнено $\rho(x_n, x_m) < \varepsilon$, значит последовательность $\{x_k\}_{k \in \mathbb{N}}$ фундаментальна. В силу полноты рассматриваемого пространства она сходится к некоторой точке $x \in X$. В силу замкнутости рассматриваемых шаров она принадлежит каждому из них, следовательно $x \in \bigcap_{k=1}^{\infty} B_k$

Достаточность.

Пусть последовательность $\{x_n\}_{n \in \mathbb{N}}$ элементов X фундаментальна. Рассмотрим бесконечно малые последовательности $\{\varepsilon_k\}_{k \in \mathbb{N}}$ $\{r_k\}_{k \in \mathbb{N}}$, которые будут определены ниже.

Из определения фундаментальной последовательности следует, что существует строго возрастающая последовательность натуральных чисел $\{N_k\}_{k \in \mathbb{N}}$ такая, что для каждого $k \in \mathbb{N}$ для всех $m, n \geq N_k$ выполнено $\rho(x_n, x_m) < \varepsilon_k$.

Рассмотрим последовательность замкнутых шаров B_k с центрами x_{N_k} и радиусами r_k . Чтобы эта последовательность была вложенной достаточно, чтобы

$$r_{k+1} + \varepsilon_k \leq r_k$$

Действительно, в этом случае для любого $y \in B_{k+1}$ выполнено

$$\rho(y, x_k) \leq \rho(y, x_{k+1}) + \rho(x_{k+1}, x_k) < r_{k+1} + \varepsilon_k < r_k$$

значит $y \in B_k$. Тогда $B_{k+1} \subset B_k$.

Чтобы (1) выполнялось, достаточно взять $r_k = 2^{-k}$ и $\varepsilon_k = 2^{-k-1}$.

Итак, мы построили стягивающуюся последовательность вложенных замкнутых шаров B_k . По условию эта последовательность имеет единственную общую точку $x \in X$. По построению $x_{N_k} \rightarrow x$ при $k \rightarrow \infty$. Тогда в силу фундаментальности вся последовательность $\{x_n\}_{n \in \mathbb{N}}$ сходится к x .

Существенно, что в теореме Кантора идёт речь о стягивающейся последовательности вложенных замкнутых шаров. Если от этого условия отказаться, то она перестанет быть верной. Например, если на \mathbb{R} рассмотреть метрику

$$\rho(x, y) = \begin{cases} 0, & x = y \\ \pi + |\arctg x - \arctg y|, & x \neq y \end{cases}$$

то в метрическом пространстве (\mathbb{R}, ρ) можно построить последовательность вложенных замкнутых шаров, не имеющих ни одной общей точки.

Теорема Бэра для метрических пространств Пусть (X, ρ) - метрическое пространство (причём $X \neq \emptyset$). Так же как в случае \mathbb{R}^d , множество $A \subset X$ называется всюду плотным во множестве $B \subset X$,

если $B \subset \bar{A}$. Доказательство теоремы Бэра, приведённое выше для $X = \mathbb{R}$, можно практически без изменений повторить для полных метрических пространств:

Теорема 9. (Бэр, для открытых множеств в метрическом пространстве). Пусть метрическое пространство (X, ρ) полно. Если для каждого $n \in \mathbb{N}$ множество $G_n \subset X$ открыто и всюду плотно в X , то $\bigcap_{n=1}^{\infty} G_n$ всюду плотно в X .

Доказательство:

Рассмотрим произвольный открытый шар B_0 в X . По условию в нём найдётся хотя бы одна точка x_1 , принадлежащая множеству G_1 . Вместе с ней в G_1 войдёт и некоторый открытый шар B_1 с центром в x_1 . Этот шар можно выбрать так, что $\bar{B}_1 \subset B_0 \cap G_1$ и радиус B_1 не превосходит 1.

Продолжая этот процесс мы получим последовательность открытых шаров $\{B_k\}_{k \in \mathbb{N}}$ таких, что для каждого $k \in \mathbb{N}$ выполнено $\bar{B}_k \subset B_{k-1} \cap G_k$ и длина шара B_k не превосходит $1/k$. По теореме Кантора существует единственная точка $x \in \bigcap_{k=1}^{\infty} \bar{B}_k$. Так как для каждого $k \in \mathbb{N}$ по построению $x \in \bar{B}_k \subset G_k \cap B_0$, то $x \in \bigcap_{k=1}^{\infty} G_k \cap B_0$.

Контрапозицией доказанной теоремы получаем более распространённую версию теоремы Бэра:

Теорема 10. (Бэр, для замкнутых множеств в метрическом пространстве). Пусть метрическое пространство (X, ρ) полно. Пусть для каждого $n \in \mathbb{N}$ множество $F_n \subset X$ замкнуто и $X = \bigcup_{n=1}^{\infty} F_n$. Тогда существуют одно из множеств F_n содержит некоторый шар, то есть существуют $\varepsilon > 0, n \in \mathbb{N}$ и $x \in X$ такие, что $U_\varepsilon(x) \subset F_n$.

Доказательство:

Если для любых $\varepsilon > 0$ и $x \in X$ выполнено $U_\varepsilon(x) \not\subset F_n$, то $G_n := F_n^c$ всюду плотно в X . Если это верно для любого $n \in \mathbb{N}$, то по теореме Бэра $\bigcap_{n=1}^{\infty} G_n \neq \emptyset$. Переходя к дополнению получаем $\bigcup_{n=1}^{\infty} F_n \neq X$, что противоречит условию. $\emptyset \neq E \cap (a, b) \subset F_n$

Приложения теоремы Бэра

Задача 11. Доказать, что множество иррациональных чисел нельзя представить в виде не более чем счётного объединения замкнутых множеств.

Задача 12. Существует ли функция $f : \mathbb{R} \rightarrow \mathbb{R}$, непрерывная во всех рациональных точках и разрывная во всех иррациональных точках?

▽ Указание

Функция f непрерывна в точке x тогда и только тогда, когда этой точке x её колебание

$$\omega_f(x) = \inf_{\delta > 0} \sup_{y, z \in U_\delta(x)} |f(y) - f(z)|$$

равно нулю. Кроме того, для каждого $n \in \mathbb{N}$ множество $\{x : \omega_f(x) < 1/n\}$ открыто.

Задача 13. Пусть функция $f : \mathbb{R} \rightarrow \mathbb{R}$ бесконечно дифференцируема и для каждой точки $x \in \mathbb{R}$ существует $n = n(x) \in \mathbb{N}$ такое, что $f^{(n)}(x) = 0$. Верно ли, что f — многочлен?

▽ Указание

Пусть H — множество всех точек $x \in \mathbb{R}$ таких, что для каждого $x \in H$ существует $\varepsilon = \varepsilon_x > 0$, для которого на интервале $U_\varepsilon(x)$ функция f совпадает с некоторым многочленом степени $m = m_x \in \mathbb{N} \cup \{0\}$.

Обозначим $E := \mathbb{R} \setminus H$. Предположим что E непусто. Тогда

1. Если $(\alpha, \beta) \subset H$, то на всём интервале (α, β) функция f совпадает с некоторым многочленом;

Пояснение:

Это следует из леммы Гейне-Бореля, так как любой отрезок, лежащий в (α, β) , покрывается интервалами $U_{\varepsilon_x}(x)$.

2. Множество E замкнуто и не содержит изолированных точек (такие множества называются совершенными); В силу предыдущего пункта в правой и левой окрестности любой изолированной точки функция совпадает с некоторыми многочленами, что возможно только в случае когда их коэффициенты равны, так как $f \in C^\infty(\mathbb{R})$.
3. Для каждого $n \in \mathbb{N}$ множество $F_n := \{x \in \mathbb{R} : f^{(n)}(x) = 0\}$ замкнуто;
4. Существуют $n \in \mathbb{N}$ и интервал (a, b) такие, что $\emptyset \neq E \cap (a, b) \subset F_n$;

Пояснение:

Это следует из теоремы Бэра.

5. Для всех $x \in (a, b) \cap E$ и всех целых $k \geq n$ выполнено $f^{(k)}(x) = 0$;

Пояснение: Это следует из того, что любая точка $x \in E$ является пределом последовательности точек $x_j \in E$, отличных от x (так как E - совершенное). А по теореме Ролля между x_j и x_{j+1} всегда найдётся точка ξ_j такая, что $f^{(n+1)}(\xi_j) = 0$. Аналогично для производных высших порядков.

6. Для всех $x \in (a, b) \setminus E$ имеет место неравенство $m_x < n$, следовательно, $f^{(n)}(x) \neq 0$

Из двух последних пунктов получаем, что $(a, b) \subset F_n$, а значит $E \cap (a, b) = \emptyset$. Противоречие.

МА-33. Определённый интеграл Римана

Предположим, что на отрезке $[a, b]$ задана вещественная функция f . (Здесь $a < b$.)

Определение 1. Разбиением отрезка $[a, b]$ будем называть любой упорядоченный набор точек $p = \{x_i\}_{i=0}^n$ такой, что

$$a = x_0 < x_1 < \dots < x_n = b$$

Мелкостью разбиения p будем называть число

$$\ell(p) := \max_{i \in \overline{1, n}} \Delta x_i$$

где

$$\Delta x_i := x_i - x_{i-1}$$

Под H_i будем понимать отрезок $[x_{i-1}, x_i]$. Под $P_\delta[a, b]$ будем понимать семейство всех разбиений отрезка $[a, b]$, для которых $\ell(p) < \delta$.

Определение 2. Если $p = \{x_i\}_{i=0}^n$ - разбиение отрезка $[a, b]$, то любой набор точек $\xi = \{\xi_i\}_{i=1}^n$ таких, что для всех $i \in \overline{1, n}$ выполнено

$$\xi_i \in [x_{i-1}, x_i]$$

называется выборкой из разбиения p . Под V_p будем понимать совокупность всевозможных выборок из разбиения p .

Определение 3. Интегральной суммой Римана для функции f , разбиения $p = \{x_i\}_{i=0}^n$ и выборки ξ из разбиения p будем называть сумму

$$\sigma(f, p, \xi) := \sum_{i=1}^n f(\xi_i) \Delta x_i$$

Определение 4. Если существует число $I \in \mathbb{R}$ такое, что для любого $\varepsilon > 0$ найдётся $\delta > 0$ такое, что для любого разбиения p отрезка $[a, b]$ с мелкостью $\ell(p) < \delta$ и для любой выборки ξ из разбиения p выполнено

$$|\sigma(f, p, \xi) - I| < \varepsilon$$

В этом случае число I называется интегралом (Римана) от функции f по отрезку $[a, b]$ и обозначается

$$\int_a^b f(x) dx := I$$

а функция f называется интегрируемой по Риману на отрезке $[a, b]$.

Под $\mathcal{R}[a, b]$ будем понимать множество всех функций, интегрируемых по Риману на отрезке $[a, b]$

Если $b < a$, то по определению считаем, что $\int_a^b f dx := -\int_b^a f dx$.

Из определения интеграла Римана легко видеть, что число I определено однозначно. Действительно, если число J также удовлетворяет определению, то из неравенства

$$|I - J| \leq |I - \sigma(f, p, \xi)| + |\sigma(f, p, \xi) - J|$$

легко видеть, что для любого $\varepsilon > 0$ выполнено $|I - J| < 2\varepsilon$.

Пользуясь определением, несложно проверить, что для любых α, β удовлетворяющих неравенствам $a \leq \alpha < \beta \leq b$ функция $1_{[\alpha, \beta]}$ (индикатор множества $[\alpha, \beta]$) интегрируема по Риману на $[a, b]$ и

$$\int_a^b 1_{[\alpha, \beta]} dx = \beta - \alpha$$

Аналогичное верно для полуинтервалов $[\alpha, \beta), (\alpha, \beta]$ и интервала (α, β) .

Основные необходимые и достаточные условия интегрируемости

Теорема 5. (критерий Коши интегрируемости по Риману). Функция f интегрируема по Риману на $[a, b]$ тогда и только тогда, когда для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что для любых разбиений p, q мелкостью меньше δ и любых выборок ξ, η из p и q выполнено

$$|\sigma(f, p, \xi) - \sigma(f, q, \eta)| < \varepsilon$$

Доказательство:

Необходимость очевидна из неравенства

$$|\sigma(f, p, \xi) - \sigma(f, q, \eta)| \leq |\sigma(f, p, \xi) - I| + |I - \sigma(f, q, \eta)|$$

где $I = \int_a^b f dx$

Для доказательства достаточности сначала рассмотрим последовательность разбиений p_n и выборок ξ_n из p_n таких, что $\ell(p_n) \rightarrow 0$ при $n \rightarrow \infty$. Тогда по условию числовая последовательность

$$\sigma_n := \sigma(f, p_n, \xi_n)$$

фундаментальна, а значит сходится (по критерию Коши для числовой последовательности). Пусть $I = \lim_{n \rightarrow \infty} \sigma_n$. Тогда по определению предела для любого $\varepsilon > 0$ существует $N \in \mathbb{N}$ такое, что $|\sigma_n - I| < \varepsilon$ при $n \geq N$. Для данного ε выберем $\delta > 0$ так, что выполнено (1). Тогда для любого разбиения p с $\ell(p) < \delta$ выполнено

$$|\sigma(f, p, \xi) - I| \leq |\sigma(f, p, \xi) - \sigma_N| + |\sigma_N - I| < 2\varepsilon$$

Значит по определению $f \in \mathcal{R}[a, b]$

Из критерия Коши следует, что функция Дирихле $1_{\mathbb{Q}}$ не интегрируема по Риману на $[a, b]$. Однако для исследования большинства конкретных функций применять критерий Коши напрямую как

правило нецелесообразно. Для этого используются другие критерии и необходимые и достаточные условия, которые мы далее получим как следствия.

Напомним, что колебанием функции f на множестве $A \subset [a, b]$ называется значение

$$\omega(f, A) := \sup_{x, y \in A} |f(x) - f(y)| = \sup_A f - \inf_A f$$

Определение 6. Пусть $p = \{x_i\}_{i=0}^n$ - разбиение отрезка $[a, b]$. Тогда взвешенным колебанием f на разбиении p будем называть значение

$$\Omega(f, p) := \sum_{i=1}^n \omega(f, [x_{i-1}, x_i]) \Delta x_i$$

Теорема 7. (критерий интегрируемости). Функция f интегрируема по Риману на $[a, b]$ тогда и только тогда, когда для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что для любого разбиения p мелкостью меньше δ выполнено

$$\Omega(f, p) < \varepsilon$$

Доказательство:

Необходимость.

По определению супремума для любого разбиения $p = \{x_i\}_{i=0}^n$ отрезка $[a, b]$ выполнено

$$\Omega(f, p) = \sum_{i=1}^n \sup_{\xi_i, \eta_i \in [x_{i-1}, x_i]} |f(\xi_i) - f(\eta_i)| \Delta x_i = \sup_{\xi, \eta} |\sigma(f, p, \xi) - \sigma(f, p, \eta)|$$

где супремум берётся по всевозможным выборкам ξ, η из разбиения p . Если $f \in \mathcal{R}[a, b]$, то для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что при $\ell(p) < \delta$ для любых выборок ξ, η из разбиения p выполнено $|\sigma(f, p, \xi) - \sigma(f, p, \eta)| < \varepsilon$. Тогда $\Omega(f, p) \leq \varepsilon$.

Достаточность.

Мы снова воспользуемся критерием Коши. Для этого потребуется оценить $|\sigma(f, p, \xi) - \sigma(f, q, \eta)|$ для различных разбиений $p = \{x_i\}_{i=0}^n$ и $q = \{y_j\}_{j=0}^m$ отрезка $[a, b]$. Начнём с более простого случая, когда разбиение q является измельчением разбиения p , то есть $q \subset p$. В этом случае для каждого $i \in \overline{0, n}$ существует $n_i \in \overline{0, m}$ такое, что

$$x_i = y_{n_i}$$

причём $n_{i-1} < n_i$. Тогда

$$\Delta x_i = \sum_{j=n_{i-1}+1}^{n_i} \Delta y_j$$

и, следовательно,

$$\begin{aligned}
|\sigma(f, p, \xi) - \sigma(f, q, \eta)| &= \left| \sum_{i=1}^n f(\xi_i) \Delta x_i - \sum_{i=1}^n \sum_{j=n_{j-1}+1}^{n_j} f(\eta_j) \Delta y_j \right| \\
&= \left| \sum_{i=1}^n \sum_{j=n_{j-1}+1}^{n_j} (f(\xi_i) - f(\eta_j)) \Delta y_j \right| \\
&\leq \sum_{i=1}^n \sum_{j=n_{j-1}+1}^{n_j} |f(\xi_i) - f(\eta_j)| \Delta y_j \\
&\leq \sum_{i=1}^n \omega(f, [x_{i-1}, x_i]) \sum_{j=n_{j-1}+1}^{n_j} \Delta y_j \\
&= \Omega(f, p).
\end{aligned}$$

В общем случае рассмотрим общее измельчение разбиений p и q :

$$r := p \cup q$$

Тогда в силу полученного выше неравенства для любой выборки ζ из разбиения r имеем

$$\begin{aligned}
|\sigma(f, p, \xi) - \sigma(f, q, \eta)| &\leq |\sigma(f, p, \xi) - \sigma(f, r, \zeta)| + |\sigma(f, r, \zeta) - \sigma(f, q, \eta)| \\
&\leq \Omega(f, p) + \Omega(f, q).
\end{aligned}$$

Отсюда видно, что если $\ell(p) < \delta$ и $\ell(q) < \delta$, то

$$|\sigma(f, p, \xi) - \sigma(f, q, \eta)| < 2\varepsilon$$

Значит f интегрируема по критерию Коши.

Теорема 8. (необходимое условие интегрируемости). Если функция f интегрируема по Риману на $[a, b]$, то она ограничена.

Доказательство:

По критерию интегрируемости существует разбиение $p = \{x_i\}_{i=0}^n$ отрезка $[a, b]$ такие, что

$$\Omega(f, p) < 1.$$

Отсюда следует, что для каждого i выполнено $\omega(f, [x_{i-1}, x_i]) \leq \frac{1}{\Delta x_i}$, значит f ограничена на $[x_{i-1}, x_i]$. Тогда f ограничена на $[a, b] = \bigcup_{i=1}^n [x_{i-1}, x_i]$.

Теорема 9. (достаточное условие интегрируемости). Если функция f непрерывна на отрезке $[a, b]$, то она интегрируема по Риману на $[a, b]$. По теореме Кантора f равномерно непрерывна на $[a, b]$, а значит для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что $|f(x) - f(y)| < \varepsilon$ при $|x - y| < \delta$. Тогда для любого разбиения $p = \{x_i\}_{i=0}^n$ отрезка $[a, b]$ при $\ell(p) < \delta$ для каждого i выполнено

$$\omega(f, [x_{i-1}, x_i]) \leq \varepsilon$$

Значит

$$\Omega(f, p) \leq \varepsilon(b - a)$$

откуда ясно, что $f \in \mathcal{R}[a, b]$.

Теорема 10.. Если функция f монотонна на отрезке $[a, b]$, то она интегрируема по Риману на $[a, b]$

Доказательство:

Не уменьшая общности будем считать, что f не убывает, причём $f(a) < f(b)$. Тогда для любого разбиения $p = \{x_i\}_{i=0}^n$ для каждого i выполнено

$$\omega(f, [x_{i-1}, x_i]) = f(x_i) - f(x_{i-1})$$

$$\Omega(f, p) = \sum_{i=1}^n (f(x_i) - f(x_{i-1})) \Delta x_i \leq \ell(p) \sum_{i=1}^n (f(x_i) - f(x_{i-1})) = \ell(p)(f(b) - f(a))$$

Отсюда видно, что для любого $\varepsilon > 0$ существует $\delta = \varepsilon/(f(b) - f(a))$ такое, что из $\ell(p) < \delta$ следует $\Omega(f, p) < \varepsilon$. Значит $f \in \mathcal{R}[a, b]$ по критерию интегрируемости.

Отметим, что из монотонности функции f на интервале (a, b) интегрируемость f на отрезке $[a, b]$ вообще говоря не следует. Например, $f(x) = 1/x$ монотонна на $(0, 1)$, но независимо от значений в концах $[0, 1]$ она не может быть интегрируема на $[0, 1]$, так как неограничена.

Теорема 11.. Если $f, g \in \mathcal{R}[a, b]$, то $f \cdot g \in \mathcal{R}[a, b]$

Доказательство:

Так как $f, g \in \mathcal{R}[a, b]$, то существует константа $M > 0$ такая, что $|f| \leq M$ и $|g| \leq M$ на $[a, b]$. Тогда для любых $x, y \in [a, b]$

$$|f(x)g(x) - f(y)g(y)| \leq |f(x)||g(x) - g(y)| + |f(x) - f(y)||g(y)|$$

значит для любого $I \subset [a, b]$

$$\omega(fg, I) \leq M(\omega(f, I) + \omega(g, I))$$

Тогда для любого разбиения p отрезка $[a, b]$

$$\Omega(fg, p) \leq M(\Omega(f, p) + \Omega(g, p)).$$

Значит по критерию интегрируемости $fg \in \mathcal{R}[a, b]$.

Следствие 12. Если $f \in \mathcal{R}[a, b]$, то для любого отрезка $[c, d] \subset [a, b]$ (где $c \leq d$) выполнено $1_{[c, d]}f \in \mathcal{R}[a, b]$, причём

$$\int_a^b 1_{[c, d]} f dx = \int_c^d f dx$$

(Аналогичное верно для полуинтервалов $[c, d), (c, d]$ и интервала (c, d) .)

Доказательство:

Так как любое разбиение отрезка $[c, d]$ можно дополнить до разбиения $[a, b]$, то из интегрируемости

f на $[a, b]$ следует интегрируемость f на $[c, d]$ (например, можно воспользоваться критерием интегрируемости, или критерием Коши).

Так как $f \in \mathcal{R}[a, b]$ и $1_{[c, d]} \in \mathcal{R}[a, b]$, то $1_{[c, d]}f \in \mathcal{R}[a, b]$.

Для вычисления $\int_a^b 1_{[c, d]}f dx$ можно использовать любую последовательность разбиений $[a, b]$, мелкость которых стремится к нулю. Если дополнительно потребовать, чтобы точка c была одним из узлов разбиения, то соответствующие интегральные суммы будут совпадать с интегральными суммами для f на $[c, d]$. Переходя к пределу, получим искомое равенство.

Аддитивность интеграла Римана

Утверждение 13 (линейность интеграла Римана).

Если $f, g \in \mathcal{R}[a, b]$, то $f + g \in \mathcal{R}[a, b]$ и

$$\int_a^b (f + g) dx = \int_a^b f dx + \int_a^b g dx$$

Доказательство:

Так как $\omega(f + g, I) \leq \omega(f, I) + \omega(g, I)$, то по критерию интегрируемости $f + g \in \mathcal{R}[a, b]$. Тогда для любой последовательности разбиений p_n с $\ell(p_n) \rightarrow 0$ при $n \rightarrow \infty$ и для любых выборок ξ_n из p_n

$$\sigma(f + g, p_n, \xi_n) = \sigma(f, p_n, \xi_n) + \sigma(g, p_n, \xi_n)$$

Утверждение 14 (аддитивность интеграла Римана).

Если $c \in (a, b)$ и $f \in \mathcal{R}[a, c] \cap \mathcal{R}[c, b]$, то $f \in \mathcal{R}[a, b]$ и

$$\int_a^b f dx = \int_a^c f dx + \int_c^b f dx$$

Доказательство:

Функции $1_{[a, c]}f$ и $1_{(c, b]}f$ интегрируемы по Риману (как произведение интегрируемых по Риману функций). Тогда в силу линейности определённого интеграла получаем

$$\int_a^b f dx = \int_a^b 1_{[a, c]}f dx + \int_a^b 1_{(c, b]}f dx$$

и остаётся дважды воспользоваться равенством (2).

Утверждение 15. Если f ограничена на $[a, b]$ и для любых $c, d \in (a, b)$ выполнено $f \in \mathcal{R}[c, d]$, то $f \in \mathcal{R}[a, b]$

Доказательство:

Пусть $\delta_1 > 0$ и $p = \{x_i\}_{i=0}^n$ - разбиение $[a, b]$ с мелкостью $\ell(p) < \delta_1$. Обозначим $c := a + \delta_1$ и $d := b - \delta_1$ (при достаточно малом δ_1 выполнено $c < d$).

Обозначим

$$\begin{aligned} I_1 &:= \{i : [x_{i-1}, x_i] \subset [c, d]\}, \\ I_2 &:= \{i : [x_{i-1}, x_i] \not\subset [c, d]\}. \end{aligned}$$

Так как $f \in \mathcal{R}[c, d]$, то для любого $\varepsilon > 0$ найдётся $\delta_2 > 0$ такое, что при $\ell(p) < \delta_2$ выполнено

$$\sum_{i \in I_1} \omega(f, [x_{i-1}, x_i]) \Delta x_i < \varepsilon/2$$

С другой стороны, так как по условию существует константа M такая, что $|f| \leq M$ на $[a, b]$, то

$$\sum_{i \in I_1} \omega(f, [x_{i-1}, x_i]) \Delta x_i \leq 4M(\delta_1 + \ell(p))$$

Таким образом, при $\delta_1 < \varepsilon/(16M)$ и $\ell(p) < \min(\delta_1, \delta_2)$ получаем

$$\Omega(f, p) < \varepsilon$$

По критерию интегрируемости получаем $f \in \mathcal{R}[a, b]$.

Следствие 16. Если f ограничена на $[a, b]$ и множество всех точек разрыва функции f конечно, то $f \in \mathcal{R}[a, b]$.

Отметим, что интегрируемая по Риману функция может иметь бесконечное множество точек разрыва. Например, функция Римана

$$f(x) = \begin{cases} \frac{1}{q}, & x = p/q \\ 0, & \text{иначе} \end{cases}$$

а также функции 1_A (где $A = \{1/n\}_{n \in \mathbb{N}}$) и 1_C (где C - множество Кантора) интегрируемы по Риману, но имеют бесконечно много точек разрыва.

МА-34. Определённый интеграл Римана (продолжение)

Интегрирование неравенств и первая теорема о среднем

Теорема 1. Пусть $f, g \in \mathcal{R}[a, b]$ и $f \leq g$ на $[a, b]$. Тогда

$$\int_a^b f dx \leq \int_a^b g dx$$

Если, кроме того, f и g непрерывны в точке $x_0 \in [a, b]$ и $f(x_0) < g(x_0)$, то

$$\int_a^b f dx < \int_a^b g dx$$

Доказательство:

Рассмотрим последовательность разбиений p_n отрезка $[a, b]$ такую, что $\ell(p_n) \rightarrow 0$ при $n \rightarrow \infty$, а также последовательность выборок ξ_n из p_n . Так как $f \leq g$, то для каждого $n \in \mathbb{N}$ выполнено

$$\sigma(f, p_n, \xi_n) \leq \sigma(g, p_n, \xi_n)$$

Переходя в этом неравенстве к пределу $n \rightarrow \infty$ получим (1).

Для доказательства (2) достаточно показать, что $\int_a^b (g - f) dx > 0$. Для этого заметим, что $(g - f)(x_0) > 0$, поэтому ввиду непрерывности найдётся $\delta > 0$ такое, что

$$g(x) - f(x) \geq \frac{1}{2} (g(x_0) - f(x_0))$$

для всех $x \in [x_0 - \delta, x_0 + \delta] \cap [a, b]$. Иначе говоря, для всех $x \in [a, b]$

$$g(x) - f(x) \geq \frac{1}{2} (g(x_0) - f(x_0)) 1_{[x_0 - \delta, x_0 + \delta]}(x)$$

Интегрируя данное неравенство с помощью уже доказанной части теоремы, при достаточно малых δ получим

$$\int_a^b (g(x) - f(x)) dx \geq \frac{\delta}{2} (g(x_0) - f(x_0)) > 0$$

Теорема 2. (первая теорема о среднем). Пусть $f \in C[a, b]$ и $g \in \mathcal{R}[a, b]$, причём $g \geq 0$. Тогда существует $\xi \in [a, b]$ такое, то

$$\int_a^b f(x)g(x)dx = f(\xi) \int_a^b g(x)dx$$

Доказательство:

Отметим, что функция f интегрируема по Риману, так как непрерывна на $[a, b]$. Тогда произведение fg интегрируемо по Риману как произведение интегрируемых по Риману функций.

Так как функция f непрерывна, то

$$m := \inf_{[a,b]} f \quad \text{и} \quad M := \sup_{[a,b]} f$$

конечны и неотрицательны. Так как $g \geq 0$, то из неравенства $m \leq f \leq M$ следует, что

$$mg(x) \leq f(x)g(x) \leq Mg(x) \quad \forall x \in [a, b]$$

Интегрируя данное неравенство получим

$$m \int_a^b g dx \leq \int_a^b f(x)g(x) dx \leq M \int_a^b g dx.$$

Значит существует $\lambda \in [m, M]$ такое, что

$$\int_a^b f(x)g(x) dx = \lambda \int_a^b g(x) dx$$

По теореме о промежуточных значениях существует $\xi \in [a, b]$ такое, что $f(\xi) = \lambda$.

Теорема 3. (теорема о дифференцировании интеграла по верхнему пределу.)

Пусть $f \in C[a, b]$. Тогда функция

$$\Phi(x) = \int_a^x f(t) dt$$

дифференцируема на $[a, b]$ и для каждого $x \in [a, b]$ выполнено

$$\Phi'(x) = f(x)$$

(при $x = a$ или b имеется в виду соответствующая односторонняя производная.)

Доказательство:

Докажем, что для каждого $x \in [a, b]$ существует

$$\Phi'_+(x) = f(x)$$

По теореме о среднем для каждого $h > 0$ (при достаточно малых h) существует $\xi \in [x, x+h]$ такое, что

$$\frac{\Phi(x+h) - \Phi(x)}{h} = \frac{1}{h} \int_x^{x+h} f(t) dt = f(\xi)$$

Пусть $\varepsilon > 0$. Так как f непрерывна в x , то существует $\delta > 0$ такое, что при $|\xi - x| < \delta$ выполнено $|f(\xi) - f(x)| < \varepsilon$. Тогда

$$\left| \frac{\Phi(x+h) - \Phi(x)}{h} - f(x) \right| = |f(\xi) - f(x)| < \varepsilon$$

при $0 < h < \delta$. Отсюда по определению правосторонней производной получаем $\exists \Phi'_+(x) = f(x)$

Для левосторонних производных доказательство аналогично.

Определение 4. Функция F называется первообразной функции f , заданной на промежутке $I \subset \mathbb{R}$, если в каждой точке этого промежутка существует $F'(x) = f(x)$ (если x является одним из концов I , то F' понимается как соответствующая односторонняя производная).

Теорема 5. Если F и Φ - первообразные функции f на промежутке I , то существует константа $C \in \mathbb{R}$ такая, что $F(x) = \Phi(x) + C$ для всех $x \in I$. Пусть $x_0 \in I$. Тогда по теореме Лагранжа для любого $x \in I$ существует $\xi \in I$ такое, что

$$\begin{aligned} (F(x) - \Phi(x)) - (F(x_0) - \Phi(x_0)) &= (F'(\xi) - \Phi'(\xi))(x - x_0) \\ &= (f(\xi) - f(\xi))(x - x_0) = 0. \end{aligned}$$

Значит достаточно взять $C := F(x_0) - \Phi(x_0)$.

Теорема 6. (формула Ньютона-Лейбница). Если F является первообразной непрерывной на отрезке $[a, b]$ функции f , то

$$\int_a^b f(x)dx = F(b) - F(a).$$

Доказательство:

По теореме 3 функция

$$\Phi(x) := \int_a^x f(x)dx$$

является одной из первообразных функции f . По теореме 5 существует константа C такая, что $F = \Phi + C$. Тогда

$$F(b) - F(a) = \Phi(b) - \Phi(a) = \int_a^b f(x)dx$$

по определению Φ .

Следствие 7. Для любой непрерывно дифференцируемой на $[a, b]$ функции F выполнено

$$\int_a^b F'(x)dx = F(b) - F(a).$$

Отметим, что непрерывность функции f существенна в теореме 6. Например, функция

$$F(x) = \begin{cases} 0, & x \leq 0 \\ x^2 \sin \frac{1}{x^2}, & x > 0 \end{cases}$$

дифференцируема в каждой точке $x \in \mathbb{R}$, но её производная не интегрируема по Риману на $[0,1]$, так как неограничена на этом отрезке. Тем не менее, если дополнительно потребовать интегрируемость производной, то формула (4) будет справедлива:

Теорема 8. (обобщённая формула Ньютона-Лейбница). Если функция f интегрируема по Риману на отрезке $[a, b]$ и имеет на этом отрезке первообразную F , то справедлива формула Ньютона-Лейбница (3).

Доказательство:

Пусть $p = \{x_i\}_{i=0}^n$ - разбиение $[a, b]$. Тогда по теореме Лагранжа для каждого $i \in \overline{0, n}$ существует $\xi_i \in [x_{i-1}, x_i]$ такое, что

$$F(x_i) - F(x_{i-1}) = F'(\xi_i) \Delta x_i = f(\xi_i) \Delta x_i$$

Тогда

$$F(b) - F(a) = \sum_{i=1}^n (F(x_i) - F(x_{i-1})) = \sigma(f, p, \xi)$$

где $\xi = \{\xi_i\}_{i=1}^n$. В силу интегрируемости по Риману функции f для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что при $\ell(p) < \delta$ имеем $|\sigma(f, p, \xi) - \int_a^b f dx| < \varepsilon$. Тогда для любого $\varepsilon > 0$ выполнено $|F(b) - F(a) - \int_a^b f dx| < \varepsilon$, значит $F(b) - F(a) - \int_a^b f dx = 0$

Теорема 9. (о замене переменной). Пусть функция $g : [\alpha, \beta]$ непрерывно дифференцируема, а f непрерывна на $g([\alpha, \beta])$. Тогда

$$\int_{\alpha}^{\beta} f(g(t))g'(t)dt = \int_{g(\alpha)}^{g(\beta)} f(x)dx$$

Доказательство:

По теореме о промежуточных значениях существует отрезок $[a, b]$ такой, что $f([\alpha, \beta]) = [a, b]$. Пусть F - первообразная функции f на $[a, b]$ (в силу непрерывности f по теореме о дифференцировании интеграла по верхнему пределу можно определить $F(y) := \int_a^y f(t)dt$.) По теореме о производной сложной функции

$$(F(g(t)))' = f(g(t))g'(t)$$

значит по формуле Ньютона-Лейбница

$$\int_{\alpha}^{\beta} f(g(t))g'(t)dt = \int_{\alpha}^{\beta} (F(g(t)))'dt = F(g(\beta)) - F(g(\alpha))$$

и, с другой стороны,

$$F(g(\beta)) - F(g(\alpha)) = \int_{g(\alpha)}^{g(\beta)} F'(x)dx$$

Теорема 10. (интегрирование по частям). Пусть функции $f, g \in C^1[a, b]$. Тогда

$$\int_a^b f(x)g'(x)dx = (fg)|_a^b - \int_a^b f'(x)g(x)dx,$$

где

$$(fg)|_a^b = f(b)g(b) - f(a)g(a)$$

Доказательство:

Проинтегрировав правило Лейбница

$$(f(x)g(x))' = f'(x)g(x) + f(x)g'(x)$$

получим

$$(fg)|_a^b = \int_a^b f'(x)g(x)dx + \int_a^b f(x)g'(x)dx.$$

Вторая теорема о среднем

Теорема 11. (вторая теорема о среднем). Пусть $f, g \in \mathcal{R}[a, b]$, причём g монотонна. Тогда существует $\xi \in [a, b]$ такое, что

$$\int_a^b f(x)g(x)dx = g(a) \int_a^\xi f(x)dx + g(b) \int_\xi^b f(x)dx.$$

Доказательство (в случае когда f и g непрерывно дифференцируемы)

Если f и g непрерывно дифференцируемы, то рассмотрим

$$F(x) := \int_a^x f(t)dt$$

$$\int_a^b F'(x)g(x)dx = F(b)g(b) - \int_a^b F(x)g'(x)dx$$

Так как g монотонна, то g' знакопостоянна. Тогда по первой теореме о среднем существует $\xi \in [a, b]$ такое, что

$$\int_a^b F(x)g'(x)dx = F(\xi) \int_a^b g'(x)dx = F(\xi)(g(b) - g(a)).$$

Подставив это в (5) получим

$$\int_a^b F'(x)g(x)dx = g(a)F(\xi) + g(b)(F(b) - F(\xi))$$

что и требовалось доказать.

Неравенства Юнга, Гёльдера и Минковского

Ранее было доказано неравенство Юнга: для любого $p > 1$ и любых чисел $a, b \geq 0$ выполнено

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

где $q = p/(p-1)$ (иначе говоря, $\frac{1}{p} + \frac{1}{q} = 1$).

При $p \in (0, +\infty)$ для любой $f \in \mathcal{R}[a, b]$ определим

$$\|f\|_p := \left(\int_a^b |f|^p dx \right)^{1/p}$$

Ясно, что для любого $\alpha \in \mathbb{R}$ выполнено

$$\|\alpha f\|_p = |\alpha| \|f\|_p$$

Кроме того, если $f \in C[a, b]$ то $\|f\|_p = 0$ тогда и только тогда, когда $f = 0$.

Отметим, что неравенство треугольника

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p$$

может нарушаться при $p \in (0, 1)$. Например, пусть $a = 0, b = 1, f = 1_{[0, 1/2]}$ и $g = 1_{(1/2, 1]}$. Тогда $f + g = 1_{[0, 1]}$, поэтому

$$\|f\|_p + \|g\|_p = \frac{1}{\sqrt[p]{2}} + \frac{1}{\sqrt[p]{2}} = 2^{1-\frac{1}{p}} < 2^0 = 1 = \|f + g\|_p$$

Теорема 12. (неравенство Гёльдера). Пусть $f, g \in C[a, b], p \in (1, +\infty)$ $\frac{1}{p} + \frac{1}{q} = 1$. Тогда

$$\int_a^b f(x)g(x)dx \leq \|f\|_p \|g\|_q$$

Доказательство

Если $\|f\|_p = 0$ или $\|g\|_q = 0$, то $f \cdot g = 0$, поэтому достаточно рассмотреть случай $\|f\|_p > 0$ и $\|g\|_q > 0$. Определим

$$u = |f|/\|f\|_p, \quad v = |g|/\|g\|_q$$

Так как $\int_a^b f g dx \leq \int_a^b |f| |g| dx$, то достаточно доказать, что

$$\int_a^b u v dx \leq 1$$

В силу неравенства (6) имеем $uv \leq \frac{u^p}{p} + \frac{v^q}{q}$. Проинтегрируем это неравенство:

$$\int_a^b u v dx \leq \frac{1}{p} \|u\|_p^p + \frac{1}{q} \|v\|_q^q = 1$$

что и требовалось доказать.

МА-35. Несобственный интеграл Римана

Пусть вещественная функция f задана на промежутке $I \subset \mathbb{R}$.

Определение 1.

1. Если для любого отрезка $J \subset I$ функция f интегрируема по Риману на J , то f называется локально интегрируемой (по Риману) на I .
2. Точка $x \in I \cup \{\inf I, \sup I\}$ называется особенностью функции f , если $x \notin I$ или для любого $\delta > 0$ функция f не интегрируема по Риману на $[x - \delta, x + \delta] \cap \bar{I}$.

Например, если $I = (0, +\infty)$ и

$$f(x) = \begin{cases} \frac{1}{\ln x}, & x \neq 1 \\ 0, & x = 1 \end{cases}$$

то особенностями будут точки $x = 0$, $x = 1$ и $x = +\infty$.

Далее будем считать, что f локально интегрируема на $I = [a, b)$, где $b \in (a, +\infty]$. В этом случае b является единственной особенностью рассматриваемой функции.

Определение 2. Предел (конечный или бесконечный)

$$\int_a^{\rightarrow b} f(x)dx := \lim_{y \rightarrow b-0} F(y), \quad F(y) := \int_a^y f(x)dx$$

называется несобственным интегралом Римана от функции f по промежутку $I = [a, b)$ (с особенностью в точке b .) Если несобственный интеграл конечен, то говорят, что он сходится.

При этом обычный интеграл Римана (по отрезку) называется собственным. Покажем, что если функция интегрируема в собственном смысле, то она интегрируема и в несобственном смысле, причём интегралы в смысле этих двух определений совпадают.

Утверждение 3. Если $f \in \mathcal{R}[a, b]$, то существует

$$\int_a^{\rightarrow b} f(x)dx = \int_a^b f(x)dx.$$

Так как $f \in \mathcal{R}[a, b]$, то f ограничена на $[a, b]$ некоторой константой $C > 0$. Тогда в силу аддитивности интеграла Римана

$$\left| \int_a^b f(x)dx - F(y) \right| = \left| \int_y^b f(x)dx \right| \leq \int_y^b |f(x)|dx \leq C(b - y) \rightarrow 0$$

при $y \rightarrow b$

Ввиду утверждения 3 для обозначения несобственного интеграла можно не выделять особенность стрелкой, то есть вместо $\int_a^{\rightarrow b}$ можно писать \int_a^b .

Утверждение 4 (о замене переменной в несобственном интеграле). Пусть $f : [a, b) \rightarrow \mathbb{R}$ непрерывна и $g : [\alpha, \beta) \rightarrow [a, b)$ непрерывно дифференцируема, причём $g(\alpha) = a$ и $g(t) \rightarrow b$ при $t \rightarrow \beta$. Тогда

$$\int_a^b f(x)dx = \int_\alpha^\beta f(g(t))g'(t)dt$$

в том смысле, что если одна из частей (1) существует (конечная или бесконечная), то существует и другая и выполнено равенство (1).

Доказательство:

Обозначим

$$F(y) = \int_a^y f(x)dx, \quad \Phi(t) = \int_\alpha^t f(g(s))g'(s)ds$$

По теореме о замене переменной в собственном интеграле Римана для любого $t \in [\alpha, \beta)$

$$\Phi(t) = F(g(t))$$

Если существует предел $\lim_{y \rightarrow b} F(y) = A$, то по теореме о замене переменной под знаком предела существует и $\lim_{t \rightarrow \beta} F(g(t)) = A$. Обратно, пусть существует предел $\lim_{t \rightarrow \beta} \Phi(t) = B$. Тогда для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что для всех $t \in \dot{U}_\delta^-(\beta)$

$$F(g(t)) \in U_\varepsilon(B)$$

Так как g непрерывна и $g(t) \rightarrow b$ при $t \rightarrow \beta$, то по теореме о промежуточных значениях существует $r > 0$ такое, что

$$g\left(\dot{U}_\delta^-(\beta)\right) \supset \dot{U}_r^-(b)$$

Тогда для всех $x \in \dot{U}_r^-(b)$ выполнено

$$F(x) \in U_\varepsilon(B)$$

что и требовалось доказать.

Утверждение 5. (аддитивность несобственного интеграла). Если f локально интегрируема на $[a, b)$, то для любого $c \in (a, b)$

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx$$

в том смысле что если один из несобственных интегралов $\int_a^b f(x)dx$ и $\int_c^b f(x)dx$ существует, то существует и другой и имеет место записанное выше равенство.

Доказательство:

Данная теорема непосредственно следует из аналогичной теоремы для собственного интеграла Римана и предельного перехода по верхнему пределу интегрирования.

Утверждение 6. (линейность несобственного интеграла). Если f, g локально

интегрируемы на $[a, b)$ и несобственный интеграл $\int_a^b g dx$ сходится, то

$$\int_a^b (f + g) dx = \int_a^b f dx + \int_a^b g dx$$

в том смысле что если несобственный интеграл $\int_a^b f dx$ существует (конечный или бесконечный), то существует и несобственный интеграл $\int_a^b (f + g) dx$, причём имеет место записанное выше равенство.

Доказательство:

Данная теорема непосредственно следует из аналогичной теоремы для собственного интеграла Римана и предельного перехода по верхнему пределу интегрирования.

Определение 7. Если функции f и g локально интегрируемы на $[a, b)$ и $[\alpha, \beta)$ соответственно и несобственные интегралы $\int_a^b f dx$ и $\int_\alpha^\beta g dx$ сходятся или расходятся одновременно, то говорят, что эти интегралы эквивалентны в смысле сходимости. Пишут

$$\int_a^b f dx \overset{\text{cx}}{\sim} \int_\alpha^\beta g dx$$

Если $\alpha = a$ и $\beta = b$, то пишут также $f \overset{\text{cx}}{\sim} g$

Следствие 8.. Если f локально интегрируема на $[a, b)$, то для любого $c \in (a, b)$

$$\int_a^b f dx \overset{\text{ccx}}{\sim} \int_c^b f dx$$

Следствие 9.. Если f, g локально интегрируемы на $[a, b)$ и несобственный интеграл $\int_a^b g dx$ сходится, то

$$\int_a^b (f + g) dx \overset{\text{cx}}{\sim} \int_a^b f dx$$

Теорема 10. (критерий Коши). Для сходимости интеграла $\int_a^b f dx$ необходимо и достаточно, чтобы для любого $\varepsilon > 0$ существовало $\beta \in (a, b)$ такое, что для всех $s, t \in (\beta, b)$ выполнялось

$$\left| \int_s^t f dx \right| < \varepsilon$$

Доказательство:

Данная теорема непосредственно следует из критерия Коши существования конечного предела функции $F(y) = \int_a^y f(x) dx$.

Теорема 11. (признак сравнения). Пусть функции f и g локально интегрируемы на $[a, b)$, причём

$$0 \leq f \leq g$$

Тогда

$$\int_a^b f(x)dx \leq \int_a^b g(x)dx$$

В частности,

1. Если $\int_a^b g dx$ сходится, то и $\int_a^b f dx$ сходится;
2. Если $\int_a^b f dx$ расходится, то и $\int_a^b g dx$ расходится.

Доказательство:

Функции $y \mapsto \int_a^y f(x)dx$ и $y \mapsto \int_a^y g(x)dx$ нестрого возрастают ввиду неотрицательности f и g . Значит интегралы $\int_a^b f(x)dx$ и $\int_a^b g(x)dx$ существуют (конечные или бесконечные). Интегрируя неравенство $f \leq g$ по отрезку $[a, y]$ и переходя к пределу $y \rightarrow b$ получаем (2).

Определение 12. Пусть f локально интегрируема на $[a, b)$.

- Если несобственный интеграл $\int_a^b |f|dx$ сходится, то говорят, что несобственный интеграл $\int_a^b f dx$ сходится абсолютно.
- Если несобственный интеграл $\int_a^b |f|dx$ сходится, то говорят, что несобственный интеграл $\int_a^b f dx$ сходится условно.

Утверждение 13. Если $\int_a^b f dx$ сходится абсолютно, то он сходится. При этом

$$\left| \int_a^b f dx \right| \leq \int_a^b |f| dx$$

Доказательство:

Так как при $s \leq t$

$$\left| \int_s^t f dx \right| \leq \int_s^t |f| dx$$

то по критерию Коши из сходимости $\int_a^b |f|dx$ следует сходимость $\int_a^b f dx$. Так как функция $t \mapsto |t|$ непрерывна, то переходя к пределу $y \rightarrow b$ в неравенстве $\left| \int_a^y f dx \right| \leq \int_a^y |f| dx$ (только одна) из следующих трёх ситуаций:

1. интеграл сходится абсолютно; 2. интеграл сходится условно; 3. интеграл расходится.

Утверждение 14 (асимптотический признак сравнения). Если f, g локально

интегрируемы на $[a, b)$ и $f(x) = O(g(x))$ при $x \rightarrow b$, то из абсолютной сходимости интеграла $\int_a^b g dx$ следует абсолютная сходимость интеграла $\int_a^b f dx$.

Доказательство:

Так как $f(x) = O(g(x))$ при $x \rightarrow b$, то существуют константа $C > 0$ и число $x_0 \in (a, b)$ такие, что $|f(x)| \leq C|g(x)|$ при $x \in [x_0, b)$. Тогда интеграл $\int_a^b |f|dx \stackrel{\text{ex}}{\sim} \int_{x_0}^b |f|dx$ сходится по признаку сравнения.

Теорема 15. (признак Дирихле). Рассмотрим несобственный интеграл

$$I = \int_a^b f(x)g(x)dx$$

где функции f и g локально интегрируемы на $[a, b)$. Пусть

1. f непрерывна на $[a, b)$, а g непрерывно дифференцируема на $[a, b)$;
2. g нестрого убывает на $[a, b)$ и $g(x) \rightarrow 0$ при $x \rightarrow b$,
3. существует $M > 0$ такое, что для любого $x \in [a, b)$

$$\left| \int_a^x f(t)dt \right| \leq M$$

Тогда несобственный интеграл I сходится.

Доказательство:

Обозначим $F(x) := \int_a^x f(t)dt$. По второй теореме о среднем при $a \leq s \leq t < b$ существует $\xi \in [s, t]$ такое, что

$$\int_s^t f(x)g(x)dx = g(s)(F(\xi) - F(s)) + g(t)(F(t) - F(\xi))$$

По условию существует константа $M > 0$ такая, что $|F(x)| \leq M$ для всех $x \in [a, b)$. Тогда

$$\left| \int_s^t f(x)g(x)dx \right| \leq 2M(g(s) + g(t))$$

Так как $g(t) \rightarrow 0$ при $t \rightarrow b$, то по критерию Коши отсюда заключаем, что I сходится.

Теорема 16. (признак Абеля). Пусть функция f непрерывна на $[a, b)$, а функция g непрерывно дифференцируема на $[a, b)$. Если интеграл $\int_a^b f dx$ сходится, а функция g монотонна на $[a, b)$ и имеет конечный предел $A = \lim_{x \rightarrow b} g(x)$, то интеграл $\int_a^b f(x)g(x)dx$ сходится.

Доказательство:

Не уменьшая общности пусть g нестрого убывает. Тогда

$$\int_a^b f(x)g(x)dx = \int_a^b f(x)(g(x) - A)dx + A \int_a^b f(x)dx$$

так как несобственные интегралы в правой части сходятся: первый сходится по признаку Дирихле, а второй по условию.

«Табличные» несобственные интегралы

Утверждение 17.

Интеграл

$$\int_1^{+\infty} \frac{dx}{x^\alpha}$$

сходится тогда и только тогда, когда $\alpha > 1$.

Доказательство

$$F(y) = \int_1^y \frac{dx}{x^\alpha} = \begin{cases} \frac{y^{1-\alpha}-1}{1-\alpha}, & \alpha \neq 1 \\ \ln y, & \alpha = 1 \end{cases}$$

Отсюда видно, что для существования конечного предела $\lim_{y \rightarrow +\infty} F(y)$ необходимо и достаточно чтобы $1 - \alpha < 0$.

$$\int_0^1 \frac{dx}{x^\alpha}$$

сходится тогда и только тогда, когда $\alpha < 1$.

Доказательство:

Для доказательства можно сделать замену $x = 1/t$ и воспользоваться предыдущим утверждением.

Утверждение 19.

Интеграл

$$I = \int_2^{+\infty} \frac{dx}{x^\alpha (\ln x)^\beta}$$

сходится тогда и только тогда, когда

$$\begin{cases} \alpha > 1 \\ \alpha = 1, \quad \beta > 1 \end{cases}$$

Доказательство:

Предположим что $\alpha = 1 + 2\varepsilon$, где $\varepsilon > 0$. Так как $x^\varepsilon (\ln x)^\beta \rightarrow +\infty$ при $x \rightarrow +\infty$ для любого $\beta \in \mathbb{R}$ (в этом можно убедиться, например, с помощью правила Лопиталья), то

$$\frac{1}{x^\alpha (\ln x)^\beta} = \frac{1}{x^{1+\varepsilon}} \frac{1}{x^\varepsilon (\ln x)^\beta} = O\left(\frac{1}{x^{1+\varepsilon}}\right), \quad x \rightarrow +\infty$$

Так как $\int_1^{+\infty} \frac{dx}{x^{1+\varepsilon}} < +\infty$, то по признаку сходимости интеграл I сходится.

Предположим, что $\alpha = 1 - 2\varepsilon$, где $\varepsilon > 0$. Так как $x^\varepsilon (\ln x)^{-\beta} \rightarrow +\infty$ при $x \rightarrow \infty$ для любого $\beta \in \mathbb{R}$, то

$$\frac{1}{x^{1-\varepsilon}} = \frac{1}{x^{1-2\varepsilon} (\ln x)^\beta} \frac{1}{x^\varepsilon (\ln x)^{-\beta}} = O\left(\frac{1}{x^\alpha (\ln x)^\beta}\right), \quad x \rightarrow +\infty$$

Так как $\int_1^{+\infty} \frac{dx}{x^{1-\varepsilon}} = +\infty$, то по признаку сравнения I расходится.

Осталось рассмотреть случай $\alpha = 1$. В этом случае по теореме о замене переменной в несобственном интеграле

$$I = \int_{\ln 2}^{+\infty} \frac{dt}{t^\beta}$$

Утверждение 20.

Интеграл

$$I = \int_1^{+\infty} \frac{\sin x}{x^\alpha}$$

сходится тогда и только тогда, когда $\alpha > 0$. При $\alpha > 1$ данный интеграл сходится абсолютно, а при $0 < \alpha \leq 1$ условно.

Доказательство:

При $\alpha \leq 0$ интеграл I расходится по критерию Коши. Действительно, для любого $n \in \mathbb{N}$ при $2\pi n + \pi/4 \leq x \leq 2\pi n + 3\pi/4$ выполнено $\sin x \geq 1/2$ и $1/x^\alpha \geq 1$, поэтому

$$\int_{2\pi n + \pi/4}^{2\pi n + 3\pi/4} \frac{\sin x}{x^\alpha} dx \geq \frac{\pi}{4}$$

Пусть $\alpha > 0$. Так как

$$\int_1^x \sin x dx = \cos 1 - \cos x$$

по модулю не превосходит 2, а функция $1/x^\alpha$ убывает на $(1, +\infty)$ и $1/x^\alpha \rightarrow 0$ при $x \rightarrow +\infty$, то при $\alpha > 0$ интеграл I сходится по признаку Дирихле.

При $\alpha > 1$ выполнено

$$\left| \frac{\sin x}{x^\alpha} \right| \leq \frac{1}{x^\alpha}$$

значит по признаку сравнения I сходится абсолютно.

Наконец, при $0 < \alpha \leq 1$ имеем

$$\frac{|\sin x|}{x^\alpha} \geq \frac{\sin^2 x}{x^\alpha} = \frac{1}{2x^\alpha} - \frac{\cos 2x}{2x^\alpha}$$

Так как $\int_1^{+\infty} \frac{\cos 2x}{x^\alpha} dx$ сходится по признаку Дирихле, а интеграл $\int_1^{+\infty} \frac{1}{x^\alpha} dx$ расходится (так как $0 < \alpha \leq 1$). Значит $\int_1^{+\infty} \frac{\sin^2 x}{x^\alpha} dx$ расходится. Тогда $\int_1^{+\infty} \frac{|\sin x|}{x^\alpha} dx$ расходится по признаку сравнения.

МА-36. Знакопостоянные числовые ряды

Определение 1. Пусть $\{a_k\}_{k \in \mathbb{N}}$ - последовательность вещественных или комплексных чисел. Выражение

$$\sum_{k=1}^{\infty} a_k$$

называется числовым рядом (с общим членом a_k). Числа

$$A_n = \sum_{k=1}^n a_k$$

называются частичными суммами ряда (1). Если существует предел $A = \lim_{n \rightarrow \infty} A_n$, то он называется суммой ряда (1), пишут $\sum_{k=1}^{\infty} a_k = A$. (В случае когда все a_k вещественны, допускаются значения $A = +\infty$ или $-\infty$).

Если последовательность частичных сумм ряда (1) имеет конечный предел, то говорят, что данный ряд сходится. В противном случае говорят, что данный ряд расходится.

Легко видеть, что изменение любого конечного числа членов ряда не влияет на его сходимость. Кроме того, если ряд (1) сходится, то для каждого $n \in \mathbb{N}$ ряд

$$R_n := \sum_{k=n+1}^{\infty} a_k$$

тоже сходится, причём в силу арифметических свойств предела $S = A_n + R_n$. По этой причине ряд R_n называется n -ым остатком ряда (1).

Пример 2. Пусть $q \in \mathbb{C}$. Ряд $\sum_{k=1}^{\infty} q^k$ сходится тогда и только тогда, когда $|q| < 1$.

Доказательство:

Действительно, при $q = 1$ имеем $\sum_{k=1}^n q^k = n$ откуда ясно, что ряд расходится. При $q \neq 1$

$$\sum_{k=1}^n q^k = q \frac{q^n - 1}{q - 1}$$

Теорема 3. (критерий Коши сходимости числового ряда). Ряд $\sum_{k=1}^{\infty} a_k$ сходится тогда и только тогда, когда для любого $\varepsilon > 0$ существует $N \in \mathbb{N}$ такое, что для всех $n \geq N$ и всех $p \in \mathbb{N}$ выполнено

$$\left| \sum_{k=n+1}^{n+p} a_k \right| < \varepsilon.$$

Доказательство:

Достаточно применить критерий Коши сходимости последовательности для частичных сумм рассматриваемого ряда.

Следствие 4. (необходимое условие сходимости числового ряда). Если ряд $\sum_{k=1}^{\infty} a_k$ сходится, то $\lim_{k \rightarrow \infty} a_k = 0$.

Доказательство:

Достаточно записать (2) с $p = 1$.

Пример 5. Ряд $\sum_{k=1}^{\infty} (-1)^k$ расходится.

Доказательство:

Так как $(-1)^k \not\rightarrow 0$ при $k \rightarrow \infty$, то не выполнено необходимое условие сходимости ряда.

Определение 6. Если ряд $\sum_{k=1}^{\infty} |a_k|$ сходится, то говорят, что ряд (1) сходится абсолютно.

Покажем, что всякий абсолютно сходящийся ряд сходится. Попутно мы обобщим неравенство треугольника на бесконечное число слагаемых.

Утверждение 7. Если ряд (1) сходится абсолютно, то он сходится. При этом

$$\left| \sum_{k=1}^{\infty} a_k \right| \leq \sum_{k=1}^{\infty} |a_k|.$$

Доказательство:

Пусть рассматриваемый ряд сходится абсолютно. Тогда по критерию Коши для любого $\varepsilon > 0$ существует $N \in \mathbb{N}$ такое, что для всех натуральных $n \geq N$ и $p \geq 1$ выполнено

$$\sum_{k=n+1}^{n+p} |a_k| < \varepsilon$$

Тогда

$$\left| \sum_{k=n+1}^{n+p} a_k \right| \leq \sum_{k=n+1}^{n+p} |a_k| < \varepsilon,$$

значит рассматриваемый ряд сходится по критерию Коши.

Так как функция $t \mapsto |t|$ непрерывна и для любого $n \in \mathbb{N}$ выполнено

$$\left| \sum_{k=1}^n a_k \right| \leq \sum_{k=1}^n |a_k|$$

то переходя в данном неравенстве к пределу при $n \rightarrow \infty$ получим (3).

Из сходимости ряда $\sum_{k=1}^{\infty} a_k$ не следует, что ряд $\sum_{k=1}^{\infty} |a_k|$ сходится. Например, если элементы данного ряда имеют вид

$$a_k = \frac{1}{2}, \frac{-1}{2}, \frac{1}{4}, \frac{-1}{4}, \frac{1}{4}, \frac{-1}{4}, \frac{1}{8}, \frac{-1}{8}, \frac{1}{8}, \frac{-1}{8}, \frac{1}{8}, \frac{-1}{8}, \frac{1}{8}, \frac{-1}{8}, \dots,$$

то $\sum_{k=1}^{\infty} a_k = 0$, но $\sum_{k=1}^{\infty} |a_k| = +\infty$.

Определение 8. Если ряд $\sum_{k=1}^{\infty} a_k$ сходится, но не сходится абсолютно, то говорят, что данный ряд сходится условно.

Таким образом, любой числовой ряд либо сходится абсолютно, либо сходится условно, либо расходится.

Теорема 9. (1-й признак сравнения). Пусть $0 \leq a_k \leq b_k$ для всех $k \in \mathbb{N}$. Тогда $\sum_{k=1}^{\infty} a_k \leq \sum_{k=1}^{\infty} b_k$. В частности,

1. если ряд $\sum_{k=1}^{\infty} b_k$ сходится, то ряд $\sum_{k=1}^{\infty} a_k$ тоже сходится;
2. если ряд $\sum_{k=1}^{\infty} a_k$ расходится, то ряд $\sum_{k=1}^{\infty} b_k$ тоже расходится.

Доказательство:

Так как a_k и b_k неотрицательны, то последовательности частичных сумм $A_n = \sum_{k=1}^n a_k$ и $B_n = \sum_{k=1}^n b_k$ нестрого возрастают и, следовательно, по теореме Вейерштрасса имеют пределы $\sum_{k=1}^{\infty} a_k = A$ и $B = \sum_{k=1}^{\infty} b_k$ соответственно (конечные или равные $+\infty$). Так как $a_k \leq b_k$ для всех k , то $A_n \leq B_n$ для всех $n \in \mathbb{N}$. Отсюда по теореме о видно, что если $B < +\infty$, то и $A < +\infty$. Аналогично, если $A = +\infty$, то и $B = +\infty$.

Теорема 10. (2-й признак сравнения). Пусть $b_k \geq 0$ и $a_k = O(b_k)$ при $k \rightarrow \infty$. Если ряд $\sum_{k=1}^{\infty} b_k$ сходится, то ряд $\sum_{k=1}^{\infty} a_k$ сходится абсолютно.

Доказательство:

Так как существуют константа $C > 0$ и число $N > 0$ такие, что для всех $k \geq N$

$$|a_k| \leq C b_k$$

то по первому признаку сравнения

$$\sum_{k=N}^{\infty} |a_k| \leq C \sum_{k=N}^{\infty} b_k < +\infty,$$

следовательно $\sum_{k=1}^{\infty} |a_k| < +\infty$.

Теорема 11. (интегральный критерий сходимости). Пусть на $[1, +\infty)$ задана неотрицательная нестрого убывающая функция f . Тогда для сходимости ряда $\sum_{k=1}^{\infty} f(k)$ необходимо и достаточно, чтобы несобственный интеграл $\int_1^{+\infty} f(x)dx$ сходится.

Доказательство

Интегрируя неравенство

$$f(k+1) \leq f(x) \leq f(k) \quad \forall x \in [k, k+1]$$

по x получим

$$f(k+1) \leq \int_k^{k+1} f(x)dx \leq f(k)$$

Просуммировав по k от 1 до n получаем

$$A_{n+1} - f(1) \leq \int_1^{n+1} f(x)dx \leq A_n$$

где $A_n = \sum_{k=1}^n f(k)$. Так как $f \geq 0$ то последовательность A_n неотрицательна и, следовательно, по теореме Вейерштрасса имеет предел (конечный или равный $+\infty$). Аналогично, функция $A(y) = \int_1^y f(x)dx$ нестрого возрастает и, следовательно, имеет предел при $y \rightarrow +\infty$. Поэтому переходя в (4) к пределу при $n \rightarrow \infty$ получаем

$$\sum_{k=1}^{\infty} f(k) \leq f(1) + \int_1^{+\infty} f(x)dx, \quad \int_1^{+\infty} f(x)dx \leq \sum_{k=1}^{\infty} f(k).$$

Отсюда видно, что $\int_1^{+\infty} f(x)dx < +\infty$ тогда и только тогда, когда $\sum_{k=1}^{\infty} f(k) < +\infty$, что и требовалось доказать.

Отметим, что в интегральном критерии сходимости условие нестрого убывания функции f можно заменить на монотонность, так как в случае нестрогого возрастания ряд $\sum_{k=1}^{\infty} f(k)$ и интеграл $\int_1^{+\infty} f(x)dx$ расходятся.

Пример 12. Ряд $\sum_{k=1}^{\infty} \frac{1}{k^\alpha}$ сходится тогда и только тогда, когда $\alpha > 1$.

Доказательство:

Действительно, так как $f(x) = \frac{1}{x^\alpha}$ монотонна, то по интегральному критерию сходимости

$$\sum_{k=1}^{\infty} \frac{1}{k^\alpha} \overset{\text{cx}}{\sim} \int_1^{+\infty} \frac{1}{x^\alpha} dx$$

Интеграл в правой части сходится тогда и только тогда, когда $\alpha > 1$.

Выше мы видели, что исследование сходимости ряда особенно просто в случае геометрической прогрессии. Естественнo ожидать, что исследование сходимости рядов, «близких» к геометрической прогрессии, можно свести к исследованию сходимости последней. Так как для геометрической прогрессии $a_k = q^{k-1}a_1$ выполнено

$$\frac{a_{k+1}}{a_k} = q, \quad \sqrt[k]{a_k} = q \sqrt[k]{a_1/q}$$

то «близость» ряда a_k к геометрической прогрессии естественно понимать в следующем смысле: при достаточно больших k выполнено $\frac{a_{k+1}}{a_k} \approx q$ или $\sqrt[k]{a_k} \approx q$. В первом случае эти наводящие соображения приводят нас к признаку Даламбера, а во втором - к признаку Коши.

Теорема 13. (признак Даламбера). Пусть $a_k > 0$ для всех $k \in \mathbb{N}$ и существует $\lim_{k \rightarrow \infty} \frac{a_{k+1}}{a_k} = q \in [0, +\infty]$. Тогда

1. если $q < 1$, то ряд $\sum_{k=1}^{\infty} a_k$ сходится;
2. если $q > 1$, то ряд $\sum_{k=1}^{\infty} a_k$ расходится; Пусть $q < 1$. Возьмём $r \in (q, 1)$. Тогда при достаточно больших k выполнено

$$a_{k+1} \leq r a_k$$

Следовательно, $a_k \leq r^{k-1}a_1$. Так как $\sum_{k=1}^{\infty} r^k < +\infty$, то ряд $\sum_{k=1}^{\infty} a_k$ сходится по признаку сравнения.

Пусть $q > 1$. Возьмём $r \in (1, q)$. Тогда при достаточно больших k выполнено

$$a_{k+1} \geq r a_k$$

Следовательно, $a_k \geq r^{k-1} a_1$. Так как $\sum_{k=1}^{\infty} r^k = +\infty$, то ряд $\sum_{k=1}^{\infty} a_k$ расходится по признаку сравнения.

Пример 14. Если $a_k = \frac{k!}{k^k}$, то $\lim_{k \rightarrow \infty} \frac{a_{k+1}}{a_k} = \frac{1}{e} < 1$. Значит ряд $\sum_{k=1}^{\infty} a_k$ сходится (по признаку Даламбера).

Теорема 15. (признак Коши). Пусть $a_k \geq 0$ для всех $k \in \mathbb{N}$ и $\overline{\lim}_{k \rightarrow \infty} \sqrt[k]{a_k} = q \in [0, +\infty]$. Тогда

1. если $q < 1$, то ряд $\sum_{k=1}^{\infty} a_k$ сходится;
2. если $q > 1$, то ряд $\sum_{k=1}^{\infty} a_k$ расходится;

Доказательство:

Пусть $q < 1$. Тогда для любого $r \in (q, 1)$ при достаточно больших k выполнено

$$a_k \leq r^k$$

значит $\sum_{k=1}^{\infty} a_k$ сходится по признаку сравнения.

Пусть $q > 1$. Тогда существуют $r \in (1, q)$ и подпоследовательность a_{n_k} такие, что

$$\sqrt[n_k]{a_{n_k}} \rightarrow r$$

при $k \rightarrow \infty$. Тогда $a_{n_k} \not\rightarrow 0$ при $k \rightarrow \infty$ и, следовательно, $a_n \not\rightarrow 0$ при $n \rightarrow \infty$. Значит $\sum_{k=1}^{\infty} a_k$ расходится.

Пример 16. Если $a_k = (1 - 1/k)^{k^2}$, то $\lim_{k \rightarrow \infty} \sqrt[k]{a_k} = \frac{1}{e} < 1$. Значит ряд $\sum_{k=1}^{\infty} a_k$ сходится (по признаку Коши). Если в признаке Даламбера или в признаке Коши $q = 1$, то ряд $\sum_{k=1}^{\infty} a_k$ может как сходиться, так и расходиться. Например, для любого α при $a_k = 1/k^\alpha$ имеем

$$\frac{a_{k+1}}{a_k} = \left(\frac{k}{k+1} \right)^\alpha \rightarrow 1, \quad \sqrt[k]{a_k} = (\sqrt[k]{k})^\alpha \rightarrow 1$$

при $k \rightarrow \infty$. Но ряд $\sum_{k=1}^{\infty} 1/k^\alpha$ сходится при $\alpha > 1$ и расходится при $\alpha \leq 1$.

Отметим, что признак Коши сильнее признака Даламбера в следующем смысле: если сходимость ряда можно доказать с помощью признака Даламбера, то можно её доказать и с помощью признака Коши. Действительно, если $\lim_{k \rightarrow \infty} \frac{a_{k+1}}{a_k} = q \in [0, 1)$, то для любого $r \in (q, 1)$ при достаточно больших k выполнено $a_{k+1} \leq r a_k$, следовательно $\sqrt[k]{a_k} \leq r \sqrt[k]{a_1/r} \rightarrow r$ при $k \rightarrow \infty$. В то же время, сходимость ряда с общим членом

$$a_k = 2^{-k} \cdot \begin{cases} 1, & k = 2m - 1 \\ m, & k = 2m \end{cases}$$

можно непосредственно доказать с помощью признака Коши, но не нельзя непосредственно доказать с помощью признака Даламбера.

МА-37. Знакопеременные числовые ряды

Для сумм вида $\sum_{k=1}^n a_k b_k$ имеется аналог интегрирования по частям, известный под названием «преобразование Абеля»:

Лемма 1 (преобразование Абеля). Для любых вещественных или комплексных чисел A_0, A_1, \dots, A_n и b_1, \dots, b_n, b_{n+1} выполнено

$$\sum_{k=1}^n (A_k - A_{k-1}) b_k = \sum_{k=1}^n A_k (b_k - b_{k+1}) + A_n b_{n+1} - A_0 b_1.$$

Доказательство:

Действительно,

$$\begin{aligned} \sum_{k=1}^n (A_k - A_{k-1}) b_k &= \sum_{k=1}^n A_k b_k - \sum_{k=1}^n A_{k-1} b_k \\ &= \sum_{k=1}^n A_k b_k - \sum_{\ell=0}^{n-1} A_\ell b_{\ell+1} \\ &= \sum_{k=1}^n A_k b_k - \sum_{\ell=1}^n A_\ell b_{\ell+1} + A_n b_{n+1} - A_0 b_1 \\ &= \sum_{k=1}^n A_k (b_k - b_{k+1}) + A_n b_{n+1} - A_0 b_1, \end{aligned}$$

что и требовалось доказать.

Теорема 2. (признак Дирихле). Если частичные суммы ряда $\sum_{k=1}^{\infty} a_k$ ограничены, а последовательность b_k монотонно стремится к нулю при $k \rightarrow \infty$, то ряд $\sum_{k=1}^{\infty} a_k b_k$ сходится. При этом

$$\left| \sum_{k=1}^{\infty} a_k b_k \right| \leq C |b_1|$$

где $C = \sup_n \left| \sum_{k=1}^n a_k \right|$. Не уменьшая общности пусть b_k нестрого убывает (иначе можно заменить b_k на $-b_k$).

Тогда $b_k \geq 0$ для всех $k \in \mathbb{N}$. Пользуясь преобразованием Абеля запишем

$$S_n = \sum_{k=1}^n a_k b_k = \sum_{k=1}^n A_k (b_k - b_{k+1}) + A_n b_{n+1}$$

где $A_n = \sum_{k=1}^n a_k$ и $A_0 = 0$.

Заметим, что для каждого $k \in \mathbb{N}$ выполнено

$$|A_k(b_k - b_{k+1})| \leq C(b_k - b_{k+1})$$

и $\sum_{k=1}^n (b_k - b_{k+1}) = b_1 - b_{n+1} \rightarrow b_1$ при $n \rightarrow \infty$. Тогда по признаку сравнения ряд

$$\sum_{k=1}^{\infty} A_k(b_k - b_{k+1})$$

сходится абсолютно, причём

$$\left| \sum_{k=1}^{\infty} A_k(b_k - b_{k+1}) \right| \leq C \sum_{k=1}^{\infty} (b_k - b_{k+1}) = Cb_1$$

С другой стороны, так как $\sup_n |A_n| = C$, то $|A_n b_{n+1}| \leq Cb_{n+1} \rightarrow 0$ при $n \rightarrow \infty$. Поэтому из (2) следует, что последовательность S_n имеет конечный предел. Переходя к пределу в (2) при $n \rightarrow \infty$ получаем (1).

Теорема 3. (признак Лейбница). Если последовательность b_n монотонно сходится к нулю, то ряд $\sum_{k=1}^{\infty} (-1)^k b_k$ сходится.

Доказательство:

Обозначим $A_n = \sum_{k=1}^n (-1)^k$, тогда $|A_n| \leq 1$ для всех n , следовательно рассматриваемый ряд сходится по признаку Дирихле.

Например, ряд $\sum_{k=1}^{\infty} \frac{(-1)^k}{k}$ сходится по признаку Лейбница, причём сходимость является условной, так как ряд $\sum_{k=1}^{\infty} 1/k$ расходится.

Лемма 4. Если $x \notin \pi\mathbb{Z}$, то для любого $n \in \mathbb{N}$

$$\left| \sum_{k=1}^n \sin kx \right| \leq \frac{1}{\left| \sin \frac{x}{2} \right|}, \quad \left| \sum_{k=1}^n \cos kx \right| \leq \frac{1}{\left| \sin \frac{x}{2} \right|}$$

$$2 \sin \frac{x}{2} \sin(kx) = \cos \left(k - \frac{1}{2} \right) x - \cos \left(k + \frac{1}{2} \right) x$$

то

$$2 \sin \frac{x}{2} \sum_{k=1}^n \sin kx = \cos \frac{1}{2}x - \cos \left(n + \frac{1}{2} \right) x$$

откуда получаем первое неравенство в (4). Второе неравенство аналогичным образом следует из равенства $2 \sin \frac{x}{2} \cos(kx) = \sin \left(k + \frac{1}{2} \right) x - \sin \left(k - \frac{1}{2} \right) x$.

Другой способ доказательства заключается в использовании формулы Эйлера и формулы суммы геометрической прогрессии:

$$\sum_{k=0}^n e^{ikx} = \frac{e^{i(n+1)x} - 1}{e^{ix} - 1} = \frac{e^{i(n+\frac{1}{2})x} - e^{i\frac{x}{2}}}{e^{i\frac{x}{2}} - e^{-i\frac{x}{2}}}$$

Пример 5. Ряд

$$S = \sum_{k=1}^{\infty} \frac{\sin kx}{k^\alpha}, \quad x \notin \pi\mathbb{Z},$$

сходится абсолютно при $\alpha > 1$, сходится условно при $0 < \alpha \leq 1$ и расходится при $\alpha \leq 0$

Доказательство:

Докажем, что при $\alpha \leq 0$ ряд S расходится. Для этого покажем, что его общий член не стремится к нулю при $k \rightarrow \infty$. В противном случае $\sin kx = k^\alpha o(1) = o(1)$ при $k \rightarrow \infty$, следовательно

$$\cos kx = \frac{\sin(k+1)x - \cos kx \sin x}{\sin x} \rightarrow 0$$

при $k \rightarrow \infty$. Тогда $1 = \sin^2 kx + \cos^2 kx \rightarrow 0$ при $k \rightarrow \infty$, что невозможно.

Теперь докажем, что при $\alpha > 0$ ряд S сходится. В силу (4) частичные суммы ряда $\sum_{k=1}^{\infty} \sin kx$ ограничены некоторой константой, зависящей только от x .

Последовательность $1/k^\alpha$ монотонно сходится к нулю. Значит по признаку Дирихле S сходится. Так как $|\sin kx|/k^\alpha \leq 1/k^\alpha$ и ряд $\sum_{k=1}^{\infty} 1/k^\alpha$ сходится, то при $\alpha > 1$ ряд S сходится абсолютно (по признаку сравнения).

Остаётся показать, что при $0 < \alpha \leq 1$ ряд

$$\tilde{S} = \sum_{k=1}^{\infty} \frac{|\sin kx|}{k^\alpha}$$

расходится. Заметим, что $|\sin kx| \geq \sin^2 kx = \frac{1}{2}(1 - \cos 2kx)$. При этом ряд $\sum_{k=1}^{\infty} \frac{\cos 2kx}{k^\alpha}$ сходится по признаку Дирихле, а ряд $\sum_{k=1}^{\infty} 1/k^\alpha$ расходится. Значит ряд $\sum_{k=1}^{\infty} (1 - \cos 2kx)/k^\alpha$ расходится. Следовательно, ряд \tilde{S} расходится по признаку сравнения.

Теорема 6. (признак Абеля). Если ряд $\sum_{k=1}^{\infty} a_k$ сходится, а последовательность b_k монотонно сходится к числу $b \in \mathbb{R}$, то ряд $\sum_{k=1}^{\infty} a_k b_k$ сходится.

Доказательство:

Ряд $\sum_{k=1}^{\infty} a_k (b_k - b)$ сходится по признаку Дирихле, а ряд $\sum_{k=1}^{\infty} a_k$ сходится по условию. Значит

$$\sum_{k=1}^{\infty} a_k b_k = \sum_{k=1}^{\infty} a_k (b_k - b) + b \sum_{k=1}^{\infty} a_k$$

Перестановки знакопеременных рядов Как известно, сумма конечного числа слагаемых не зависит от их порядка. Покажем, что в случае бесконечного числа слагаемых это в общем случае не так. Для этого рассмотрим ряд

$$S = \frac{1}{1} - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \dots = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k}$$

Частичные суммы этого ряда равны

$$S_n = \sum_{k=1}^n \int_0^1 (-x)^{k-1} dx = \int_0^1 \sum_{k=1}^n (-x)^{k-1} dx = \int_0^1 \frac{1 + (-x)^n}{1+x} dx.$$

$$\left| \int_0^1 \frac{(-x)^n}{1+x} dx \right| \leq \int_0^1 x^n dx = \frac{1}{n+1} \rightarrow 0$$

при $n \rightarrow \infty$. По формуле Ньютона-Лейбница

$$\int_0^1 \frac{dx}{1+x} = \ln 2$$

Отсюда заключаем, что $S = \ln 2$.

Заметим, что в ряде S чередуются положительные и отрицательные слагаемые. Переставим члены S так, чтобы после каждого положительного слагаемого шло два отрицательных:

$$\tilde{S} := \frac{1}{1} - \frac{1}{2} - \frac{1}{4} + \frac{1}{3} - \frac{1}{6} - \frac{1}{8} + \frac{1}{5} - \frac{1}{10} - \frac{1}{12} + \dots$$

Приведём к общему знаменателю каждую пару из положительного слагаемого и следующего сразу после него отрицательного:

$$\begin{aligned} \tilde{S} &= \frac{1}{2} - \frac{1}{4} + \frac{1}{6} - \frac{1}{8} + \frac{1}{10} - \frac{1}{12} + \dots \\ &= \frac{1}{2} \left(\frac{1}{1} - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \dots \right) = \frac{1}{2} S = \frac{1}{2} \ln 2. \end{aligned}$$

Таким образом, в результате изменения порядка слагаемых сумма ряда уменьшилась в два раза.

Оказывается, что для абсолютно сходящихся рядов сумма не зависит от порядка слагаемых:

Теорема 7.. Для любой числовой последовательности $\{a_k\}$ и любой биекции $f: \mathbb{N} \rightarrow \mathbb{N}$

$$\sum_{k=1}^{\infty} |a_{f(k)}| = \sum_{k=1}^{\infty} |a_k|$$

Если ряд $\sum_{k=1}^{\infty} a_k$ сходится абсолютно, то ряд $\sum_{k=1}^{\infty} a_{f(k)}$ тоже сходится абсолютно, причём

$$\sum_{k=1}^{\infty} a_{f(k)} = \sum_{k=1}^{\infty} a_k.$$

Для доказательства равенства (5) достаточно проверить, что для ряда $b_k = |a_k|$ имеет место неравенство

$$\sum_{k=1}^{\infty} b_{f(k)} \leq \sum_{k=1}^{\infty} b_k$$

Действительно, если применить (7) к ряду $\tilde{b}_k = b_{f(k)}$ и биекции $\tilde{f} = f^{-1}$, то получим

$$\sum_{k=1}^{\infty} b_k = \sum_{k=1}^{\infty} \tilde{b}_{\tilde{f}(k)} \leq \sum_{k=1}^{\infty} \tilde{b}_k = \sum_{k=1}^{\infty} b_{f(k)}$$

Пусть $M := \sum_{k=1}^{\infty} b_k$. Если $M = +\infty$, то неравенство (7) очевидно. Пусть $M < +\infty$, $n \in \mathbb{N}$ и $m = \max(f(1), \dots, f(n))$. Так как f — биекция, то каждое из чисел $f(1), \dots, f(n)$ входит в $1, \dots, m$, причём ровно один раз. Поэтому

$$\sum_{k=1}^n b_{f(k)} \leq \sum_{\ell=1}^m b_{\ell} \leq M$$

Так как $b_k \geq 0$, то отсюда заключаем $\sum_{k=1}^{\infty} b_{f(k)} \leq M$, что и требовалось доказать.

Пусть $\sum_{k=1}^{\infty} |a_k| < +\infty$. В силу доказанного тогда и $\sum_{\ell=1}^{\infty} |a_{f(\ell)}| < +\infty$. Обозначим $m_n := \max(f(1), \dots, f(n))$. Ясно, что $m_n \rightarrow \infty$ при $n \rightarrow \infty$. Если $\ell \in \{1, \dots, m_n\}$ и $\ell \notin \{f(1), \dots, f(n)\}$, то $f^{-1}(\ell) \geq n+1$. Поэтому

$$\sum_{\ell=1}^{m_n} a_{\ell} = \sum_{k=1}^n a_{f(k)} + O\left(\sum_{k=n+1}^{\infty} |a_{f(k)}|\right) = \sum_{k=1}^n a_{f(k)} + o(1)$$

при $n \rightarrow \infty$. Переходя к пределу при $n \rightarrow \infty$ получаем (6).

Выше мы уже видели, что сумма условно сходящегося ряда может измениться, если изменить его порядок слагаемых. Оказывается, что слагаемые любого условно сходящегося ряда всегда можно переставить так, чтобы получилось любая заданная наперёд сумма (конечная или равная плюс или минус бесконечности):

Теорема 8. (Риман). Если ряд $\sum_{k=1}^{\infty} a_k$ сходится условно, то для любого $S \in \overline{\mathbb{R}}$ существует биекция $f: \mathbb{N} \rightarrow \mathbb{N}$ такая, что существует $\sum_{k=1}^{\infty} a_{f(k)} = S$.

Доказательство:

Для простоты будем считать, что $S \in \mathbb{R}$ и $a_k \neq 0$ для всех k . Так как $\sum_{k=1}^{\infty} a_k$ сходится, то $a_k \rightarrow 0$ при $k \rightarrow \infty$. Обозначим $a_k^+ = \max(a_k, 0)$ и $a_k^- = \max(-a_k, 0)$. Тогда $a_k = a_k^+ - a_k^-$ и, следовательно,

$$\sum_{k=1}^{\infty} a_k^+ = \sum_{k=1}^{\infty} a_k^- = +\infty$$

Действительно, если обе суммы конечны, то $\sum_{k=1}^{\infty} |a_k| < +\infty$, что противоречит условию. Если же конечна только одна из них, то $\sum_{k=1}^{\infty} a_k = \sum_{k=1}^{\infty} a_k^+ - \sum_{k=1}^{\infty} a_k^-$ расходится, что также противоречит условию.

Пусть $\{\alpha_1, \alpha_2, \dots\}$ и $\{\beta_1, \beta_2, \dots\}$ — соответственно все положительные и все отрицательные члены последовательности $\{a_1, a_2, \dots\}$.

В силу (8)

$$\sum_{k=1}^{\infty} \alpha_k = +\infty, \quad \sum_{k=1}^{\infty} \beta_k = -\infty$$

Так как $a_k \rightarrow 0$ при $k \rightarrow \infty$, то и $\alpha_k \rightarrow 0$ и $\beta_k \rightarrow 0$ при $k \rightarrow \infty$.

Пусть N_1 - наименьшее натуральное или равное нулю число такое, что $\sum_{k=1}^{N_1} \alpha_k \geq S$.

Пусть M_1 - наименьшее натуральное или равное нулю число такое, что $\sum_{k=1}^{N_1} \alpha_k + \sum_{k=1}^{M_1} \beta_k < S$

При $k > 1$ определим N_k и M_k соответственно как наименьшие натуральные числа, превосходящие N_{k-1} и M_{k-1} соответственно, такие что

$$\sum_{j=1}^{N_k} \alpha_j + \sum_{j=1}^{M_{k-1}} \beta_j \geq S, \quad \sum_{j=1}^{N_k} \alpha_j + \sum_{j=1}^{M_k} \beta_j < S.$$

Рассмотрим последовательность

$$\alpha_1, \dots, \alpha_{N_1}, \beta_1, \dots, \beta_{M_1}, \alpha_{N_1+1}, \dots, \alpha_{N_2}, \beta_{M_1+1}, \dots, \beta_{M_2}, \dots$$

Обозначим её k -й член как u_k . По построению существует биекция $f : \mathbb{N} \rightarrow \mathbb{N}$ такая, что для каждого $k \in \mathbb{N}$ выполнено $u_k = a_{f(k)}$.

$$r_k = \begin{cases} |\beta_{M_1}|, & N_1 + M_1 \leq k < N_2 + M_1 \\ \alpha_{N_2}, & N_2 + M_1 \leq k < N_2 + M_2 \\ |\beta_{M_2}|, & N_2 + M_2 \leq k < N_3 + M_2 \\ \alpha_{N_3}, & N_3 + M_2 \leq k < N_3 + M_3 \\ \dots \end{cases}$$

Так как $\alpha_k \rightarrow 0$ и $\beta_k \rightarrow 0$ при $k \rightarrow \infty$, то $r_k \rightarrow 0$ при $k \rightarrow \infty$.

МА-38. Поточечная и равномерная сходимость функциональных последовательностей

Многие функции, возникающие в приложениях, задаются как суммы рядов вида $f(x) = \sum_{k=1}^{\infty} u_k(x)$. Такие функции часто требуется дифференцировать или интегрировать, но их редко удаётся выразить через элементарные функции. Так как исходная функция f задана как сумма ряда, то и её производную или интеграл естественно представить как сумму некоторого ряда. При этом естественно ожидать, что

$$\left(\sum_{k=1}^{\infty} u_k(x) \right)' = \sum_{k=1}^{\infty} u_k'(x)$$
$$\int_a^b \left(\sum_{k=1}^{\infty} u_k(x) \right) dx = \sum_{k=1}^{\infty} \int_a^b u_k(x) dx$$

или другими словами, что

$$\left(\lim_{n \rightarrow \infty} f_n \right)' = \left(\lim_{n \rightarrow \infty} f_n' \right)$$

и

$$\int_a^b \lim_{n \rightarrow \infty} f_n dx = \lim_{n \rightarrow \infty} \int_a^b f_n dx$$

где $f_n(x) = \sum_{k=1}^n u_k(x)$

Несложно проверить, что равенствам (1) и (2) удовлетворяет, например, функциональная последовательность $f_n(x) = e^{x/n}$. Однако ниже мы приведём примеры ситуаций, в которых равенства (1) и (2) не имеют места. Поэтому далее мы получим достаточные условия, при выполнении которых равенства (1) и (2) верны.

Поточечная сходимость функциональных последовательностей Пусть на множестве $E \subset \mathbb{R}$ для каждого $n \in \mathbb{N}$ задана функция $f_n : E \rightarrow \mathbb{R}$. Тогда говорят, что на E задана функциональная последовательность $\{f_n\}$. Пусть на E также задана функция $f : E \rightarrow \mathbb{R}$.

Определение 1. Будем говорить, что последовательность $\{f_n\}$ сходится к f поточечно на E (пишут $f = \lim_{n \rightarrow \infty} f_n$ на E), если для любого $x \in E$ выполнено $f_n(x) \rightarrow f(x)$ при $n \rightarrow \infty$. При этом f называется поточечным пределом функциональной последовательности $\{f_n\}$, а также предельной функцией.

Для обозначения поточечной сходимости $\{f_n\}$ к f на E также пишут $f_n \xrightarrow{E} f$ или $f_n \rightarrow f$ на E при $n \rightarrow \infty$. Кроме того, используется обозначение $f_n(x) \rightarrow f(x), x \in E$, при $n \rightarrow \infty$.

Из единственности предела числовой последовательности следует единственность поточечного предела функциональной последовательности. Действительно пусть $f_n \rightarrow f$ и $f_n \rightarrow g$ на E при $n \rightarrow \infty$.

Тогда для любого $x \in E$ имеем $f_n(x) \rightarrow f(x)$ и $f_n(x) \rightarrow g(x)$ при $n \rightarrow \infty$. В силу единственности предела числовой последовательности $f(x) = g(x)$. Это верно для любого $x \in E$, значит $f = g$ на E .

Рассмотрим на вещественной прямой $E = \mathbb{R}$ функциональную последовательность

$$f_n(x) = \frac{x}{1 + n^2 x^2}, \quad n \in \mathbb{N}$$

Легко видеть, что f_n сходится поточечно на E к функции

$$f \equiv 0$$

но последовательность производных f_n

$$f'_n(x) = \frac{1 - n^2 x^2}{(1 + n^2 x^2)^2}, \quad n \in \mathbb{N}$$

сходится поточечно на E к функции

$$g(x) = \begin{cases} 1, & x = 0 \\ 0, & x \neq 0 \end{cases}$$

Таким образом, равенство (1) может нарушаться даже если предельная функция f оказалось дифференцируемой.

Теперь рассмотрим на отрезке $[0,1]$ функциональную последовательность

$$f_n(x) = \begin{cases} (n+1)x^n, & 0 \leq x < 1 \\ 0, & x = 1 \end{cases}$$

Легко видеть, что $f_n \rightarrow 0$ (поточечно на $[0,1]$) при $n \rightarrow \infty$. Однако при этом

$$\int_0^1 f_n(t) dt = 1 \not\rightarrow 0, \quad n \rightarrow \infty$$

Таким образом, равенство (2) может нарушаться даже если предельная функция f оказалось интегрируемой.

Равномерная сходимость функциональных последовательностей

Определение 2. Будем говорить, что последовательность $\{f_n\}$ сходится к f равномерно на E , если для любого $\varepsilon > 0$ существует $N \in \mathbb{N}$ такое, что при $n > N$ для всех $x \in E$ выполнено

$$|f_n(x) - f(x)| < \varepsilon$$

Для обозначения равномерной сходимости $\{f_n\}$ к f на E также пишут $f_n \xrightarrow[E]{} f$ при $n \rightarrow \infty$ или $f_n \rightrightarrows f$ на E при $n \rightarrow \infty$. Кроме того, используется обозначение $f_n(x) \rightrightarrows f(x), x \in E$, при $n \rightarrow \infty$.

Для сравнения запишем в кванторах определение поточечной сходимости $f_n \rightarrow f$: для любого $x \in E$ для любого $\varepsilon > 0$ существует $N \in \mathbb{N}$ такое, что при $n > N$ выполнено (4). Таким образом в

случае равномерной сходимости число N может зависеть лишь от ε и не зависит от x , в то время как в случае поточечной сходимости число N может зависеть как от ε , так и от x .

Равномерная сходимость сильнее поточечной в том смысле, что последняя следует из первой:

Утверждение 3. Если $f_n \rightrightarrows f$ при $n \rightarrow \infty$, то $f_n \xrightarrow{E} f$ при $n \rightarrow \infty$.

Доказательство:

Пусть $\varepsilon > 0$. Тогда существует $N \in \mathbb{N}$ такое, что для всех $n > N$ для всех $x \in E$ выполнено $|f_n(x) - f(x)| < \varepsilon$. Значит для любого $x \in E$ при $n > N$ имеем $|f_n(x) - f(x)| < \varepsilon$. Значит $f_n \rightarrow f$ на E при $n \rightarrow \infty$.

Обратное в общем случае неверно, то есть равномерная сходимость не следует из поточечной. Например, последовательность (3) сходится поточечно к нулю, поэтому если она сходится к некоторой функции f равномерно на $[0,1]$, то по утверждению 3 имеем $f \equiv 0$. Допустим, что $f_n \rightrightarrows 0$ на $[0,1]$ при $n \rightarrow \infty$. Тогда существует $N \in \mathbb{N}$ такое, что при $n > N$ для всех $x \in E$ $|f_n(x)| < 1$. Следовательно при $n > N$ для всех $x \in [0,1]$ выполнено $(n+1)x^n < 1$. Зафиксировав n и перейдя к пределу $x \rightarrow 1-0$ получим $n+1 \leq 1$. Последнее неравенство не может выполняться для всех $n > N$.

Свойства равномерно сходящихся функциональных последовательностей

Определение 4. Функциональную последовательность $\{f_n\}$ будем называть равномерно ограниченной на E , если существует константа $C > 0$ такая, что для всех $n \in \mathbb{N}$ и всех $x \in E$ выполнено

$$|f_n(x)| \leq C$$

Как известно, всякая сходящаяся числовая последовательность ограничена. Однако если функциональная последовательность $\{f_n\}$ сходится равномерно на множестве E , то отсюда не следует что $\{f_n\}$ равномерно ограничена на E . Например, функциональная последовательность $f_n(x) := x + 2^{-n}$ сходится к $f(x) = x$ равномерно на любом множестве E . Но если E неограничено, то данная последовательность f_n не является равномерно ограниченной. Однако в общем случае если $f_n \rightrightarrows f$ при $n \rightarrow \infty$ и f ограничена на E , то последовательность f_n равномерно ограничена на E .

Покажем, что равномерная сходимость (как и поточечная) согласована с арифметическими операциями:

Утверждение 5.. Пусть $f_n \rightrightarrows f$ на E при $n \rightarrow \infty$ и $g_n \rightrightarrows g$ на E при $n \rightarrow \infty$. Тогда

1. $f_n + g_n \rightrightarrows f + g$ на E при $n \rightarrow \infty$;
2. если последовательности $\{f_n\}$ и $\{g_n\}$ равномерно ограничены на E , то также $f_n \cdot g_n \rightrightarrows f \cdot g$

Доказательство:

Первое утверждение следует из неравенства $|f_n + g_n - f - g| \leq |f_n - f| + |g_n - g|$, второе следует из неравенства $|f_n g_n - f g| \leq |(f_n - f) \cdot g_n| + |f \cdot (g_n - g)|$.

Покажем, что если последовательность непрерывных функций сходится равномерно, то её предел также является непрерывной функцией:

Теорема 6.. Если для каждого $n \in \mathbb{N}$ функция $f_n : E \rightarrow \mathbb{R}$ непрерывна на E и $f_n \rightrightarrows f$ на E при $n \rightarrow \infty$, то f непрерывна на E .

Доказательство:

Зафиксируем $x \in E$ и докажем, что f непрерывна в x . Пусть $\varepsilon > 0$.

Так как $f_n \rightrightarrows f$ при $n \rightarrow \infty$, то существует $N \in \mathbb{N}$ такое, что при $n > N$ для всех $x \in E$ выполнено

(4). Зафиксируем некоторое $n > N$. Тогда в силу непрерывности f_n в точке x существует $\delta > 0$ такое, что $|f_n(x) - f_n(y)| < \varepsilon$ для всех $y \in U_\delta(x) \cap E$. Тогда для всех таких y

$$|f(x) - f(y)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)| < 3\varepsilon$$

Аналогичным образом доказывается, что если последовательность интегрируемых по Риману функций сходится равномерно, то её предел также интегрируем по Риману. Кроме того, в случае равномерной сходимости можно переходить к пределу под знаком интеграла, то есть выполнено (2):

Теорема 7. (о перестановке предельного перехода и интегрирования). Пусть для каждого $n \in \mathbb{N}$ функция f_n интегрируема по Риману на $[a, b]$ и $f_n \Rightarrow f$ на $[a, b]$ $n \rightarrow \infty$. Тогда f интегрируема по Риману на $[a, b]$

$$\int_a^x f_n(t) dt \Rightarrow \int_a^x f(t) dt, \quad x \in [a, b]$$

при $n \rightarrow \infty$

Доказательство:

Пусть $\varepsilon > 0$. По определению равномерной сходимости существует $N \in \mathbb{N}$ такое, что при $n > N$ для всех $x \in [a, b]$.

$$|f_n(x) - f(x)| < \frac{\varepsilon}{2(b-a)}$$

Тогда для любого разбиения $p = \{x_i\}_{i=0}^m$ отрезка $[a, b]$ при $n > N$ выполнено $\Omega(f - f_n, p) < \varepsilon$. Зафиксируем $n > N$. Так как $f_n \in \mathcal{R}[a, b]$, то существует $\delta > 0$ такое, что $\Omega(f_n, p) < \varepsilon$. Тогда при $\ell(p) < \delta$

$$\Omega(f, p) \leq \Omega(f - f_n, p) + \Omega(f_n, p) < 2\varepsilon$$

Значит $f \in \mathcal{R}[a, b]$.

$$\left| \int_a^x f_n(t) dt - \int_a^x f(t) dt \right| \leq \int_a^x |f_n(t) - f(t)| dt \leq \int_a^b |f_n(t) - f(t)| dt < \varepsilon$$

откуда следует

Теперь докажем, что если производные последовательности функций сходятся равномерно, а последовательность функций сходится хотя бы в одной точке, то переход к пределу можно менять местами с дифференцированием, то есть выполнено (1):

Теорема 8. (о перестановке предельного перехода и дифференцирования). Пусть для каждого $n \in \mathbb{N}$ функция f_n непрерывно дифференцируема на $[a, b]$ и $f'_n \Rightarrow g$ на $[a, b]$ при $n \rightarrow \infty$. Пусть, кроме того, существует $x_0 \in [a, b]$ такое, что числовая последовательность $f_n(x_0)$ имеет конечный предел. Тогда существует непрерывно дифференцируемая функция $f : [a, b] \rightarrow \mathbb{R}$ такая, что $f_n \Rightarrow f$ на $[a, b]$ $n \rightarrow \infty$ и $f' = g$

Доказательство:

Пусть $A := \lim_{n \rightarrow \infty} f_n(x_0)$. Так как $f'_n \Rightarrow g$ на $[a, b]$, то по теореме 6 функция g непрерывна на $[a, b]$. Кроме того, по теореме 7

$$\int_{x_0}^x f'_n(t)dt \Rightarrow \int_{x_0}^x g(t)dt, \quad x \in [a, b]$$

при $n \rightarrow \infty$. Следовательно, последовательность f_n сходится равномерно на $[a, b]$ к функции

$$f(x) := A + \int_{x_0}^x g(t)dt$$

Так как g непрерывна на $[a, b]$, то по теореме о дифференцировании интеграла по верхнему пределу функция непрерывно дифференцируема и $f'(x) = g(x)$ для всех $x \in [a, b]$

МА-39. Исследование функциональных последовательностей и рядов на равномерную сходимость

Необходимые условия, признаки и критерии равномерной сходимости.

Пусть на множестве $E \subset \mathbb{R}$ для каждого $k \in \mathbb{N}$ задана функция $f_k : E \rightarrow \mathbb{R}$. Пусть также задана функция $f : E \rightarrow \mathbb{R}$.

Утверждение 1.. Для того, чтобы $f_n \rightrightarrows f$ на E при $n \rightarrow \infty$ необходимо и достаточно, чтобы

$$r_n := \sup_{x \in E} |f_n(x) - f(x)| \rightarrow 0$$

при $n \rightarrow \infty$

Доказательство:

Необходимость.

По определению равномерной сходимости для любого $\varepsilon > 0$ существует $N = N(\varepsilon)$ такое, что при $n > N(\varepsilon)$ для всех $x \in E$ выполнен $|f_n(x) - f(x)| < \varepsilon$. Тогда $\sup_{x \in E} |f_n(x) - f(x)| \leq \varepsilon$. Значит $r_n \rightarrow 0$ при $n \rightarrow \infty$.

Достаточность.

По определению предела для любого для любого $\varepsilon > 0$ существует $N = N(\varepsilon)$ такое, что при $n > N(\varepsilon)$ выполнено $r_n < \varepsilon$. Тогда по определению супремума для всех $x \in E$ $|f_n(x) - f(x)| \leq r_n < \varepsilon$. Значит $f_n \rightrightarrows f$ на E при $n \rightarrow \infty$

Утверждение 2. Для того, чтобы $f_n \rightrightarrows f$ на E при $n \rightarrow \infty$ необходимо и достаточно, чтобы для любой последовательности $\{x_n\}_{n \in \mathbb{N}} \subset E$ выполнялось

$$|f_n(x_n) - f(x_n)| \rightarrow 0$$

при $n \rightarrow \infty$

Доказательство:

Необходимость очевидна, так как по определению супремума для каждого $n \in \mathbb{N}$ для всех $x \in E$ выполнено $|f_n(x) - f(x)| \leq r_n$, где r_n определено в (1).

$$\rho_n := \min \left(\frac{1}{2} r_n, 1 \right)$$

По определению супремума для каждого $n \in \mathbb{N}$ существует $x_n \in E$ такое, что

$$|f_n(x_n) - f(x_n)| \geq \rho_n$$

Так как левая часть данного неравенства стремится к нулю при $n \rightarrow \infty$, то $\rho_n \rightarrow 0$ при $n \rightarrow \infty$. Следовательно, $r_n \rightarrow 0$ при $n \rightarrow \infty$.

Теорема 3. (критерий Коши сходимости функциональной последовательности). Для того, чтобы функциональная последовательность $\{f_n\}_{n \in \mathbb{N}}$ сходилась равномерно на множестве E необходимо и

достаточно, чтобы для любого $\varepsilon > 0$ существовало $N \in \mathbb{N}$ такое, что при $n > N$ для всех $p \in \mathbb{N}$ и всех $x \in E$ выполнялось

$$|f_{n+p}(x) - f_n(x)| < \varepsilon$$

Доказательство:

Необходимость.

Пусть $f_n \rightrightarrows f$ на E при $n \rightarrow \infty$. По определению равномерной сходимости для любого $\varepsilon > 0$ выберем N такое, что при $n > N$ для всех $x \in E$ выполнено $|f_n(x) - f(x)| < \varepsilon/2$. Тогда для всех $n, m > N$ и всех $x \in E$ выполнено

$$|f_n(x) - f_m(x)| \leq |f_n(x) - f(x)| + |f(x) - f_m(x)| < \varepsilon.$$

Достаточность.

Из условия (2) следует, что для каждого $x \in E$ числовая последовательность $\{f_n(x)\}_{n \in \mathbb{N}}$. Значит по критерию Коши для числовой последовательности существует конечный предел

$$f(x) := \lim_{n \rightarrow \infty} f_n(x)$$

Теперь покажем, что последовательность f_n сходится к f не только поточечно, но и равномерно на E . Для этого для каждого $x \in E$ перейдём в (2) к пределу при $p \rightarrow \infty$:

$$|f(x) - f_n(x)| \leq \varepsilon$$

Данное неравенство выполнено при $n > N$ для всех $x \in E$ (подчеркнём, что N не зависит от x). Значит $f_n \rightrightarrows f$ при $n \rightarrow \infty$.

Функциональные ряды

Пусть на множестве $E \subset \mathbb{R}$ для каждого $k \in \mathbb{N}$ задана функция $u_k : E \rightarrow \mathbb{R}$. Тогда говорят, что на множестве E задан функциональный ряд

$$S(x) := \sum_{k=1}^{\infty} u_k(x)$$

При этом для каждого $x \in E$ под суммой ряда $S(x)$ понимается предел последовательности угочастичных сумм

$$S_n(x) := \sum_{k=1}^n u_k(x)$$

при $n \rightarrow \infty$, если такой предел существует (конечный или равный $\pm\infty$).

Если для каждого $x \in E$ существует конечный предел $S(x)$, то говорят что ряд (3) сходится поточечно на E . В этом случае

$$R_n(x) := S(x) - S_n(x) = \sum_{k=n+1}^{\infty} u_k(x)$$

называется n -ым остатком ряда.

Утверждение 4. Для того, чтобы ряд (3) сходиллся равномерно на E необходимо и достаточно, чтобы он сходиллся поточечно и его остатки сходились к нулю равномерно на E .

Доказательство:

Пусть ряд сходится поточечно. Тогда последовательность $S_n = S - R_n$ сходится к S равномерно тогда и только тогда, когда R_n сходится к нулю равномерно.

Теорема 5. (критерий Коши равномерной сходимости функционального ряда). Для того, чтобы функциональный ряд $\sum_{k=1}^{\infty} u_k(x)$ сходиллся равномерно на множестве E необходимо и достаточно, чтобы для любого $\varepsilon > 0$ существовало $N \in \mathbb{N}$ такое, что для всех $n > N, p \in \mathbb{N}$ и $x \in E$ выполнялось

$$\left| \sum_{k=n+1}^{n+p} u_k(x) \right| < \varepsilon$$

Достаточно применить критерий Коши для функциональной последовательности $S_n(x) = \sum_{k=1}^n u_k(x)$

Следствие 6. Если существует последовательность $\{x_n\} \subset E$ такая, что

$$\sum_{k=n+1}^{2n} u_k(x_n) \not\rightarrow 0$$

при $n \rightarrow \infty$, то ряд (3) не сходится равномерно на множестве E .

Теорема 7. (признак Вейерштрасса). Пусть для всех $k \in \mathbb{N}$ и всех $x \in E$ выполнено неравенство

$$|u_k(x)| \leq b_k$$

причём числовой ряд $\sum_{k=1}^{\infty} b_k$ сходится. Тогда функциональный ряд $\sum_{k=1}^{\infty} u_k(x)$ сходится равномерно на E .

Доказательство:

По критерию Коши для каждого $\varepsilon > 0$ существует N такое, что при $n > N$ для всех $p \in \mathbb{N}$ выполнено

$$\sum_{k=n+1}^{n+p} b_k < \varepsilon.$$

Тогда в силу (6) для всех $x \in E$ выполнено (5), а значит ряд $\sum_{k=1}^{\infty} u_k(x)$ сходится равномерно на множестве E .

Отметим, что при выполнении неравенства (6) говорят, что последовательность $\{b_k\}$ мажорирует функциональную последовательность $\{u_k\}$ (на множестве E). При этом последовательность $\{b_k\}$ также называют мажорантой функциональной последовательности $\{u_k\}$.

МА-40. Признаки Дирихле и Абеля для функциональных рядов

Теорема 11. (признак Дирихле равномерной сходимости ряда). Если частичные суммы ряда $\sum_{k=1}^{\infty} a_k(x)$ равномерно ограничены на множестве E , последовательность $\{b_k\}_{k \in \mathbb{N}}$ сходится к нулю равномерно и для каждого $x \in E$ числовая последовательность $\{b_k(x)\}_{k \in \mathbb{N}}$ монотонна, то ряд $\sum_{k=1}^{\infty} a_k(x)b_k(x)$ сходится равномерно на множестве E .

Доказательство:

Равномерная ограниченность частичных сумм ряда $\sum_{k=1}^{\infty} a_k(x)$ означает, что существует $C > 0$ такое, что для всех $n \in \mathbb{N}$ и всех $x \in E$ выполнено $|A_n(x)| \leq C$, где $A_n(x) = \sum_{k=1}^n a_k(x)$ — n -ая частичная сумма данного ряда.

В силу сделанных предположений по признаку Дирихле сходимости числового ряда функциональный ряд $\sum_{k=1}^{\infty} a_k(x)b_k(x)$ сходится поточечно на E , причём для всех $x \in E$ и $n \in \mathbb{N}$ выполнено

$$|R_n(x)| \leq C |b_{n+1}(x)|$$

где $R_n(x) = \sum_{k=n+1}^{\infty} a_k(x)b_k(x)$ — n -й остаток ряда $\sum_{k=1}^{\infty} a_k(x)b_k(x)$. Так как $b_n \Rightarrow 0$ на E при $n \rightarrow \infty$, то $R_n \Rightarrow 0$ на E , значит $\sum_{k=1}^{\infty} a_k(x)b_k(x)$ сходится равномерно на множестве E .

В качестве примера рассмотрим ряд

$$S(x) = \sum_{k=1}^{\infty} \frac{\sin kx}{k^{\alpha}}$$

где $\alpha > 0$. (Заметим, что при $x \notin \pi\mathbb{Z}$ и $\alpha \leq 0$ данный ряд не сходится даже поточечно.)

Так как

$$\left| \frac{\sin kx}{k^{\alpha}} \right| \leq \frac{1}{k^{\alpha}}$$

и при $\alpha > 1$ ряд $\sum_{k=1}^{\infty} 1/k^{\alpha}$ сходится, то по признаку Вейерштрасса при $\alpha > 1$ ряд $S(x)$ сходится равномерно на любом множестве $E \subset \mathbb{R}$. Пусть теперь $\alpha \in (0, 1]$.

Зафиксируем $\delta > 0$ и рассмотрим $E_{\delta} = [\delta, 2\pi - \delta]$. Так как для всех $x \in E_{\delta}$ и $n \in \mathbb{N}$ выполнено

$$\left| \sum_{k=1}^n \sin kx \right| \leq \frac{1}{\sin \frac{x}{2}} \leq \frac{1}{\sin \frac{\delta}{2}}$$

а последовательность $b_k(x)$ поточечно монотонна и равномерно на E_{δ} сходится к нулю при $k \rightarrow \infty$, то по признаку Дирихле ряд $S(x)$ сходится равномерно на E_{δ} .

Исследуем ряд $S(x)$ на равномерную сходимость на множестве $E = (0, 2\pi)$. Заметим, что если $x_n = \frac{1}{2n}$, то при $n+1 \leq k \leq 2n$

$$\sin \frac{1}{2} \leq \sin(kx_n) \leq \sin 1$$

а также $\frac{1}{k^\alpha} \geq \frac{1}{2^\alpha n^\alpha}$ поэтому

$$\sum_{k=n+1}^{2n} \frac{\sin x_n}{k^\alpha} \geq n \cdot \frac{\sin \frac{1}{2}}{2^\alpha n^\alpha} = C n^{(1-\alpha)} \not\rightarrow 0$$

при $n \rightarrow \infty$ (здесь C — константа, не зависящая от n). Значит по следствию из критерия Коши ряд $S(x)$ не сходится равномерно на E .

Теорема 2. (признак Абеля равномерной сходимости ряда). Если ряд $\sum_{k=1}^{\infty} a_k(x)$ сходится на множестве E , последовательность $\{b_k\}_{k \in \mathbb{N}}$ равномерно ограничена на E и для каждого $x \in E$ числовая последовательность $\{b_k(x)\}_{k \in \mathbb{N}}$ монотонна, то ряд $\sum_{k=1}^{\infty} a_k(x)b_k(x)$ сходится равномерно на множестве E .

Доказательство (через признак Дирихле)

Если сумма $\sum_{k=1}^{\infty} a_k(x)$ ограничена на E , то признак Абеля можно получить как следствие признака Дирихле. Действительно, в этом случае последовательность частичных сумм ряда $\sum_{k=1}^{\infty} a_k(x)$ равномерно ограничена на E . Так как последовательность b_k равномерно ограничена на E то существует константа M такая, что для всех $k \in \mathbb{N}$ и всех $x \in E$ выполнено $|b_k(x)| \leq M$. Так как для каждого $x \in E$ числовая последовательность $b_k(x)$ монотонна, то по теореме Вейерштрасса она имеет предел $b(x)$, причём $|b(x)| \leq M$ по теореме о предельном переходе в неравенстве. Тогда по признаку Дирихле ряд

$$\sum_{k=1}^{\infty} a_k(x) (b_k(x) - b(x))$$

сходится равномерно на множестве E . Ряд

$$\sum_{k=1}^{\infty} a_k(x)b(x)$$

тоже сходится равномерно, так как его остатки

$$\left| b(x) \sum_{k=n+1}^{\infty} a_k(x) \right| \leq M \left| \sum_{k=n+1}^{\infty} a_k(x) \right|$$

сходятся у нулю равномерно на E .

Так как

$$a_k(x)b_k(x) = a_k(x) (b_k(x) - b(x)) + a_k(x)b(x),$$

то ряд $\sum_{k=1}^{\infty} a_k(x)b_k(x)$ сходится равномерно на E .

Доказательство (через преобразование Абеля и сходимост остатков к нулю)

Заметим, что

$$a_k(x) = R_{k-1}(x) - R_k(x)$$

где $R_n(x) = \sum_{k=n+1}^{\infty} a_k(x)$ — n -й остаток ряда $\sum_{k=1}^{\infty} a_k(x)$. Запишем преобразование Абеля:

$$\sum_{k=n+1}^{\infty} (R_{k-1}(x) - R_k(x)) b_k(x) = \sum_{k=n+1}^{\infty} R_k(x) (b_{k+1}(x) - b_k(x)) + R_n(x) b_{n+1}(x)$$

Так как последовательность b_k равномерно ограничена на E то существует константа M такая, что для всех $k \in \mathbb{N}$ и всех $x \in E$ выполнено $|b_k(x)| \leq M$.

Так как для любого $x \in E$ последовательность $\{b_k(x)\}$ монотонна, то

$$\sum_{k=n+1}^{\infty} |b_{k+1}(x) - b_k(x)| = \left| \sum_{k=n+1}^{\infty} (b_{k+1}(x) - b_k(x)) \right| = |b_{n+1}(x)| \leq M$$

Так как $R_k \Rightarrow 0$ на E при $k \rightarrow \infty$, то для любого $\varepsilon > 0$ существует N такое, что для всех $x \in E$ и $k > N$ выполнено

$$|R_k(x)| < \varepsilon.$$

Тогда из (2) для всех $n > N$ получим

$$\left| \sum_{k=n+1}^{\infty} a_k(x) b_k(x) \right| \leq 2M\varepsilon.$$

Значит $\sum_{k=n+1}^{\infty} a_k(x) b_k(x) \Rightarrow 0$ при $x \in E$ и $n \rightarrow \infty$.

Доказательство (через преобразование Абеля и критерий Коши)

Упражнение.

Из признака Абеля следует, что если числовой ряд $\sum_{k=1}^{\infty} a_k$ сходится, то функциональный ряд $\sum_{k=1}^{\infty} a_k x^k$ сходится равномерно на $E = [0,1]$. Например, $\sum_{k=1}^{\infty} \frac{(-1)^k x^k}{k}$ сходится равномерно на $E = [0,1]$.

О частичных пределах последовательности $\sin nx$

Теорема 3. Пусть $x \notin \pi\mathbb{Q}$. Тогда множество всех частичных пределов последовательности $\sin(nx)$ совпадает с отрезком $[-1,1]$.

Доказательство:

Мы докажем, что множество всех частичных пределов последовательности комплексных чисел $z_n = e^{inx}$ совпадает с единичной окружностью $S = \{z \in \mathbb{C} : |z| = 1\}$. Так как отображение $z \mapsto \operatorname{Im} z$ непрерывно, то отсюда следует, что любое число $a \in [-1,1]$ является частичным пределом последовательности $x_n = \operatorname{Im} z_n$. Так как $|x_n| \leq 1$ для всех $n \in \mathbb{N}$, то у последовательности x_n нет

частичных пределов, не принадлежащих $[-1, 1]$. Для любого $\varepsilon \in (0, \pi)$ и любого $z \in S$ под $V_\varepsilon(z)$ будем понимать дугу длины 2ε с центром в точке z :

$$V_\varepsilon(z) := \{ze^{i\varphi} : \varphi \in (-\varepsilon, \varepsilon)\}$$

Зафиксируем $\varepsilon > 0$ и покажем, что существует $\theta \in (-\varepsilon, \varepsilon)$ такое, что

$$e^{ipx} = e^{i\theta}$$

то есть $e^{ipx} \in V_\varepsilon(1)$. Для этого рассмотрим множества

$$V^n := \{e^{inx}\zeta : \zeta \in V_\varepsilon(1)\}$$

такие, что

$$z = e^{i(n+p)x}e^{i\varphi} = e^{inx}e^{i\psi}$$

Значит $e^{ipx} = e^{i\theta}$, где $\theta \in (-2\varepsilon, 2\varepsilon)$. Поэтому $e^{ipx} \in V_{2\varepsilon}(1)$, то есть промежуточное утверждение доказано (ввиду произвольности $\varepsilon > 0$). Пусть $\theta = 0$. Тогда существует $k \in \mathbb{Z}$ такое, что $px = 2\pi k$. Последнее невозможно, так как $x \notin \pi\mathbb{Z}$. Значит $\theta \neq 0$. Не уменьшая общности далее считаем, что $\theta \in (0, \varepsilon)$. Теперь покажем, что для любого $\varphi \in [0, 2\pi]$ и любого $\varepsilon > 0$ в $V_\varepsilon(e^{i\varphi})$ содержится бесконечно много членов последовательности z_n . Для каждого $k \in \mathbb{N}$ будем брать

$$n_k = p \left\lfloor \frac{\varphi + 2\pi k}{\theta} \right\rfloor.$$

Тогда для всех $k \in \mathbb{N}$

$$|e^{in_k x} - e^{i\varphi}| < \varepsilon$$

то есть в $V_\varepsilon(e^{i\varphi})$ содержится бесконечно много членов последовательности e^{inx} .

МА-41. Степенные ряды

В данном разделе мы будем работать не только с вещественными, но и с комплексными числовыми рядами. Несложно проверить, что большинство результатов, доказанных в вещественном случае, сохраняется и в комплексном случае:

- (критерий Коши) ряд $\sum_{k=1}^{\infty} a_k$ (где $a_k \in \mathbb{C}$) сходится тогда и только тогда, когда для любого $\varepsilon > 0$ существует N такое, что для всех $n > N$ и всех $p \in \mathbb{N}$ выполнено $|\sum_{k=n+1}^{n+p} a_k| < \varepsilon$
- (необходимое условие сходимости ряда) если ряд $\sum_{k=1}^{\infty} a_k$ (где $a_k \in \mathbb{C}$) сходится, то $a_k \rightarrow 0$ при $k \rightarrow \infty$
- если ряд $\sum_{k=1}^{\infty} a_k$ (где $a_k \in \mathbb{C}$) сходится абсолютно (то есть $\sum_{k=1}^{\infty} |a_k| < +\infty$), то он сходится;
- (признак Вейерштрасса) если ряд $\sum_{k=1}^{\infty} b_k$ (где $b_k \in \mathbb{R}$) сходится и для всех $k \in \mathbb{N}$ выполнено $|a_k| \leq b_k$ (где $a_k \in \mathbb{C}$), то ряд $\sum_{k=1}^{\infty} a_k$ сходится;
- (признак Дирихле) если частичные суммы ряда $\sum_{k=1}^{\infty} a_k$ (где $a_k \in \mathbb{C}$) ограничены (то есть существует $C > 0$ такое, что для всех n выполнено $|\sum_{k=1}^n a_k| \leq C$), а последовательность $b_k \in \mathbb{R}$ монотонна и стремится к нулю при $k \rightarrow \infty$, то ряд $\sum_{k=1}^{\infty} a_k b_k$ сходится;
- (признак Абеля) если ряд $\sum_{k=1}^{\infty} a_k$ (где $a_k \in \mathbb{C}$) сходится, а последовательность $b_k \in \mathbb{R}$ монотонна и имеет конечный предел при $k \rightarrow \infty$, то ряд $\sum_{k=1}^{\infty} a_k b_k$ сходится;

Аналогичным образом на комплексный случай обобщаются критерии и признаки равномерной сходимости функциональных рядов, члены которых принимают комплексные значения. Члены таких функциональных рядов могут быть функциями не только вещественного, но и комплексного переменного.

Определение 1: Функциональный ряд вида

$$\sum_{n=0}^{\infty} a_n (z - z_0)^n \quad (1)$$

Теорема 2: (первая теорема Абеля). Если ряд (1) сходится в точке $z_1 \in \mathbb{C}$, то он сходится абсолютно в любой точке $z \in \mathbb{C}$, удовлетворяющей неравенству $|z - z_0| < |z_1 - z_0|$. При этом для любого $r \in [0, |z_1|)$ ряд (1) сходится равномерно в круге

$$B_r(z_0) := \{z \in \mathbb{C} : |z - z_0| < r\}$$

Доказательство:

Не уменьшая общности будем считать, что $z_0 = 0$. Так как ряд $\sum_{n=0}^{\infty} a_n z_1^n$ сходится, то $a_n z_1^n \rightarrow 0$ при $n \rightarrow \infty$. Следовательно найдётся $C > 0$ такое, что

$$|a_n z_1^n| \leq C$$

для всех $n \in \mathbb{N}$

Пусть $0 < r < |z_1|$. Обозначим $q = r/|z_1|$, тогда при $|z| \leq r$ выполнено

$$|a_n z^n| = |a_n z_1^n| \left| \frac{z}{z_1} \right|^n \leq |a_n z_1^n| q^n \leq C q^n$$

Значит по признаку Вейерштрасса ряд $\sum_{n=0}^{\infty} |a_n z^n|$ сходится равномерно на круге $\{z \in \mathbb{C} : |z| \leq r\}$. В каждой точке этого круга ряд $\sum_{n=0}^{\infty} a_n z^n$ сходится абсолютно, следовательно, сходится. Ввиду произвольности $r \in (0, |z_1|)$ отсюда следует, что в каждой точке круга $\{z \in \mathbb{C} : |z| < |z_1|\}$ рассматриваемый ряд сходится абсолютно.

Определение 3: Радиусом сходимости ряда (??) называется

$$R := \sup \left\{ |z_1 - z_0| : \sum_{n=0}^{\infty} a_n (z_1 - z_0)^n \text{ сходится, } z_1 \in \mathbb{C} \right\}.$$

Круг $B_R(z_0) := \{z \in \mathbb{C} : |z - z_0| < R\}$ называется кругом сходимости ряда (1).

Легко видеть, что радиус сходимости любого ряда принадлежит $[0, +\infty]$. По первой теореме Абеля любой степенной ряд сходится внутри своего круга сходимости, а вне его замыкания расходится. На границе круга сходимости могут быть как точки, в которых ряд сходится, так и точки, в которых он расходится. Вне замыкания круга сходимости (т.е. при $|z - z_0| > R$) ряд (1) расходится, так как в противном случае мы получили бы противоречие определению супремума. В качестве примера рассмотрим геометрическую прогрессию. А именно, пусть $z_0 = 0$ и $a_n = 1$ для всех n . Тогда при $|z| < 1$ ряд (1) сходится по признаку Вейерштрасса, а при $|z| > 1$ ряд (1) расходится, так как $z^n \not\rightarrow 0$ при $n \rightarrow \infty$. Значит $R = 1$.

Методы вычисления радиуса сходимости Несмотря на то, что в общем случае из сходимости ряда не следует его абсолютная сходимость, при нахождении радиуса (или круга) сходимости степенного ряда можно ограничиться исследованием этого ряда на абсолютную сходимость. А именно, круг сходимости степенного ряда является наибольшим (по вложению) кругом, в каждой точке которого данный ряд сходится абсолютно:

Теорема 4: Радиус сходимости степенного ряда (1) удовлетворяет равенству

$$R := \sup \left\{ |z_1 - z_0| : \sum_{n=0}^{\infty} a_n (z_1 - z_0)^n \text{ сходится абсолютно, } z_1 \in \mathbb{C} \right\} \quad (2)$$

Доказательство:

Пусть R - радиус сходимости степенного ряда, а R_1 - супремум в правой части формулы (2). Так как всякий абсолютно сходящийся ряд сходится, то $R_1 \leq R$. С другой стороны, по первой теореме Абеля в каждой точке круга $B_R(z_0)$ ряд (??) сходится абсолютно, значит $R_1 \geq R$. Следовательно, $R = R_1$.

Теорема 5: Если $a_n \sim b_n$ при $n \rightarrow \infty$, то ряды $\sum_{n=0}^{\infty} a_n z^n$ и $\sum_{n=0}^{\infty} b_n z^n$ имеют один и тот же радиус сходимости.

Доказательство:

Так как $|a_n z^n| \sim |b_n z^n|$ при $n \rightarrow \infty (z \neq 0)$ то ряды $\sum_{n=0}^{\infty} |a_n z^n|$ и $\sum_{n=0}^{\infty} |b_n z^n|$ сходятся или расходятся одновременно. Тогда в силу формулы (2) ряды $\sum_{n=0}^{\infty} a_n z^n$ и $\sum_{n=0}^{\infty} b_n z^n$ имеют один и тот же радиус сходимости

Теорема 6: Пусть ряды $\sum_{n=0}^{\infty} a_n z^n$ и $\sum_{n=0}^{\infty} b_n z^n$ имеют радиусы сходимости R_1 и R_2 соответственно. Тогда ряд $\sum_{n=0}^{\infty} (a_n + b_n) z^n$ имеет радиус сходимости $R \geq \min(R_1, R_2)$

Доказательство:

Так как в каждой точке круга $\{z \in \mathbb{C} : |z| < \min(R_1, R_2)\}$ оба степенных ряда сходятся, то и их сумма сходится. Значит радиус сходимости суммы не меньше

$\min(R_1, R_2)$. Он может оказаться строго больше $\min(R_1, R_2)$, например, если $a_n = 1 = -b_n$ для всех $n \in \mathbb{N}$.

Пусть $R > 0$ и $a_n = R^{-n}$. Сравнивая (1) с геометрической прогрессией, легко увидеть, что радиус сходимости данного ряда равен R . При этом $1/R = a_{n+1}/a_n$. Обобщением этого примера является следующая теорема:

Теорема 7: Если существует $\lim_{n \rightarrow \infty} |a_{n+1}|/|a_n| = q \in [0, +\infty]$, то радиус сходимости ряда (??) равен $R = 1/q$. (Здесь $1/0 = +\infty, 1/+\infty = 0$.)

Доказательство:

Так как

$$\lim_{n \rightarrow \infty} \frac{|a_{n+1} z^{n+1}|}{|a_n z^n|} = q|z|$$

то по признаку Даламбера ряд $\sum_{n=0}^{\infty} |a_n z^n|$ сходится при $q|z| < 1$ и расходится при $q|z| > 1$. Значит его радиус сходимости равен $1/q$.

Пример 8. Пусть $\alpha \in \mathbb{R}$ и $a_n = C_\alpha^n$ для всех $n \in \mathbb{N}$. Тогда ряд (1) имеет радиус сходимости $R = 1$.

Доказательство:

Действительно,

$$\frac{C_\alpha^{n+1}}{C_\alpha^n} = \frac{\alpha - n}{n + 1} \rightarrow -1$$

при $n \rightarrow \infty$, значит по теореме 7 радиус сходимости данного ряда равен 1.

Теорема 9: (формула Коши-Адамара). Радиус сходимости ряда (??) удовлетворяет равенству

$$\frac{1}{R} = \overline{\lim}_{n \rightarrow \infty} \sqrt[n]{|a_n|}$$

Доказательство:

$$\overline{\lim}_{n \rightarrow \infty} \sqrt[n]{|a_n z^n|} = q|z|$$

то по признаку Коши ряд $\sum_{n=0}^{\infty} |a_n z^n|$ сходится при $q|z| < 1$ и расходится при $q|z| > 1$. Значит его радиус сходимости равен $1/q$.

Дифференцирование и интегрирование степенных рядов

Ряд

$$\sum_{n=1}^{\infty} a_n n \cdot (z - z_0)^{n-1} \quad (3)$$

- называется формально продифференцированным рядом (1), а ряд

$$\sum_{n=1}^{\infty} \frac{a_n}{n+1} \cdot (z - z_0)^n \quad (4)$$

- называется формально проинтегрированным рядом (1). Радиус сходимости степенного ряда не меняется при почленном интегрировании или дифференцировании:

Теорема 11: При почленном интегрировании или дифференцировании степенного ряда его радиус сходимости не меняется. Иначе говоря, ряды (1), (3) и (4) имеют один и тот же радиус сходимости.

Доказательство

О поведении степенных рядов на границе круга сходимости

Как мы уже говорили, на границе круга сходимости могут быть как точки, в которых степенной ряд сходится, так и точки, в которых он расходится. Рассмотрим три простых примера:

$$S_1(z) = \sum_{n=0}^{\infty} z^n, \quad S_2(z) = \sum_{n=0}^{\infty} \frac{1}{n} z^n, \quad S_3(z) = \sum_{n=0}^{\infty} \frac{1}{n^2} z^n$$

Несложно проверить, что данные ряды имеют радиус сходимости $R = 1$, но по-разному ведут себя на единичной окружности $\{z \in \mathbb{C} : |z| = 1\}$. А именно,

1. ряд S_1 расходится во всех её точках,
2. ряд S_2 сходится во всех её точках кроме $z = 1$,
3. ряд S_3 сходится во всех её точках.

Сходимость ряда $S_2|z| = 1$ и $z \neq 1$ следует из признака Дирихле.

На границе своего круга сходимости степенной ряд может вести себя достаточно сложным образом. Например, Лузин построил пример степенного ряда, имеющего радиус сходимости $R = 1$, расходящегося во всех точках единичной окружности, и при этом имеющего коэффициенты a_n монотонно стремящиеся к нулю при $n \rightarrow \infty$.

Серпинский построил пример степенного ряда, имеющего радиус сходимости $R = 1$ и сходящегося лишь в одной точке единичной окружности.

О равномерной сходимости степенного ряда вблизи границы круга сходимости

По первой теореме Абеля в любом круге с центром в z_0 , радиус которого строго меньше радиуса сходимости, ряд (??) сходится равномерно. Однако в общем случае сходимость ряда на всём круге сходимости не является равномерной. В связи с этим важную роль играет следующая теорема:

Теорема 12: (вторая теорема Абеля). Если ряд $\sum_{n=0}^{\infty} a_n z^n$ имеет радиус сходимости R и сходится в точке $z = R$, то на отрезке $[0, R]$ данный ряд сходится равномерно.

Доказательство

Не уменьшая общности пусть $R = 1$. По условию ряд $\sum_{n=0}^{\infty} a_n$ сходится. Кроме того, последовательность x^n монотонна для каждого $x \in [0, 1]$. Значит по признаку Абеля ряд $\sum_{n=0}^{\infty} a_n x^n$ сходится равномерно на $[0, 1]$.

Дифференцирование и интегрирование функциональных рядов Следующие две теоремы непосредственно вытекают из соответствующих теорем (о перестановке предельного перехода и интегрирования и о перестановке предельного перехода и дифференцирования) для функциональных последовательностей:

Теорема 1: (об интегрировании функционального ряда). Пусть для всех $u_k \in C[a, b]$ (для всех $k \in \mathbb{N}$) и ряд $S(x) = \sum_{k=1}^{\infty} u_k(x)$ сходится равномерно на $[a, b]$. Тогда его сумма S непрерывна на $[a, b]$. Кроме того, «почленно проинтегрированный» функциональный ряд $\sum_{k=1}^{\infty} \int_a^x u_k(t) dt$ сходится равномерно на $[a, b]$ и для любого $x \in [a, b]$ выполнено

$$\int_a^x \sum_{k=1}^{\infty} u_k(t) dt = \sum_{k=1}^{\infty} \int_a^x u_k(t) dt$$

Теорема 2: (о дифференцировании функционального ряда). Пусть $u_k \in C^1[a, b]$ (для всех $k \in \mathbb{N}$) и «почленно продифференцированный» функциональный ряд $\sum_{k=1}^{\infty} u'_k(x)$ сходится равномерно на $[a, b]$, а ряд $S(x) = \sum_{k=1}^{\infty} u_k(x)$ сходится в некоторой точке $x = x_0 \in [a, b]$. Тогда ряд S сходится равномерно на $[a, b]$, его сумма S непрерывно дифференцируема на $[a, b]$ и для любого $x \in [a, b]$ выполнено

$$\left(\sum_{n=1}^{\infty} u_n \right)'(x) = \sum_{n=1}^{\infty} u'_n(x)$$

Дифференцирование и интегрирование степенных рядов Если степенной ряд имеет вещественные коэффициенты (и $z_0 \in \mathbb{R}$), то пересечение его круга сходимости с вещественной осью называется интервалом сходимости данного ряда. Покажем, что внутри интервала сходимости сумма степенного ряда бесконечно дифференцируема и что её можно почленно интегрировать и дифференцировать:

Теорема 3: (об интегрировании и дифференцировании суммы степенного ряда). Пусть $z_0 = x_0 \in \mathbb{R}$, $\{a_n\}_{n=0}^{\infty} \subset \mathbb{R}$ и ряд имеет радиус сходимости $R > 0$. Тогда его сумма

$$S(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n$$

Теорема 3: (об интегрировании и дифференцировании суммы степенного ряда). Пусть $z_0 = x_0 \in \mathbb{R}$, $\{a_n\}_{n=0}^{\infty} \subset \mathbb{R}$ и ряд имеет радиус сходимости $R > 0$. Тогда его сумма

$$S(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n$$

бесконечно дифференцируема на интервале $(x_0 - R, x_0 + R)$ и для любого $x \in (x_0 - R, x_0 + R)$ выполнено

$$\int_{x_0}^x S(t)dt = \sum_{n=0}^{\infty} a_n \frac{(x-x_0)^{n+1}}{n+1} \quad (5)$$

$$S'(x) = \sum_{n=1}^{\infty} n a_n (x-x_0)^{n-1}$$

Доказательство:

Мы уже знаем, что формально продифференцированный ряд имеет тот же радиус сходимости, что и исходный. Зафиксируем $x \in (x_0 - R, x_0 + R)$. Возьмём $r \in (|x|, R)$. Тогда по первой теореме Абеля на отрезке $[x_0 - r, x_0 + r]$ продифференцированный ряд сходится равномерно. Исходный ряд в любом случае сходится в точке x_0 , значит по теореме 2 имеет место равенство (2). Аналогично, по теореме 1 имеет место равенство (1).

Вещественно-аналитические функции Вещественная функция f называется вещественно-аналитической (регулярной) в точке $x_0 \in \mathbb{R}$, если в некоторой окрестности этой точки она представляется в виде суммы некоторого степенного ряда по степеням $x - x_0$: существуют $r > 0$ и последовательность $\{a_n\}_{n=0}^{\infty} \subset \mathbb{R}$ такие, что для всех $x \in (x_0 - r, x_0 + r)$ выполнено

$$f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n$$

По теореме о дифференцировании степенного ряда всякий вещественный степенной ряд $f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n$ представляет собой вещественно-аналитическую в точке x_0 функцию. Оказывается, что коэффициенты данного ряда можно выразить через производные функции f :

Утверждение 4: Если вещественно-аналитическая в точке x_0 функция f удовлетворяет (3) в некоторой окрестности x_0 , то для всех $n \in \mathbb{N} \cup \{0\}$ имеет место равенство

$$a_n = \frac{f^{(n)}(x_0)}{n!}$$

Доказательство:

Применив n раз формулу (5), получим

$$f^{(n)}(x) = \sum_{k=n}^{\infty} k! C_k^n (x - x_0)^k$$

Подставив $x = x_0$ получим (4).

Определение 5: Если функция f бесконечно дифференцируема в точке $x_0 \in \mathbb{R}$, то ряд

$$S(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n$$

называется рядом Тейлора функции f в точке x_0 . В случае $x_0 = 0$ ряд Тейлора называется рядом Маклорена.

По теореме 3 всякая вещественно-аналитическая в точке x_0 функция f является бесконечно дифференцируемой. Обратное в общем случае не имеет места:

1. Бесконечно дифференцируемая функция может отличаться от суммы своего ряда Тейлора во всех точках, за исключением точки x_0 ;
2. Ряд Тейлора бесконечно дифференцируемой функции может иметь нулевой радиус сходимости.

Приведём соответствующие контпримеры:

Пример 6. Функция

$$f(x) = \begin{cases} 0, & x \leq 0 \\ e^{-1/x}, & x > 0 \end{cases}$$

бесконечно дифференцируема на \mathbb{R} , но её ряд Маклорена тождественно равен нулю.

$$g(x) = \sum_{k=-m}^m a_k x^k,$$

где $m \in \mathbb{N}$ и $a_k \in \mathbb{R}$ для всех $k = -m, -m+1, \dots, m$, будем называть многочленом Лорана. Иначе говоря, многочлен Лорана — это линейная комбинация мономов x^k с целыми степенями.

Так как при $x > 0$ для любого $k \in \mathbb{Z}$ выполнено

$$(x^k e^{-1/x})' = kx^{k-1} e^{-1/x} + x^{k-2} e^{-1/x},$$

то для любого многочлена Лорана $p = p(x)$ найдётся многочлен Лорана $q = q(x)$ такой, что при $x > 0$

$$(p(x)e^{-1/x})' = q(x)e^{-1/x}.$$

Отсюда следует, что для любого $n \in \mathbb{N}$ найдётся многочлен Лорана $p = p(x)$ (зависящий от n) такой, что при $x > 0$

$$f^{(n)}(x) = p(x)e^{-1/x}$$

Так как по правилу Лопиталья для любого $k \in \mathbb{Z}$ выполнено

$$\lim_{x \rightarrow 0} x^k e^{-1/x} = \lim_{t \rightarrow +\infty} t^{-k} e^t = 0,$$

то в силу для любых $n \in \mathbb{N}$ и $m \in \mathbb{N}$ имеем

$$f^{(n)}(x) = o(x^m)$$

при $x \rightarrow 0$.

Индукцией по n покажем, что тогда для всех $n \in \mathbb{N}$

$$f^{(n)}(0) = 0$$

При $n = 0$ данное равенство следует из того, что $e^{-1/x} \rightarrow 0$ при $x \rightarrow 0 + 0$. Если $f^{(n)}(0) = 0$, то по определению производной

$$f^{n+1}(0) = \lim_{x \rightarrow 0} \frac{f^{(n)}(x)}{x} = 0$$

Таким образом, если бы f была вещественно-аналитической в точке $x_0 = 0$, то для всех x из некоторой окрестности нуля имело бы место равенство

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n \equiv 0$$

Это невозможно так как $f(x) > 0$ при $x > 0$.

Пример 7. Функция

$$f(x) = \sum_{n=0}^{\infty} e^{-n} (\sin(n^2 x) + \cos(n^2 x))$$

бесконечно дифференцируема, но её ряд Маклорена имеет нулевой радиус сходимости.

Доказательство:

Так как для любого $m \in \mathbb{N}$ ряд $\sum_{n=0}^{\infty} n^m e^{-n}$ сходится, то по признаку Вейерштрасса ряд f (и его почленные производные любого порядка) сходится равномерно на \mathbb{R} . Значит по теореме 2 функция f бесконечно дифференцируема на \mathbb{R} . При этом для любого $k \in \mathbb{N}$

$$|f^{(k)}(0)| = \sum_{n=0}^{\infty} e^{-n} n^{2k} \geq e^{-n} n^{2n}$$

значит

$$|a_n| = \frac{|f^{(n)}(0)|}{n!} \geq \frac{e^{-n} n^{2n}}{n!} \geq \frac{e^{-n} n^{2n}}{n^n} = (n/e)^n,$$

следовательно по формуле Коши-Адамара $1/R = \overline{\lim}_{n \rightarrow \infty} \sqrt[n]{|a_n|} \geq \lim_{n \rightarrow \infty} (n/e) = +\infty$.

Теорема 8: Если f и g - вещественно-аналитические в точке x_0 функции, то $f \pm g$ также является вещественно-аналитической в точке x_0 . Если, кроме того, $f^{(n)}(x_0) = g^{(n)}(x_0)$ для всех $n \in \mathbb{N} \cup \{0\}$, то f совпадает с g в некоторой окрестности точки x_0 .

Доказательство:

Пусть f представляется в виде суммы степенного ряда на $(x_0 - r_1, x_0 + r_1)$ а g представляется в виде суммы степенного ряда на $(x_0 - r_2, x_0 + r_2)$. Тогда радиус сходимости суммы степенных рядов будет не меньше $r = \min(r_1, r_2)$, а его сумма равна $f(x) + g(x)$ при $x \in (x_0 - r, x_0 + r)$. Аналогично, функция $\varphi(x) = f(x) - g(x)$ является вещественно-аналитической в точке x_0 . Но по условию $\varphi^{(n)}(x_0) = 0$ для всех $n \in \mathbb{N}$, следовательно ряд Тейлора функции φ в точке x_0 имеет нулевые коэффициенты.

МА-43. Ряд Тейлора

Напомним, что если функция f бесконечно дифференцируема в точке $x_0 \in \mathbb{R}$, то ряд

$$T[f, x_0](x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}$$

называется рядом Тейлора функции f в точке x_0 . Мы уже видели, что не всякая бесконечно дифференцируемая функция раскладывается в свой ряд Тейлора. Найдём достаточные условия, при которых такое разложение имеет место.

Чтобы сравнивать значение функции с частичными суммами ряда Тейлора, определим для любого $n \in \mathbb{N}$ многочлен Тейлора

$$T_n[f, x_0](x) := \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!}$$

и остаточный член

$$R_n[f, x_0](x) := f(x) - T_n[f](x).$$

Для краткости вместо $T[f, x_0]$, $T_n[f, x_0]$ и $R_n[f, x_0]$ мы часто будем писать $T[f]$, $T_n[f]$ и $R_n[f]$ или просто T , T_n и R_n соответственно.

Достаточное условие разложимости бесконечно дифференцируемой функции в ряд Тейлора

Лемма 1: (остаточный член в форме Лагранжа). Если f бесконечно дифференцируема в $U_r(x_0)$, то для любых $n \in \mathbb{N}$ и $x \in U_r(x_0)$ существует $\xi \in (x_0, x)$ такое, что

$$R_n[f](x) = f^{(n+1)}(\xi) \frac{(x - x_0)^{n+1}}{(n+1)!}$$

(при $x < x_0$ под (x_0, x) понимается (x, x_0)).

Доказательство:

Если n -ые производные функции f в окрестности точки x_0 не слишком быстро растут при $n \rightarrow \infty$, то f оказывается вещественно-аналитической в точке x_0 :

Теорема 2: (достаточное условие разложения функции в ряд Тейлора). Пусть f бесконечно дифференцируема на $U_r(x_0)$ и существуют константы $M > 0$, $N > 0$ и $q > 0$ такие, что для всех $x \in U_r(x_0)$ и всех натуральных $n > N$ выполнено

$$|f^{(n)}(x)| \leq Mq^n n!$$

Тогда для всех x таких, что $|x - x_0| < \min(r, 1/q)$, выполнено $f(x) = T[f](x)$.

Доказательство

Лемма 3: (о производных суммы геометрической прогрессии). Для любых $t \in (-1,1)$ и $n \in \mathbb{N}$ выполнено

$$\frac{1}{(1-t)^{n+1}} = \sum_{k=n}^{\infty} C_k^n t^{k-n}$$

Доказательство

Теорема 4: Сумма вещественного степенного ряда является вещественноаналитической функцией в каждой точке его интервала сходимости.

Доказательство

Разложения основных элементарных функций в ряд Тейлора

Для разложения в ряд Тейлора конкретных функций вместо теоремы 2 бывает удобнее использовать непосредственно лемму 1:

Теорема 5: Пусть $x_0 \in \mathbb{R}$. Тогда для любого $x \in \mathbb{R}$ имеют разложения

$$\begin{aligned} e^x &= \sum_{n=0}^{\infty} e^{x_0} \frac{(x-x_0)^n}{n!} \\ \sin x &= \sum_{n=0}^{\infty} \sin\left(x_0 + \frac{\pi n}{2}\right) \frac{(x-x_0)^n}{n!} \\ \cos x &= \sum_{n=0}^{\infty} \cos\left(x_0 + \frac{\pi n}{2}\right) \frac{(x-x_0)^n}{n!} \end{aligned}$$

В частности, при $x_0 = 0$ для любого $x \in \mathbb{R}$ выполнено

$$\begin{aligned} e^x &= \sum_{n=0}^{\infty} \frac{x^n}{n!} \\ \sin x &= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \\ \cos x &= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}. \end{aligned}$$

Доказательство:

Зафиксируем произвольное $x > x_0$. (Для $x < x_0$ рассуждения аналогичны.)

Так как для $f(x) = e^x$ при всех $n \in \mathbb{N}$ и $\xi \in (x_0, x)$ выполнено $|f^{(n)}(\xi)| \leq e^{x-x_0}$. Тогда

$$|R_{n-1}(x)| \leq e^{x-x_0} \frac{(x-x_0)^n}{n!} \rightarrow 0$$

Для $f(x) = \sin x$ и $\cos x$ рассуждения аналогичны.

Теорема 6: (ряд Тейлора для логарифма и его производной). Для любого $x \in (-1,1)$ имеют место разложения

$$\frac{1}{1+x} = \sum_{n=0}^{\infty} (-1)^n x^n$$

$$\ln(1+x) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^n}{n}$$

Доказательство:

Первое разложение следует из того, что

$$\sum_{n=0}^m (-x)^n = \frac{1 - (-x)^{m+1}}{1+x} \rightarrow \frac{1}{1+x}$$

при $m \rightarrow \infty$. Второе разложение получается из первого с помощью теоремы об интегрировании и дифференцировании суммы степенного ряда (или теоремы об интегрировании функционального ряда):

$$\begin{aligned} \ln(1+x) &= \int_0^x \frac{1}{1+t} dt = \int_0^x \sum_{k=0}^{\infty} (-1)^k t^k dt \\ &= \sum_{k=0}^{\infty} \int_0^x (-1)^k t^k dt = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} x^{k+1} = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^n}{n}. \end{aligned}$$

Пусть $\alpha \in \mathbb{R} \setminus \{\mathbb{N} \cup 0\}$. Рассмотрим функцию

$$f(x) = (1+x)^\alpha.$$

Для любого $n \in \mathbb{N}$

$$f^{(n)}(x) = n! C_\alpha^n (1+x)^{\alpha-n}$$

Из леммы 1 видно, что для исследования сходимости ряда Маклорена для f необходимо знать поведение коэффициентов C_α^n при $n \rightarrow \infty$. Как мы уже выяснили ранее (с помощью признака Даламбера), радиус сходимости ряда Маклорена для f равен 1. Поэтому для любого $x \in (-1, 1)$ выполнено $C_\alpha^n x^n \rightarrow 0$ при $n \rightarrow \infty$. Нам будет удобно немного усилить данный результат:

Лемма 7: (оценка скорости роста обобщённых биномиальных коэффициентов). Пусть $\alpha \in \mathbb{R}$. Тогда для любого $q \in (0, 1)$ и любого $k \in \mathbb{N}$ выполнено

$$n^k C_\alpha^n q^n \rightarrow 0$$

при $n \rightarrow \infty$. Кроме того, при $\alpha \geq -1$ последовательность $\{C_\alpha^n\}_{n \in \mathbb{N}}$ ограничена.

Доказательство:

Так как при дифференцировании степенного ряда k раз его радиус сходимости не меняется, то ряд

$$\sum_{n=k}^{\infty} C_\alpha^n n! C_n^k x^{n-k}$$

сходится на интервале $(-1,1)$. Так как при любом k выполнено $n!C_n^k \sim n^k$ при $n \rightarrow \infty$, то ряд

$$\sum_{n=k}^{\infty} C_{\alpha}^n n^k x^{n-k}$$

тоже имеет радиус сходимости 1. Значит для любого k выполнено $C_{\alpha}^n n^k x^{n-k} \rightarrow 0$ при $n \rightarrow \infty$

Если $n > |\alpha|$, то

$$\left| \frac{C_{\alpha}^{n+1}}{C_{\alpha}^n} \right| = \left| \frac{\alpha - n}{n+1} \right| = \frac{n - \alpha}{n+1} < 1$$

следовательно при достаточно больших n последовательность $|C_{\alpha}^n|$ убывает.

Напрямую лемма 1 позволяет установить сходимость ряда Маклорена к значению f лишь в точках $x \in (-1/2, 1)$. Действительно, если $x \in (-1, 0)$ то существует $\xi \in (x, 0)$ (зависящее от n) такое, что

$$|R_n(x)| = \left| C_{\alpha}^{n+1} (1 + \xi)^{\alpha} \left(\frac{x}{1 + \xi} \right)^{n+1} \right| \leq C \left(\frac{|x|}{1 - |x|} \right)^{n+1}$$

где константа $C > 0$ не зависит от n . Для того, чтобы правая часть стремилась к нулю при $n \rightarrow \infty$, достаточно чтобы $|x| < 1 - |x|$, то есть $|x| < 1/2$. При $x \in (0, 1)$ рассуждение аналогично. сходимости, нам понадобится новая форма остаточного члена в формуле Тейлора.

Теорема 8: (остаточный член в интегральной форме). Если f бесконечно

дифференцируема в $U_r(x_0)$, то для любых $n \in \mathbb{N}$ и $x \in U_r(x_0)$ выполнено

$$R_n[f](x) = \frac{1}{n!} \int_{x_0}^x f^{(n+1)}(t) (x - t)^n dt$$

Доказательство:

Докажем формулу (2) индукцией по n .

Так как $R_0(x) = f(x) - f(x_0) = \int_{x_0}^x f'(t) dt$, то формула (2) верна при $n = 0$. Предположим, что

$$R_{n-1}(x) = \frac{1}{(n-1)!} \int_{x_0}^x f^{(n)}(t) (x - t)^{n-1} dt$$

Так как $(x - t)^{n-1} = -\frac{1}{n} \frac{d}{dt} (x - t)^n$, то интегрируя по частям получаем

$$\begin{aligned} R_{n-1}(x) &= \frac{1}{n!} \left(-f^{(n)}(t) (x - t)^n \Big|_{t=x_0}^x + \int_{x_0}^x f^{(n+1)}(t) (x - t)^{n+1} dt \right) \\ &= \frac{1}{n!} f^{(n)}(x_0) (x - x_0)^n + \frac{1}{n!} \int_{x_0}^x f^{(n+1)}(t) (x - t)^{n+1} dt \end{aligned}$$

Остаётся заметить, что $R_n(x) = R_{n-1}(x) - \frac{1}{n!} f^{(n)}(x_0) (x - x_0)^n$.

Теорема 9: (обобщённый бином Ньютона). Пусть $\alpha \in \mathbb{R}$. Тогда для любого $x \in (-1, 1)$ выполнено

$$(1+x)^\alpha = \sum_{n=0}^{\infty} C_\alpha^n x^n$$

Доказательство 1

По теореме 8 для любого $x \in \mathbb{R}$ остаточный член в интегральной форме имеет вид

$$R_n(x) = (n+1)C_\alpha^{n+1} \int_0^x (1+t)^\alpha \left(\frac{x-t}{1+t} \right)^n dt$$

Так как t и x имеют один и тот же знак и $|x| - |t| \leq |x| - |tx|$, то

$$\left| \frac{x-t}{1+t} \right| = \frac{|x|-|t|}{|1+t|} \leq \frac{|x|-|t|}{1-|t|} \leq |x|.$$

Значит

$$|R_n(x)| \leq C \cdot (n+1)C_\alpha^{n+1}|x|^n =: r_n$$

где $C = \int_0^1 |1+t|^\alpha dt$ - константа, не зависящая от x и n . Остаётся заметить, что по лемме $7r_n \rightarrow 0$ при $n \rightarrow \infty$, так как $|x| < 1$.

Доказательство 2

Рассмотрим функцию $g(x) = (1+x)^{-\alpha} \cdot S(x)$, где $S(x) = \sum_{n=0}^{\infty} C_\alpha^n x^n$. Необходимо доказать, что $g(x) = 1$ для всех $x \in (-1,1)$. Так как $g(0) = 1$, то достаточно доказать, что $g'(x) = 0$ для всех $x \in (-1,1)$. Заметим, что

$$\begin{aligned} g'(x) \cdot (1+x)^{\alpha+1} &= (-\alpha)S(x) + (1+x)S'(x) \\ &= -(\alpha+1)S(x) + S'(x) + (xS(x))' =: A \end{aligned}$$

По теореме о об интегрировании и дифференцировании суммы степенного ряда

$$\begin{aligned} S'(x) &= \sum_{n=0}^{\infty} (n+1)C_\alpha^{n+1}x^n, \\ (xS(x))' &= \sum_{n=0}^{\infty} (n+1)C_\alpha^n x^n, \end{aligned}$$

Так как $C_\alpha^n + C_\alpha^{n+1} = \frac{\alpha+1}{n+1}C_\alpha^n$, то $A = 0$. Следовательно, $g'(x) = 0$.

Следствие 10. Пусть $\alpha \in \mathbb{R}$. Если $x+y > 0$ и $y > |x|$, то

$$(x+y)^\alpha = \sum_{n=0}^{\infty} C_\alpha^n x^n y^{\alpha-n}.$$

Для функций $f(x) = \arcsin x$ и $g(x) = \operatorname{arctg} x$ выполнено $f'(x) = \frac{1}{\sqrt{1-x^2}}$ и $g'(x) = \frac{1}{1+x^2}$. Поэтому разложения f и g в ряд Маклорена можно получить почленным интегрированием разложений f' и g' аналогично теореме 6.

Полученное выше разложение функции $f(x) = (1+x)^\alpha$ в ряд Маклорена имеет место для всех $x \in (-1,1)$. Для анализа ряда Маклорена на границе интервала сходимости необходимо уточнить поведение коэффициентов C_α^n при $n \rightarrow \infty$.

Лемма 11: Пусть $\alpha \in \mathbb{R}$. Для того, чтобы $C_\alpha^n \rightarrow 0$ при $n \rightarrow \infty$, необходимо и достаточно, чтобы выполнялось $\alpha > -1$. При $\alpha \leq -1$

$$|C_\alpha^n| = \frac{(0-\alpha)(1-\alpha) \cdot \dots \cdot (n-1-\alpha)}{n!} \geq 1$$

поэтому $C_\alpha^n \nrightarrow 0$ при $n \rightarrow \infty$. Пусть $\alpha > -1$. Тогда

$$\begin{aligned} C_\alpha^n &= (-1)^n \frac{(-\alpha)(1-\alpha)(2-\alpha)(3-\alpha) \cdot \dots \cdot (n-1-\alpha)}{n!} \\ &= (-1)^n \frac{(1-1-\alpha)(2-1-\alpha)(3-1-\alpha) \cdot \dots \cdot (n-1-\alpha)}{1 \cdot 2 \cdot 3 \cdot \dots \cdot n} \\ &= (-1)^n \left(1 - \frac{1+\alpha}{1}\right) \left(1 - \frac{1+\alpha}{2}\right) \left(1 - \frac{1+\alpha}{3}\right) \dots \left(1 - \frac{1+\alpha}{n}\right). \end{aligned}$$

Так как $1+\alpha > 0$, то существует n_0 такое, что при $n \geq n_0$ выполнено $\frac{1+\alpha}{n} < 1$. Тогда при $n \geq n_0$

$$\ln |C_\alpha^n| = C + \sum_{k=n_0+1}^n \ln \left(1 - \frac{1+\alpha}{k}\right),$$

где константа $C = \ln \left| \left(1 - \frac{1+\alpha}{1}\right) \left(1 - \frac{1+\alpha}{2}\right) \dots \left(1 - \frac{1+\alpha}{n_0}\right) \right|$ не зависит от n . При каждом $k \geq n_0 + 1$ выполнено

$$\ln \left(1 - \frac{1+\alpha}{k}\right) < -\frac{1+\alpha}{k}$$

значит

$$\ln |C_\alpha^n| < C - \sum_{k=n_0+1}^n \frac{1+\alpha}{k}$$

Правая часть данного неравенства стремится к $-\infty$ при $n \rightarrow \infty$, так как $\sum_{k=1}^{\infty} \frac{1}{k} = +\infty$. Значит $|C_\alpha^n| \rightarrow 0$ при $n \rightarrow \infty$.

Из леммы 11 можно заключить, что при $\alpha > -1$ ряд $\sum_{n=0}^{\infty} C_\alpha^n x^n$ сходится равномерно на $[0,1]$ и $2^\alpha = \sum_{n=0}^{\infty} C_\alpha^n$. Действительно, из определения C_α^n видно, что знаки C_α^n чередуются при возрастании n (начиная с некоторого $n = n_0$, зависящего от α). А по лемме 11 выполнено $|C_\alpha^n| \rightarrow 0$ при $n \rightarrow \infty$. Следовательно, ряд $\sum_{n=0}^{\infty} C_\alpha^n$ сходится по признаку Лейбница. Тогда $S(x) = \sum_{n=0}^{\infty} C_\alpha^n x^n$ сходится равномерно на $[0,1]$ по признаку Абеля. Следовательно S непрерывна на $[0,1]$. Так как при $x \in [0,1)$ выполнено $S(x) = (1+x)^\alpha$, то переходя к пределу при $x \rightarrow 1-0$ получаем $S(1) = 2^\alpha$. Рассуждая аналогично можно представить $\ln 2$ в виде суммы числового ряда.

На самом деле помощью формулы Стирлинга можно показать, что для любого $\alpha \in \mathbb{R} \setminus (\mathbb{N} \cup \{0\})$ существует константа $C \neq 0$ (зависящая от α) такая, что $C_\alpha^n \sim C \frac{(-1)^n}{n^{1+\alpha}}$ при $n \rightarrow \infty$

Иррациональность числа e

Теорема 12: Число e иррационально.

Доказательство:

Пусть $e = p/q$ для некоторых $p \in \mathbb{Z}$ и $q \in \mathbb{N}$ (таких, что дробь p/q несократима). Тогда

$$\frac{p}{q} = e^1 = 1 + \frac{1}{2!} + \frac{1}{3!} + \dots$$

Умножив это равенство на $q!$ мы получим, что число

$$m := q! \sum_{n=q+1}^{\infty} \frac{1}{n!}$$

является целым. С другой стороны

$$\begin{aligned} m &= \frac{1}{q+1} + \frac{1}{(q+1)(q+2)} + \dots \\ &< \frac{1}{q+1} + \frac{1}{(q+1)^2} + \dots \\ &= \frac{1}{q+1} \frac{1}{1 - \frac{1}{q+1}} = \frac{1}{q-1} < 1. \end{aligned}$$

Мы получили противоречие с тем, что m - целое.

Отметим, что известно доказательство иррациональности числа π с помощью интегрирования по частям (имеющее отдалённое сходство с доказательством интегральной формы остаточного члена в формуле Тейлора).

МА-44. Частные производные, производные по направлениям и дифференцируемость функций многих переменных

Пусть $U \subset \mathbb{R}^n$ открыто и задано отображение $f : U \rightarrow \mathbb{R}^m$, где $n, m \in \mathbb{N}$. Значение функции f в точке $x = (x_1, \dots, x_n) \in U$ будем обозначать как $f(x)$, а также как $f(x_1, \dots, x_n)$. Напомним, что для любого $y = (y_1, \dots, y_n) \in \mathbb{R}^n$ под $|y|$ понимается евклидова длина вектора y , то есть $|y| = \sqrt{\sum_{j=1}^n |y_j|^2}$, а под ε -окрестностью точки $x \in \mathbb{R}^n$ понимается $U_\varepsilon(x) = \{y \in \mathbb{R}^n : |y - x| < \varepsilon\}$; также $\overset{\circ}{U}_\varepsilon(x) = U_\varepsilon(x) \setminus \{x\}$.

Например, для любой вещественной матрицы размера $m \times n$ (m строк, n столбцов) $A \in \mathbb{R}^{m \times n}$ можно рассмотреть отображение $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ вида

$$f(x) = A \cdot x$$

Для каждого $i \in \overline{1, m}$ по определению i -ая компонента такого отображения f имеет вид

$$(f(x))_i \equiv f_i(x) = \sum_{j=1}^n a_{ij} x_j$$

где a_{ij} - компоненты матрицы A .

Вернёмся к произвольному отображению f , заданному на $U \subset \mathbb{R}^n$ и принимающем значения в \mathbb{R}^m .

При $m = n = 1$ отображение f называется (числовой) функцией, при $n = 1$ и произвольном m отображение f называется вектор-функцией, а при $m = 1$ и произвольном n отображение f называется (числовой) функцией многих переменных. Впрочем мы не всегда будем придерживаться данной терминологии и будем использовать термин функция вместо термина отображение.

Обобщим символы $\overset{\circ}{O}$ и $\overset{\circ}{O}$ на случай функций многих переменных:

Определение 1: Пусть функции многих переменных f и g заданы в окрестности точки $x \in \mathbb{R}^n$. Если существуют константы $C > 0$ и $\delta > 0$ такие, что для всех $y \in \overset{\circ}{U}_\delta(x)$ выполнено

$$|f(y)| \leq C|g(y)|$$

то говорят, что $f(y) = \overset{\circ}{O}(g(y))$ при $y \rightarrow x$. Аналогично, если для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что для всех $y \in \overset{\circ}{U}_\delta(x)$ выполнено

$$|f(y)| \leq \varepsilon|g(y)|$$

то говорят, что $f(y) = o(g(y))$ при $y \rightarrow x$

Обсудим связь между пределом функции двух переменных и понятием равномерной сходимости. В этом случае вместо $f(x_1, x_2)$ удобно писать $f(x, y)$, где $x, y \in \mathbb{R}$.

Утверждение 2: Для того, чтобы $f(x, y) = o(1)$ при $(x, y) \rightarrow 0$ необходимо и достаточно, чтобы

$$f(r \cos \varphi, r \sin \varphi) \Rightarrow 0$$

при $r \rightarrow 0$ (равномерно по $\varphi \in [0, 2\pi]$)

Доказательство:

Формально мы давали определение равномерной сходимости только для функциональных последовательностей, но более общее определение равномерной сходимости аналогично: $F(x, t) \Rightarrow G(x)$ при $t \rightarrow t_0$ (равномерно по $x \in E$) тогда и только тогда, когда для любого $\varepsilon > 0$ существует (не зависящее от t) $\delta > 0$ такое, что при $|t - t_0| < \delta$ для всех $x \in E$ выполнено $|F(x, t) - G(x)| < \varepsilon$

Теперь запишем определение символа o : $f(x, y) = o(1)$ при $(x, y) \rightarrow 0$ тогда и только тогда, когда для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что при $0 < \sqrt{x^2 + y^2} < \delta$ выполнено $|f(x, y)| < \varepsilon$. Перейдя к полярным координатам $x = r \cos \varphi, y = r \sin \varphi$ получаем: при $0 < r < \delta$ для всех $\varphi \in [0, 2\pi]$ выполнено $|f(r \cos \varphi, r \sin \varphi)| < \varepsilon$. А это и есть определение равномерной сходимости.

В случае отображений символы f и O следует понимать покомпонентно:

Определение 3: Пусть отображения f и g заданы в окрестности точки $x \in \mathbb{R}^n$ и принимают значения в \mathbb{R}^m . Если для каждого $i \in \overline{1, m}$ выполнено $f_i(y) = O(g_i(y))$ при $y \rightarrow x$, то пишут $f(y) = O(g(y))$ при $y \rightarrow x$. Если для каждого $i \in \overline{1, m}$ выполнено $f_i(y) = o(g_i(y))$ при $y \rightarrow x$, то пишут $f(y) = o(g(y))$ при $y \rightarrow x$.

В случае когда f принимает значения в \mathbb{R}^m , а g принимает значения в \mathbb{R} , под g_i следует понимать g .

Для любого отображения f и числовой функции g равенства $f = O(g)$ и $|f| = O(g)$ эквиваленты (то же можно сказать о равенствах $f = o(g)$ и $|f| = o(g)$):

Утверждение 4: Пусть отображение f задано в окрестности точки $x \in \mathbb{R}^n$ и принимает значения в \mathbb{R}^m , а функция многих переменных g задана в окрестности точки x (и принимает значения в \mathbb{R}). Для того, чтобы $f(y) = O(g(y))$ при $y \rightarrow x$ необходимо и достаточно, чтобы $|f(y)| = O(g(y))$ при $y \rightarrow x$. Аналогично для того, чтобы $f(y) = o(g(y))$ при $y \rightarrow x$ необходимо и достаточно, чтобы $|f(y)| = o(g(y))$ при $y \rightarrow x$.

Доказательство:

Ограничимся доказательством для символа O .

Необходимость. Для каждого $i \in \overline{1, m}$ существуют константа C и число $\delta_i > 0$ такое, что $|f_i(y)| \leq C_i |g(y)|$ при $y \in \overset{\circ}{U}_{\delta_i}(x)$. Пусть $\delta = \min(\delta_1, \dots, \delta_m)$ и $C = \sqrt{\sum_{i=1}^m C_i^2}$. Тогда при $|y - x| < \delta$ выполнено $|f(y)| \leq C |g(y)|$.

Достаточность. Пусть существуют константы $C > 0$ и $\delta > 0$ такие, что при $y \in \overset{\circ}{U}_{\delta}(x)$ выполнено $|f(y)| < C |g(y)|$. Тогда для каждого $i \in \overline{1, m}$ выполнено $|f_i(y)| \leq |f(y)| < C |g(y)|$

Лемма 5: Пусть $A \in \mathbb{R}^{m \times n}$ и $f(x) = A \cdot x$ для всех $x \in \mathbb{R}^n$. Тогда $f(x) = O(|x|)$ при $x \rightarrow 0$. Кроме того, $f(x) = o(|x|)$ при $x \rightarrow 0$ тогда и только тогда, когда матрица A — нулевая.

Доказательство:

В силу неравенства Коши-Буняковского

$$|f_i(x)| \leq \left(\sum_{j=1}^n |a_{ij}|^2 \right)^{1/2} |x|$$

значит $f(x) = O(|x|)$ при $x \rightarrow 0$.

Пусть $f(x) = o(|x|)$ при $x \rightarrow 0$. Зафиксируем $i \in \overline{1, m}$ и $j \in \overline{1, n}$. Для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что $|f_i(y)| < \varepsilon$ при $y \in \overset{\circ}{U}_\delta(x)$. В частности, при $y = te_j$, где e_j — j -й базисный вектор и $0 < |t| < \delta$, получаем

$$|a_{ij}t| < \varepsilon|t|$$

Значит $|a_{ij}| < \varepsilon$. Ввиду произвольности $\varepsilon > 0$ отсюда следует, что $a_{ij} = 0$.

Определение 6: Пусть отображение f задано в окрестности точки $x \in \mathbb{R}^n$ и принимает значения в \mathbb{R}^m . Если существует матрица $A \in \mathbb{R}^{m \times n}$ такая, что

$$f(y) = f(x) + A(y - x) + o(|y - x|)$$

при $y \rightarrow x$, то f называется дифференцируемым в точке x . При этом A называется матрицей Якоби отображения f в точке x и обозначается как $Df(x)$, а также как $\frac{D(f_1, \dots, f_m)}{D(x_1, \dots, x_n)}$.

По леммы 5 матрица Якоби определена однозначно.

Аналогично одномерному случаю дифференцируемость влечёт непрерывность:

Утверждение 7: Если отображение f дифференцируемо в точке x , то оно непрерывно в этой точке.

Как мы уже видели, обратное утверждение неверно даже в одномерном случае.

Частные производные и производные по направлению

Пусть функция многих переменных f задана в окрестности точки $x \in \mathbb{R}^n$.

Определение 8: Пусть e_j — j -й базисный вектор. Если у функции $t \mapsto f(x + te_j)$ существует конечная производная в точке $t = 0$, то эта производная называется частной производной функции f по переменной x_j (в точке x):

$$\partial_j f(x) \equiv \partial_{x_j} f(x) \equiv \frac{\partial f(x)}{\partial x_j} := \left. \frac{d}{dt} f(x + te_j) \right|_{t=0}$$

Более общей конструкцией является производная по направлению:

Определение 9: Пусть $\nu \in \mathbb{R}^n, |\nu| = 1$. Если у функции $t \mapsto f(x + t\nu)$ существует конечная производная в точке $t = 0$, то эта производная называется частной производной функции f по направлению ν (в точке x):

$$\partial_\nu f(x) \equiv \frac{\partial f(x)}{\partial \nu} := \left. \frac{d}{dt} f(x + t\nu) \right|_{t=0}$$

$$\frac{\partial f(x)}{\partial x_j} = \frac{\partial f(x)}{\partial e_j}$$

Из определения производной по направлению видно, что $\partial_\nu f(x)$ можно интерпретировать как скорость возрастания функции f в направлении ν .

Выясним связь между матрицей Якоби и частными производными.

Теорема 10: Если f дифференцируема в окрестности точки $x \in \mathbb{R}^n$, то для любого единичного вектора $\nu \in \mathbb{R}^n$ существует

$$\frac{\partial f(x)}{\partial \nu} = Df(x) \cdot \nu$$

где $Df(x)$ - матрица Якоби функции f в точке x .

Доказательство:

Рассмотрим

$$\varphi(z) := f(x+z) - f(x) - Az, \quad z \in \mathbb{R}^n.$$

По определению дифференцируемости $\varphi(z) = o(|z|)$ при $z \rightarrow 0$. Значит $\varphi(t\nu) = o(t)$ при $t \rightarrow 0$ (как функция вещественного переменного t). Таким образом

$$f(x+t\nu) = f(x) + At\nu + o(t)$$

при $t \rightarrow 0$, что и требовалось доказать.

Отметим, что теорема 10, сформулированная для функций многих переменных, непосредственно обобщается и на отображения.

В одномерном случае дифференцируемость функции эквивалентна существованию конечной производной. Однако для функций многих переменных (и отображений) это уже не так. А именно, по теореме 10 дифференцируемость функции влечёт существование частных производных, но обратное в общем случае неверно.

Покажем, что из существования частных производных не следует дифференцируемость. Рассмотрим функцию

$$f(x, y) = \sqrt[3]{x^2 y}$$

(где $x, y \in \mathbb{R}$) имеет нулевые частные производные в нуле (так как $f(0, y) = f(x, 0) = 0$ для всех $x, y \in \mathbb{R}$). Однако эта функция не является дифференцируемой в нуле. Докажем это от противного. Если f дифференцируема в нуле, то по теореме 10 имеем $Df(0,0) = 0$. Тогда $f(x, y) = o(\sqrt{x^2 + y^2})$ при $(x, y) \rightarrow 0$, а следовательно $f(t, t) = o(t)$ при $t \rightarrow 0$. Но $f(t, t) = t$, следовательно $t = o(t)$ при $t \rightarrow 0$, что невозможно. Существование производных по всем направлениям также в общем случае не влечёт дифференцируемость функции. Более того, даже если производные функции по всем направлениям равны нулю, то функция оказаться разрывной. Приведём пример такой функции:

$$f(x, y) = \begin{cases} 1, & y = x^2 > 0 \\ 0, & \text{иначе} \end{cases}$$

Наконец, приведём достаточное условие дифференцируемости:

Теорема 11: Если в окрестности точки $x \in U$ для всех $i \in \overline{1, n}$ существуют $\partial_{x_i} f$, непрерывные в точке x , то f дифференцируема в точке x и $\partial_{x_j} f(x) = \lim_{y \rightarrow x} \partial_{x_j} f(y)$.

Доказательство:

Ограничимся доказательством для функции двух переменных. Пусть в окрестности точки $(0,0)$ у функции f существуют частные производные, непрерывные по совокупности переменных в точке

$(0,0)$. Пусть $\varepsilon > 0$. По определению непрерывности существует $\delta > 0$ такое, что при $\sqrt{x^2 + y^2} < \delta$ выполнено

$$|\partial_x f(x, y) - \partial_x f(0,0)| < \varepsilon/2 \quad \text{и} \quad |\partial_y f(x, y) - \partial_y f(0,0)| < \varepsilon/2.$$

Тогда по теореме Лагранжа для любой точки (x, y) из некоторой окрестности точки $(0,0)$ существуют числа ξ, η такие, что $|\eta| \leq |y|, |\xi| \leq |x|$ и

$$f(x, y) - f(0,0) = f(x, y) - f(x, 0) + f(x, 0) - f(0,0) = \partial_y f(x, \eta) \cdot y + \partial_x f(\xi, 0) \cdot x$$

Отсюда при $\sqrt{x^2 + y^2} < \delta$ в силу (1) получаем

$$|f(x, y) - f(0,0) - \partial_x f(0,0)x - \partial_y f(0,0)y| < \varepsilon(|x| + |y|)/2 < \varepsilon\sqrt{x^2 + y^2}$$

Остается воспользоваться определением дифференцируемости.

Градиент и его геометрический смысл Снова рассмотрим дифференцируемое в точке $x \in \mathbb{R}^n$ отображение, принимающее значения в \mathbb{R}^m . Если $m = 1$, то матрица Якоби $Df(x)$ представляет собой строку длины n . В этом случае вектор (столбец)

$$\nabla f(x) = (Df(x))^T$$

называется градиентом f в точке x .

Теорема 12: (геометрический смысл градиента). Пусть функция $f = f(x_1, \dots, x_n)$ дифференцируема в точке $x \in \mathbb{R}^n$. Тогда функция

$$\nu \mapsto \partial_\nu f(x), \quad \nu \in \mathbb{R}^n, |\nu| = 1$$

принимает своё максимальное значение при $\nu = \nabla f(x)/|\nabla f(x)|$.

Доказательство:

По теореме 10 имеем $\partial_\nu f(x) = (\nabla f(x), \nu)$, где в правой части скобки (\cdot, \cdot) обозначают скалярное произведение. Тогда в силу неравенства Коши-Буняковского

$$|\partial_\nu f(x)| \leq |\nabla f(x)|$$

С другой стороны, при $\nu = \nabla f(x)/|\nabla f(x)|$ получаем

$$\partial_\nu f(x) = |\nabla f(x)|$$

Теперь рассмотрим функцию $z = f(x, y)$. По аналогии с одномерным случаем плоскость

$$z = f(x^0, y^0) + \frac{\partial f(x^0, y^0)}{\partial x} \cdot (x - x_0) + \frac{\partial f(x^0, y^0)}{\partial y} \cdot (y - y_0)$$

называется касательной плоскостью к поверхности $z = f(x, y)$ в точке (x^0, y^0) .

Определение 13: Пусть $C \in \mathbb{R}$. Множество $S_C := \{(x, y) \in \mathbb{R}^2 : f(x, y) = C\}$ называется линией уровня функции f , соответствующей значению C .

Если $\nabla f(x^0, y^0) \neq 0$, то уравнение касательной к кривой S_C в точке $(x^0, y^0) \in S_C$ имеет вид

$$\nabla f(x^0, y^0) \cdot \begin{bmatrix} x - x^0 \\ y - y^0 \end{bmatrix} = 0$$

Таким образом, градиент функции перпендикулярен её линиям уровня.

В общем случае, если поверхность S задана уравнением $F(x, y, z) = 0$, причём $\nabla F \neq 0$ в точке $(x^0, y^0, z^0) \in S$, то плоскость

$$\nabla F(x^0, y^0, z^0) \cdot \begin{bmatrix} x - x^0 \\ y - y^0 \\ z - z^0 \end{bmatrix} = 0$$

называется касательной плоскостью к поверхности S в точке (x^0, y^0, z^0) . В этом случае вектор $\nabla F(x^0, y^0, z^0)$ ортогонален к данной касательной плоскости.

МА-45. Дифференцирование композиции отображений. Теорема о неявной функции.

Теорема 1: (о дифференцируемости композиции отображений; цепное правило). Пусть отображения $f : \mathbb{R}^m \rightarrow \mathbb{R}^k$ и $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ дифференцируемы в точках $y \in \mathbb{R}^m$ и $x \in \mathbb{R}^n$ соответственно, причём $y = g(x)$. Тогда композиция $\Phi = f \circ g$ дифференцируема в точке x и

$$D\Phi(x) = Df(g(x)) \cdot Dg(x)$$

В частности, при $k = 1$ для любого $j \in \overline{1, n}$ выполнено

$$\frac{\partial \Phi}{\partial x_j} = \sum_{i=1}^m \frac{\partial f}{\partial y_i} \frac{\partial g_i}{\partial x_j}$$

где производные $\frac{\partial f}{\partial y_i}$ вычислены в точке $y = g(x)$, а остальные производные вычислены в точке x .

(В теореме 1 достаточно требовать, чтобы f и g были определены не на всём пространстве \mathbb{R}^m или \mathbb{R}^n , а лишь в некоторых окрестностях точек y и x соответственно.)

Доказательство:

По определению дифференцируемости

$$f(g(x) + z) = f(g(x)) + A \cdot z + o(|z|)$$

при $z \rightarrow 0$ в \mathbb{R}^m и

$$g(x + h) = g(x) + B \cdot h + o(|h|)$$

при $h \rightarrow 0$ в \mathbb{R}^n , где $A = Df(y)$ и $B = Dg(x)$. Тогда подставив

$$z = g(x + h) - g(x) = B \cdot h + o(|h|)$$

в (1) получим

$$\begin{aligned} f(g(x + h)) &= f(g(x) + g(x + h) - g(x)) \\ &= f(g(x) + z) \\ &= f(g(x)) + A \cdot z + o(|z|) \\ &= f(g(x)) + A \cdot B \cdot h + o(|h|) \end{aligned}$$

при $h \rightarrow 0$. По определению дифференцируемости (и матрицы Якоби) получаем, что $D\Phi(x) = A \cdot B$, что и требовалось доказать.

Более подробное обоснование

Рассмотрим остаточные члены

$$\begin{aligned} R(z) &= f(g(x) + z) - f(g(x)) - Az \\ r(h) &= g(x + h) - g(x) - Bh \end{aligned}$$

Так как $R(z) = o(|z|)$ при $z \rightarrow 0$, то для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что при $|z| < \delta$ выполнено

$$|R(z)| < \varepsilon|z|$$

Так как $Bh = O(|h|)$ при $h \rightarrow 0$, то $Bh + r(h) = O(|h|)$ при $h \rightarrow 0$. Следовательно существуют $M > 0$ и $\rho > 0$ такие, что при $|h| < \rho$ выполнено

$$|Bh + r(h)| \leq M|h|$$

Тогда при

$$|h| < \min\left(\frac{\delta}{M}, \rho\right)$$

имеем

$$|R(Bh + r(h))| < \varepsilon|Bh + r(h)| \leq M\varepsilon|h|.$$

Таким образом,

$$R(Bh + r(h)) = o(|h|)$$

при $h \rightarrow 0$

Теперь мы можем записать

$$\begin{aligned} f(g(x + h)) &= f(g(x)) + A(Bh + r(h)) + R(Bh + r(h)) \\ &= f(g(x)) + ABh + A \cdot r(h) + o(|h|) \\ &= f(g(x)) + ABh + o(|h|) \end{aligned}$$

при $h \rightarrow 0$, так как $A \cdot r(h) = o(|h|)$ при $h \rightarrow 0$.

Из рассмотренного ранее примера можно увидеть, что теорему 1 нельзя превратить в достаточное условие дифференцируемости композиции отображений. А именно, из дифференцируемости композиции $f \circ g$ для любого дифференцируемого отображения g не следует дифференцируемость f .

Определение 2: Пусть отображение $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ дифференцируемо в точке $x \in \mathbb{R}^n$. Тогда отображение

$$dg(x, h) := \left. \frac{d}{dt} g(x + th) \right|_{t=0} = Dg(x) \cdot h$$

называется дифференциалом отображения g в точке x .

Независимую переменную h , являющуюся аргументом дифференциала, обычно обозначают символом dx . Легко видеть, что при $|h| = 1$ значения дифференциала совпадают со значением производной по направлению h .

Для дифференцируемой в точке $y \in \mathbb{R}^n$ функции $f : \mathbb{R}^m \rightarrow \mathbb{R}$ дифференциал в точке y имеет вид

$$df(y, dy) = \sum_{i=1}^m \frac{\partial f}{\partial y_i} dy_i$$

где частные производные $\frac{\partial f}{\partial y_i}$ вычислены в точке y .

Из теоремы 1 следует инвариантность дифференциала 1 порядка (так же, как и в одномерном случае):

Теорема 3: (инвариантность дифференциала 1 порядка). Пусть выполнены предположения теоремы 1 и $k = 1$. Тогда для композиции отображений f и g выполнено

$$d(f \circ g, dx) = \sum_{i=1}^m \frac{df}{dy_i} dg_i(x, dx)$$

то есть дифференциал композиции получается из дифференциала «внешней» функции f заменой дифференциалов независимых переменных на дифференциалы «внутренних» dg_i

Доказательство:

Действительно,

$$\begin{aligned} d(f \circ g, dx) &= \sum_{j=1}^n \frac{\partial f \circ g}{\partial x_j} dx_j \\ &= \sum_{j=1}^n \sum_{i=1}^m \frac{\partial f}{\partial y_i} \frac{\partial g_i}{\partial x_j} dx_j \\ &= \sum_{i=1}^m \frac{\partial f}{\partial y_i} \sum_{j=1}^n \frac{\partial g_i}{\partial x_j} dx_j \\ &= \sum_{i=1}^m \frac{\partial f}{\partial y_i} dg_i(x, dx) \end{aligned}$$

Теорема о неявной функции

Лемма 4: Функция $\varphi = \varphi(h)$ есть $(|h|)$ при $h \rightarrow 0$ в \mathbb{R}^n тогда и только тогда, когда

$\varphi(h) = \sum_{i=1}^n \alpha_i(h) h_i$, где при каждом $i \in \overline{1, n}$ функция $\alpha_i = \alpha_i(h)$ является бесконечно малой при $h \rightarrow 0$.

Доказательство леммы:

Заметим, что $\varphi(h)$ есть $o(|h|)$ при $h \rightarrow 0$ тогда и только, тогда найдётся бесконечно малая $\alpha = \alpha(h)$ такая, что $\varphi(h) = \alpha(h)|h|$, где $|h| = \sqrt{\sum_{i=1}^n h_i^2}$. Тогда достаточность утверждения теоремы очевидна из того, что $h_i = O(|h|)$ при $h \rightarrow 0$.

$$\varphi(h) = \sum_{i=1}^n \frac{\varphi(h)h_i}{|h|^2} \cdot h_i = \sum_{i=1}^n \alpha_i(h)h_i$$

где функции

$$\alpha_i(h) := \frac{\varphi(h)h_i}{|h|^2}$$

являются бесконечно малыми так как $\varphi(h) = o(|h|)$ и $h_i = O(|h|)$ при $h \rightarrow 0$.

Теорема 5: Пусть $F = F(x_1, \dots, x_n, y)$ определена в окрестности $\Omega \subset \mathbb{R}^{n+1}$ точки $(x^0, y^0) = (x_1^0, \dots, x_n^0, y^0) \in \mathbb{R}^{n+1}$

$$(1) F(x^0, y^0) = 0$$

$$(2) F \text{ непрерывна в } \Omega;$$

$$(3a) \text{ частная производная } F_y \text{ существует в каждой точке множества } \Omega \text{ и непрерывна в } (x^0, y^0)$$

$$(36) F_y(x^0, y^0) \neq 0$$

Тогда существуют $\varepsilon > 0, \delta > 0$ и функция $f : B_\delta(x^0) \rightarrow B_\varepsilon(y^0)$ такие, что в цилиндре $B_\delta(x^0) \times B_\varepsilon(y^0)$ уравнение

$$F(x, y) = 0$$

имеет единственное решение $y = f(x)$, причём функция f непрерывна в x^0 .

Если, кроме того, F дифференцируема в точке (x^0, y^0) , то f дифференцируема в точке x^0 и для любого $i \in \overline{1, n}$

$$f_{x_i}(x^0) = -\frac{F_{x_i}(x^0, y^0)}{F_y(x^0, y^0)}$$

Доказательство:

В силу локальности утверждения теоремы и непрерывности F_y не уменьшая общности можно считать, что $F_y > 0$ в Ω . Тогда найдётся $h > 0$:

$$B_h(x^0) \times B_h(y^0) \subset \Omega$$

и для любого $\varepsilon \in (0, h)$

$$F(x^0, y^0 + \varepsilon) > 0, \quad F(x^0, y^0 - \varepsilon) < 0$$

В силу непрерывности F в точках $(x^0, y^0 \pm \varepsilon)$ существует $\delta > 0$ такое, что

$$F(x, y^0 + \varepsilon) > 0, \quad F(x, y^0 - \varepsilon) < 0.$$

для всех $x \in B_\delta(x^0)$. Для каждого такого x функция $y \mapsto F(x, y)$ непрерывна, строго возрастает на $[y^0 - \varepsilon, y^0 + \varepsilon]$ (так как $F_y > 0$) и принимает в концах этого отрезка значения разных знаков. Значит

по теореме Больцано-Коши о промежуточных значениях существует $y = f(x) \in (y^0 - \varepsilon, y^0 + \varepsilon)$ такое, что $F(x, y) = 0$, причём такое значение y единственно в $(y^0 - \varepsilon, y^0 + \varepsilon)$.

Так как в приведённом выше рассуждении ε можно взять сколь угодно малым, то функция f непрерывна в x^0 .

Для доказательства дифференцируемости f в x^0 не уменьшая общности будем считать что $(x^0, y^0) = (0, 0)$.

В силу леммы и дифференцируемости функции F найдутся бесконечно малые $\alpha_i = \alpha_i(x, y) (i \in \overline{1, n})$ и $\alpha = \alpha(x, y)$ (при $(x, y) \rightarrow 0$) такие, что

$$\sum_{i=1}^n F_{x_i}(0, 0)x_i + F_y(0, 0)y + \sum_{i=1}^n \alpha_i(x, y)x_i + \alpha(x, y)y = 0.$$

Подставим в это равенство $y = f(x)$ и заметим, что функции $\alpha_i(x) = \alpha_i(x, f(x))$ и $\alpha(x) = \alpha(x, f(x))$ будут бесконечно малыми при $x \rightarrow 0$ в силу непрерывности f :

$$f(x) = - \sum_{i=1}^n \frac{F_{x_i}(0, 0) + \alpha_i(x)}{F_y(0, 0) + \alpha(x)} x_i$$

Так как $(F_y(0, 0) + \alpha(x, f(x)))^{-1} = (F_y(0, 0))^{-1} + \rho(x)$ где $\rho(x) \rightarrow 0$ при $x \rightarrow 0$, то раскрыв скобки и воспользовавшись леммой получим

$$f(x) = - \sum_{i=1}^n \frac{F_{x_i}(0, 0)}{F_y(0, 0)} x_i + o(|x|)$$

при $x \rightarrow 0$.

МА-46. Теорема о системе функций, заданных неявно

Под прямоугольной окрестностью точки $x \in \mathbb{R}^N$ будем понимать декартово произведение открытых интервалов вида $(x_i - \delta_i, x_i + \delta_i)$, где $\delta_i > 0$ и индекс i принимает значения от 1 до N . В частности,

$$Q_\delta(x) \equiv Q_\delta(x_1, \dots, x_N) := \left\{ y \in \mathbb{R}^N : \max_{i \in \overline{1, N}} |y_i - x_i| < \delta \right\}$$

будем называть квадратной окрестностью точки x .

Теорема о неявной функции позволяет разрешить нелинейное уравнения вида

$$F(x_1, \dots, x_N) = 0$$

относительно одной из переменных, то есть выразить эту переменную как функцию от остальных (в некоторой окрестности заданного решения $x^0 = (x_1^0, \dots, x_N^0)$). В частности, если F непрерывно дифференцируема в окрестности x^0 , $F(x^0) = 0$ и $\frac{\partial F}{\partial x_N} \neq 0$ в точке x^0 , то существует функция $f = f(x_1, \dots, x_{N-1})$ такая, что в некоторой прямоугольной окрестности точки x^0 уравнение (1) равносильно уравнению $x_N = f(x_1, \dots, x_{N-1})$. При этом функция f оказывается непрерывно дифференцируемой в некоторой окрестности точки $(x_1^0, \dots, x_{N-1}^0)$. Так как на практике такую функцию редко удаётся выразить в явном виде через x_1, \dots, x_{N-1} , то в данном случае говорят, что функция f задана неявно уравнением (1).

Обобщим этот результат на случай системы нелинейных уравнений.

А именно, мы покажем, что если функции $F_i = F_i(x_1, \dots, x_N)$, где $i \in \overline{1, m}$, непрерывно дифференцируемы в окрестности точки x^0 , обращаются в нуль в точке x^0 и матрица Якоби отображения $F = (F_1, \dots, F_m)$ по переменным x_{j_1}, \dots, x_{j_m}

$$D_{x_{i_1}, \dots, x_{i_m}} F := \begin{pmatrix} \frac{\partial F_1}{\partial x_{j_1}} & \frac{\partial F_1}{\partial x_{j_2}} & \dots & \frac{\partial F_1}{\partial x_{j_m}} \\ \frac{\partial F_2}{\partial x_{j_1}} & \frac{\partial F_2}{\partial x_{j_2}} & \dots & \frac{\partial F_2}{\partial x_{j_m}} \\ \dots & \dots & \dots & \dots \\ \frac{\partial F_m}{\partial x_{j_1}} & \frac{\partial F_m}{\partial x_{j_2}} & \dots & \frac{\partial F_m}{\partial x_{j_m}} \end{pmatrix}$$

$$\begin{cases} F_1(x_1, \dots, x_N) = 0 \\ \dots \\ F_m(x_1, \dots, x_N) = 0 \end{cases}$$

можно разрешить относительно переменных x_{j_1}, \dots, x_{j_m} . Для простоты мы ограничимся случаем когда $j_1 = N - m + 1, \dots, j_m = N$.

Определитель матрицы Якоби отображения F по переменным x_{j_1}, \dots, x_{j_m} .

$$\frac{\partial (F_1, \dots, F_m)}{\partial (x_{j_1}, \dots, x_{j_m})} := \det D_{x_{i_1}, \dots, x_{i_m}} F$$

будем называть якобианом F по переменным x_{j_1}, \dots, x_{j_m} . Ясно, что невырожденность невырожденность матрицы Якоби равносильна тому, что соответствующий якобиан отличен от нуля.

Теорема 1: (о системе функций, заданных неявно). Пусть функции $F_i = F_i(x_1, \dots, x_n, y_1, \dots, y_m)$, где $i \in \overline{1, m}$, непрерывно дифференцируемы в окрестности точки $(x^0, y^0) \in \mathbb{R}^n \times \mathbb{R}^m$, где $x^0 = (x_1^0, \dots, x_n^0)$ и $y^0 = (y_1^0, \dots, y_m^0)$. Пусть $F(x^0, y^0) = 0$ и $\frac{\partial(F_1, \dots, F_m)}{\partial(y_1, \dots, y_m)} \neq 0$ в (x^0, y^0) , где $F = (F_1, \dots, F_m)$. Тогда существуют $\varepsilon > 0, \delta > 0$ и непрерывно дифференцируемые функции

$$f_i : Q_\delta(x^0) \rightarrow Q_\varepsilon(y^0), \quad i \in \overline{1, m}$$

такие, что для всех $x \in Q_\delta(x^0)$ и всех $y \in Q_\varepsilon(y^0)$

$$F(x, y) = 0 \Leftrightarrow y = f(x)$$

где $f = (f_1, \dots, f_m)$

Доказательство:

При $m = 1$ данная теорема следует из теоремы о неявной функции, заданной одним уравнением. Пусть доказываемое утверждение верно для системы из $m - 1$ уравнения (с произвольным числом «независимых переменных» $n \in \mathbb{N}$), докажем его для случая m уравнений. Так как $\frac{\partial(F_1, \dots, F_m)}{\partial(y_1, \dots, y_m)} \neq 0$ в (x^0, y^0) , то в разложении данного определителя по последнему столбцу хотя бы одно из алгебраических дополнений его элементов должно быть отлично от нуля. Не уменьшая общности пусть отлично от нуля алгебраическое дополнение элемента $\frac{\partial F_m}{\partial y_m}$ (иначе можно изменить порядок нумерации рассматриваемых уравнений).

Итак, пусть в точке (x^0, y^0)

$$\frac{\partial(F_1, \dots, F_{m-1})}{\partial(y_1, \dots, y_{m-1})} \neq 0$$

Тогда в силу сделанного предположения систему первых $m - 1$ уравнений можно разрешить относительно y_1, \dots, y_{m-1} . А именно, существуют $\delta_1 > 0, \varepsilon_1 > 0$ и непрерывно дифференцируемые функции

$$g_i : Q_{\delta_1}(x_1^0, \dots, x_n^0, y_m^0) \rightarrow Q_{\varepsilon_1}(y_1^0, \dots, y_{m-1}^0), \quad i \in \overline{1, m-1}$$

такие, что для любых $x \in Q_{\delta_1}(x)$, $y_m \in Q_{\delta_1}(y_m^0)$ и $(y_1, \dots, y_{m-1}) \in Q_{\varepsilon_1}(y_1^0, \dots, y_{m-1}^0)$ выполнено

$$\begin{cases} F_1(x_1, \dots, x_n, y_1, \dots, y_m) = 0, \\ \dots, \\ F_{m-1}(x_1, \dots, x_n, y_1, \dots, y_m) = 0. \end{cases} \Leftrightarrow \begin{cases} y_1 = g_1(x_1, \dots, x_n, y_m), \\ \dots, \\ y_{m-1} = g_{m-1}(x_1, \dots, x_n, y_m). \end{cases}$$

Таким образом, мы разрешили первые $m - 1$ уравнений относительно переменных y_1, \dots, y_{m-1} . Остаётся проверить, что после подстановки найденных значений y_1, \dots, y_{m-1} в последнее уравнение мы сможем разрешить его относительно y_m .

Рассмотрим функции

$$\Phi_i(x, y_m) := F_i(x, g_1(x, y_m), \dots, g_{m-1}(x, y_m), y_m), \quad i \in \overline{1, m}$$

где под x понимается (x_1, \dots, x_n) . По теореме о дифференцировании композиции отображений

$$\frac{\partial \Phi_i}{\partial y_m} = \sum_{k=1}^n \frac{\partial F_i}{\partial y_k} \cdot \frac{\partial g_k}{\partial y_m} + \frac{\partial F_i}{\partial y_m}$$

Значит m -й столбец матрицы Якоби отображения $\Phi = (\Phi_1, \dots, \Phi_m)$ (по переменным y) получается прибавлением к m -ому столбцу матрицы Якоби $D_y F$ отображения F (по переменным y) линейной комбинации первых $m-1$ столбцов $D_y F$ с коэффициентами $\frac{\partial g_k}{\partial y_m}$, где $k \in \overline{1, m-1}$. При выполнении такой операции определитель $D_y F$ не изменится, поэтому

$$\frac{\partial (F_1, \dots, F_m)}{\partial (y_1, \dots, y_m)} = \det D_y F = \begin{vmatrix} \frac{\partial F_1}{\partial y_1} & \dots & \frac{\partial F_1}{\partial y_{m-1}} & \frac{\partial F_1}{\partial y_m} \\ \frac{\partial F_2}{\partial y_1} & \dots & \frac{\partial F_2}{\partial y_{m-1}} & \frac{\partial F_2}{\partial y_m} \\ \dots & \dots & \dots & \dots \\ \frac{\partial F_m}{\partial y_1} & \dots & \frac{\partial F_m}{\partial y_{m-1}} & \frac{\partial F_m}{\partial y_m} \end{vmatrix} = \begin{vmatrix} \frac{\partial F_1}{\partial y_1} & \dots & \frac{\partial F_1}{\partial y_{m-1}} & \frac{\partial \Phi_1}{\partial y_m} \\ \frac{\partial F_2}{\partial y_1} & \dots & \frac{\partial F_2}{\partial y_{m-1}} & \frac{\partial \Phi_2}{\partial y_m} \\ \dots & \dots & \dots & \dots \\ \frac{\partial F_m}{\partial y_1} & \dots & \frac{\partial F_m}{\partial y_{m-1}} & \frac{\partial \Phi_m}{\partial y_m} \end{vmatrix}$$

В силу (4) для всех $i \in \overline{1, m-1}$, $x \in Q_{\delta_1}(x)$ и $y_m \in Q_{\delta_1}(y_m^0)$ выполнено

$$\Phi_i(x, y_m) = F_i(x, g_1(x, y_m), \dots, g_{m-1}(x, y_m), y_m) = 0$$

поэтому $\frac{\partial \Phi_1}{\partial y_m} = \dots = \frac{\partial \Phi_{m-1}}{\partial y_m} = 0$. Тогда раскладывая в (5) определитель по последнему столбцу получим

$$\frac{\partial (F_1, \dots, F_m)}{\partial (y_1, \dots, y_m)} = \frac{\partial (F_1, \dots, F_{m-1})}{\partial (y_1, \dots, y_{m-1})} \cdot \frac{\partial \Phi_m}{\partial y_m}$$

Так как в точке (x^0, y^0) выполнено $\frac{\partial (F_1, \dots, F_m)}{\partial (y_1, \dots, y_m)} \neq 0$ и $\frac{\partial (F_1, \dots, F_{m-1})}{\partial (y_1, \dots, y_{m-1})} \neq 0$, то из полученного равенства получаем

$$\frac{\partial \Phi_m}{\partial y_m} \neq 0$$

в точке (x^0, y_m^0) . Тогда по теореме о неявной функции, заданной одним уравнением, существуют $\delta_2 > 0, \varepsilon_2 > 0$ и непрерывно дифференцируемая функция $g_m : Q_{\delta_2}(x^0) \rightarrow Q_{\varepsilon_2}(y^0)$ такие, что для всех $x \in Q_{\delta_2}(x^0)$ и всех $y_m \in Q_{\varepsilon_2}(y_m^0)$

$$\Phi_m(x, y_m) = 0 \Leftrightarrow y_m = g_m(x).$$

Определим $\delta_3 := \min(\delta_1, \delta_2)$ и $\varepsilon_3 := \min(\varepsilon_1, \varepsilon_2, \delta_1)$. Для любых $x \in Q_{\delta_3}(x^0)$ определим

$$\begin{aligned} f_i(x) &:= g_i(x, g_m(x)), \quad i \in \overline{1, m-1} \\ f_m(x) &:= g_m(x). \end{aligned}$$

Отображение $f = (f_1, \dots, f_m)$ непрерывно дифференцируемо как композиция непрерывно дифференцируемых отображений.

Пусть $x \in Q_{\delta_3}(x^0)$ и $y \in Q_{\varepsilon_3}(y^0)$. Тогда последовательно применяя (4) и (6) получаем (3).

Так как отображение f непрерывно и $f(x^0) = y^0$, то для $\varepsilon := \varepsilon_3$ существует $\delta \in (0, \delta_3)$ такое, что для всех $x \in Q_\delta(x^0)$ выполнено $f(x) \in Q_\varepsilon(y^0)$, то есть f отображает $Q_\delta(x^0)$ в $Q_\varepsilon(y^0)$. При этом для всех $x \in Q_\delta(x^0) \subset Q_{\delta_3}(x^0)$ и $y \in Q_\varepsilon(y^0)$ попрежнему выполнено (3). Теорема доказана.

Заметим, что производные функций f_i можно выразить через производные функций F_i из равенств $F(x, f(x)) = 0$. Действительно, по теореме о дифференцировании композиции отображений получаем

$$\frac{\partial F_i}{\partial x_j} + \sum_{k=1}^m \frac{\partial F_i}{\partial y_k} \cdot \frac{\partial f_k}{\partial x_j} = 0$$

или в матричной записи

$$D_x F + D_y F \cdot D_x f = 0$$

(все значения вычислены в точке $(x, y) = (x, f(x))$). Отсюда получаем

$$D_x f = -(D_y F)^{-1} \cdot D_x F$$

что является естественным обобщением случая неявной функции, заданной одним уравнением.

МА-47. Смешанные частные производные и дифференциалы высших порядков

Частные производные 2 порядка

Частные производные высших порядков определяются естественным образом как частные производные от частных производных более низкого порядка. Например, если $n = 2$ и $m = 1$, то есть $f = f(x, y)$ - вещественная функция двух переменных, то

$$\frac{\partial^2 f(x, y)}{\partial x^2} \equiv f_{xx}(x, y) := \frac{\partial}{\partial x} \left(\frac{\partial f(x, y)}{\partial x} \right)$$

называется чистой частной производной второго порядка по переменной x , а

$$\frac{\partial^2 f(x, y)}{\partial x \partial y} \equiv f_{yx}(x, y) := \frac{\partial}{\partial x} \left(\frac{\partial f(x, y)}{\partial y} \right)$$

называется смешанной частной производной второго порядка по переменным x и y .

Аналогично определяются смешанные частные производные более высокого порядка, например $\frac{\partial^3 f}{\partial x \partial y^2} \equiv f_{yyx}$

Предположим что $f = f(x, y)$ определена в окрестности нуля. Обозначим (для произвольной функции f)

$$\Delta_x f(x, y) := f(x, y) - f(0, y), \quad \Delta_y f(x, y) := f(x, y) - f(x, 0)$$

Тогда

$$\Delta_x \Delta_y f(x, y) = \Delta_y \Delta_x f(x, y)$$

Действительно,

$$\begin{aligned} \Delta_x \Delta_y f(x, y) &= \Delta_x (f(x, y) - f(x, 0)) \\ &= (f(x, y) - f(x, 0)) - (f(0, y) - f(0, 0)) \\ &= (f(x, y) - f(0, y)) - (f(x, 0) - f(0, 0)) \\ &= \Delta_y (f(x, y) - f(0, y)) \end{aligned}$$

В силу (1) естественно ожидать, что $f_{xy} = f_{yx}$. Однако в общем случае смешанные частные производные, взятые в разном порядке, могут отличаться. В этом можно убедиться рассмотрев $f(x, y) = x \cdot y \cdot g(x, y)$ с подходящей функцией $g(x, y)$.

Доказательство:

Возьмём

$$g(x, y) = \frac{x^2 - y^2}{x^2 + y^2}$$

при $x^2 + y^2 > 0$ и $g(0,0) = 0$. Тогда при любом x выполнено $f(x, 0) = 0$, следовательно $f_x(0,0) = 0$. Также при $y \neq 0$

$$f_x(0, y) = \lim_{x \rightarrow 0} \frac{f(x, y) - f(0, y)}{x} = yg(0, y)$$

Поэтому

$$f_{xy}(0,0) = \lim_{y \rightarrow 0} \frac{f_x(0, y) - f_x(0,0)}{y} = \lim_{y \rightarrow 0} g(0, y) = -1$$

Аналогично

$$f_{yx}(0,0) = \lim_{x \rightarrow 0} g(x, 0) = 1$$

Таким образом $f_{xy}(0,0) \neq f_{yx}(0,0)$.

Несмотря на то, что в общем случае $f_{xy} \neq f_{yx}$, при достаточно общих предположениях оказывается, что $f_{xy} = f_{yx}$:

Теорема 1: Пусть для некоторых $i, j \in \overline{1, n}$ смешанные частные производные $\frac{\partial^2 f(x)}{\partial x_i \partial x_j}$ и $\frac{\partial^2 f(x)}{\partial x_j \partial x_i}$ существуют в окрестности точки $x^0 \in \mathbb{R}^n$ и непрерывны в точке x^0 . Тогда

$$\frac{\partial^2 f(x^0)}{\partial x_i \partial x_j} = \frac{\partial^2 f(x^0)}{\partial x_j \partial x_i}$$

Доказательство:

Для простоты мы ограничимся случаем $n = 2$ и $x^0 = (0,0)$. Итак, пусть $f = f(x, y)$ имеет частные производные f_{xy} и f_{yx} в окрестности точки $(0,0)$, причём эти производные непрерывны в $(0,0)$.

$$\varphi(x) := f(x, y) - f(x, 0)$$

По теореме Лагранжа существует ξ такое, что $0 < |\xi| < |x|$ и

$$\varphi(x) - \varphi(0) = \varphi'(\xi) \cdot x$$

При этом также по теореме Лагранжа существует η такое, что $0 < |\eta| < |y|$ и

$$\varphi'(\xi) = f_x(\xi, y) - f_x(\xi, 0) = f_{xy}(\xi, \eta) \cdot y$$

Таким образом

$$\varphi(x) - \varphi(0) = f_{xy}(\xi, \eta) \cdot y \cdot x$$

Аналогично, зафиксировав $x \neq 0$ и рассмотрев функцию

$$\psi(y) := f(x, y) - f(y, 0)$$

и дважды применив теорему Лагранжа мы получим точки $\tilde{\xi}, \tilde{\eta}$ такие, что $0 < |\tilde{\xi}| < |x|$, $0 < |\tilde{\eta}| < |y|$ и

$$\psi(y) - \psi(0) = f_{yx}(\tilde{\xi}, \tilde{\eta}) \cdot x \cdot y$$

Подставив $x = y = t$ мы получим $f_{xy}(\xi) = f_{xy}(0,0) + o(1)$ при $t \rightarrow 0$ и аналогично $f_{yx}(\eta) = f_{yx}(0,0) + o(1)$ при $t \rightarrow 0$, так как f_{xy} и f_{yx} непрерывны в нуле.

В силу (1)

$$\varphi(x) - \varphi(0) = \psi(y) - \psi(0)$$

поэтому подставив в это равенство $x = y = t$ получим

$$f_{xy}(0,0)t^2 + o(t^2) = f_{yx}(0,0)t^2 + o(t^2).$$

Разделив на t^2 и перейдя к пределу $t \rightarrow 0$ получим $f_{xy}(0,0) = f_{yx}(0,0)$.

Имеется и другое достаточное условие равенства смешанных частных производных:

Теорема 2: Пусть для некоторых $i, j \in \overline{1, n}$ частные производные $\frac{\partial f(x)}{\partial x_i}$ и $\frac{\partial f(x)}{\partial x_j}$ существуют в окрестности точки $x^0 \in \mathbb{R}^n$ и дифференцируемы в точке x^0 . Тогда имеет место равенство (2).

Доказательство:

Аналогично доказательству предыдущей теоремы рассмотрим $x \neq 0, y \neq 0$ и функции

$$\varphi(x) := f(x, y) - f(x, 0), \quad \psi(y) := f(x, y) - f(0, y).$$

По теореме Лагранжа о среднем существуют ξ и η такие, что $0 < |\xi| < |x|$, $0 < |\eta| < |y|$ и

$$\varphi(x) - \varphi(0) = (f_x(\xi, y) - f_x(\xi, 0))x, \quad \psi(y) - \psi(0) = (f_y(x, \eta) - f_y(0, \eta))y.$$

В силу дифференцируемости f_x в нуле

$$\begin{aligned} f_x(\xi, y) &= f_x(0,0) + f_{xx}(0,0) \cdot \xi + f_{xy}(0,0) \cdot y + o(\sqrt{\xi^2 + y^2}) \\ f_x(\xi, 0) &= f_x(0,0) + f_{xx}(0,0) \cdot \xi + o(\sqrt{\xi^2 + y^2}) \end{aligned}$$

при $(\xi, y) \rightarrow 0$. Следовательно,

$$\varphi(x) - \varphi(0) = f_{xy}(0,0) \cdot yx + o(x^2 + y^2)$$

при $(x, y) \rightarrow (0,0)$. Аналогично

$$\psi(y) - \psi(0) = f_{yx}(0,0) \cdot xy + o(x^2 + y^2)$$

Так как $\varphi(x) - \varphi(0) = \psi(y) - \psi(0)$, то подставив $x = y = t$ аналогично доказательству предыдущей теоремы получим

$$f_{xy}(0,0) \cdot t^2 + o(t^2) = f_{yx}(0,0) \cdot t^2 + o(t^2)$$

при $t \rightarrow 0$. Отсюда $f_{xy}(0,0) = f_{yx}(0,0)$.

Дифференциалы высших порядков

Дифференциалы высших порядков проще всего определить по индукции. Мы ограничимся рассмотрением функций многих переменных, принимающих вещественные значения.

Определение 3: Пусть $m \geq 2$ - натуральное число и множество $U \subset \mathbb{R}^n$ открыто. Функция $f : U \rightarrow \mathbb{R}$ называется m раз дифференцируемой в точке $x \in U$, если она дифференцируема в окрестности x и все частные производные $\frac{\partial f}{\partial x_i}$ дифференцируемы $(m-1)$ раз в точке x .

Заметим, что если f дифференцируема m раз в точке $x \in \mathbb{R}^n$, то все её частные производные порядка $k \leq m-1$ дифференцируемы $m-k$ раз в точке x . В этом несложно убедиться индукцией по m , так как частные производные функции f порядка (которые по определению дифференцируемы $m-1$ раз в точке x).

Утверждение 4: Если f дифференцируема m раз в точке $x \in \mathbb{R}^n$, то для каждого $h \in \mathbb{R}^n$ функция $\varphi(t) = f(x+th)$ дифференцируема m раз в точке $t=0$

Доказательство:

Докажем утверждение индукцией по m . При $m=1$ данное утверждение верно, так как из дифференцируемости следует существование производных по любым направлениям.

При $m > 1$ заметим, что для любого $h \in \mathbb{R}^n$ при достаточно малых t по теореме о дифференцировании композиции отображений

$$\varphi'(t) = \sum_{i=1}^n \frac{\partial f(x+th)}{\partial x_i} h_i$$

Если правая часть дифференцируема $m-1$ раз в точке $t=0$, то левая часть тоже. А значит φ дифференцируема m раз в нуле.

Теперь мы можем дать следующее определение:

Определение 5: Если f дифференцируема m раз в точке $x \in \mathbb{R}^n$, то

$$d^m f(x, h) := \left. \frac{d^m}{dt^m} f(x+th) \right|_{t=0}$$

называется дифференциалом порядка m функции f в точке x . (В частности, $d^0 f(x, h) := f(x)$.)

Независимую переменную h часто обозначают символом $dx = (dx_1, \dots, dx_n)$.

На практике для вычисления дифференциала порядка m удобно использовать следующее эквивалентное определение дифференциала:

Утверждение 6: Если f дифференцируема m раз в точке $x \in \mathbb{R}^n$, то для любого $h \in \mathbb{R}^n$

$$d^m f(x, h) = dg(x, h)$$

где $g(x) = d^{m-1} f(x, h)$

Доказательство:

При достаточно малых $h \in \mathbb{R}^n$ для всех $t \in (-1, 1)$ выполнено

$$g(x + th) = d^{m-1}f(x + th, h) = \frac{d^{m-1}}{dt^{m-1}}f(x + th)$$

так как $\frac{d^{m-1}}{d\tau^{m-1}}f(x + th + \tau h)\Big|_{\tau=0} = \frac{d^{m-1}}{dt^{m-1}}f(x + th)$. Значит

$$d^m f(x, h) = \frac{d}{dt} \left(\frac{d^{m-1}}{dt^{m-1}}f(x + th) \right) \Big|_{t=0} = \frac{d}{dt}g(x + th) \Big|_{t=0}$$

что и требовалось доказать.

Дифференциал порядка m можно выразить через частные производные порядка m :

Утверждение 7: Если f дифференцируема m раз в точке $x \in U$, то для любого $h \in \mathbb{R}^n$

$$d^m f(x, h) = \sum_{i_1=1}^n \dots \sum_{i_m=1}^n \frac{\partial^m f(x)}{\partial x_{i_1} \dots \partial x_{i_m}} h_{i_1} \dots h_{i_m}.$$

Доказательство:

Для дифференциалов первого порядка данное утверждение уже известно. Пусть оно верно для дифференциала порядка $m - 1$. Тогда для любого $h \in \mathbb{R}^n$ при достаточно малых t

$$\frac{d^{m-1}}{dt^{m-1}}f(x + th) = \sum_{i_1=1}^n \dots \sum_{i_{m-1}=1}^n \frac{\partial^{m-1}f(x + th)}{\partial x_{i_1} \dots \partial x_{i_{m-1}}} h_{i_1} \dots h_{i_{m-1}}$$

Дифференцируя это равенство по t при $t = 0$ получим (3).

Например, если $f = f(x, y)$ трижды дифференцируема в точке (x, y) , то в силу равенства смешанных производных, записанных в разном порядке,

$$\begin{aligned} d^3 f &= f_{xxx}dx^3 + f_{xxy}dx^2dy \\ &\quad + f_{xyx}dx^2dy + f_{xyy}dxdy^2 \\ &\quad + f_{yxx}dx^2dy + f_{yxy}dxdy^2 \\ &\quad + f_{yyx}dxdy^2 + f_{yyy}dy^3 \\ &= f_{xxx}dx^3 + 3f_{xxy}dx^2dy + 3f_{xyy}dxdy^2 + f_{yyy}dy^3. \end{aligned}$$

Аналогично в общем случае, используя теорему 2, можно записать формулу (3) в виде

$$d^m f(x, h) = \sum_{k_1 + \dots + k_n = m} \frac{m!}{k_1! \dots k_n!} \frac{\partial^m f(x)}{\partial x_1^{k_1} \dots \partial x_n^{k_n}} h_1^{k_1} \dots h_n^{k_n},$$

k_1, \dots, k_n , сумма которых равна m .

Отсюда видно, что если $p(x) = x_1^{k_1} \dots x_n^{k_n}$ - моном степени $k = k_1 + \dots + k_n$, где $k_1, \dots, k_n \in \mathbb{N} \cup \{0\}$, то для любого $m \in \mathbb{N} \cup \{0\}$

$$d^m p(0, h) = \begin{cases} k! p(h), & m = k \\ 0, & \text{иначе.} \end{cases}$$

Формула Тейлора для функций многих переменных

Теорема 8: (формула Тейлора с остаточным членом в форме Лагранжа). Если f дифференцируема $m \in \mathbb{N}$ раз в некотором шаре $U_r(x) \subset \mathbb{R}^n$, то для любого $h \in U_r(0)$ найдётся $\theta \in (0,1)$ такое, что

$$f(x+h) = \sum_{k=0}^{m-1} \frac{1}{k!} d^k f(x, h) + \frac{1}{m!} d^m f(x + \theta h, h).$$

Доказательство:

По формуле Тейлора с остаточным членом в форме Лагранжа для функции $\varphi(t) = f(x + th)$ существует $\theta \in (0,1)$ такое, что

$$\varphi(1) = \sum_{k=0}^{m-1} \frac{1}{k!} \varphi^{(k)}(0) + \frac{1}{m!} \varphi^{(m)}(\theta)$$

По определению дифференциала порядка k имеем $\varphi^{(k)}(0) = d^k f(x, h)$ и $\varphi^{(m)}(\theta) = d^m f(x + \theta h, h)$

Определение 9: Многочлен

$$T_m[f](h) := \sum_{k=0}^m \frac{1}{k!} d^k f(x, h)$$

называется многочленом Тейлора порядка m функции f в точке x .

Наиболее общая теорема о разложении функции по формуле Тейлора с остаточным членом в форме Пеано является непосредственным обобщением определения дифференцируемости: **Теорема 10:** (формула Тейлора с остаточным членом в форме Пеано). Пусть $x \in \mathbb{R}^n$ и $U \subset \mathbb{R}^n$ открыто и $f : U(x) \rightarrow \mathbb{R}$ дифференцируема $m \in \mathbb{N}$ раз в точке x . Тогда

$$f(x+h) = \sum_{k=0}^m \frac{1}{k!} d^k f(x, h) + o(|h|^m)$$

при $h \rightarrow 0$ в \mathbb{R}^n

Доказательство:

По определению дифференцируемости существует $r > 0$ такое, что f дифференцируема $m-1$ раз в $U_r(x)$ для некоторого $r > 0$.

Рассмотрим функцию $g(h) = f(x+h) - \sum_{k=0}^m \frac{1}{k!} d^k f(x, h)$. В силу (5) дифференциалы функции g всех порядков от 0 до m включительно в точке $h = 0$ тождественно равны нулю. Значит все частные производные функции g порядка не выше m в точке $h = 0$ также равны нулю.

Применив теорему 8 к функции g в нуле для каждого $h \in U_r(0)$ получим $\theta \in (0,1)$ такое, что

$$g(h) = \frac{1}{(m-1)!} d^{m-1} g(\theta h, h)$$

По определению дифференцируемости для любых $i_1, \dots, i_{m-1} \in \overline{1, n}$

$$\frac{\partial^{m-1} g(h)}{\partial h_{i_1} \dots \partial h_{i_{m-1}}} = o(|h|)$$

при $h \rightarrow 0$. Тогда из (6) и (7) (используя утверждение 7) получаем, что

$$g(h) = \psi(\theta h) \cdot O(|h|^{m-1})$$

где $\psi(h) = o(|h|)$ при $h \rightarrow 0$.

Остаётся проверить, что $\psi(|\theta h|) = o(|h|)$ при $h \rightarrow 0$. Так как $\psi(h) = o(|h|)$ при $h \rightarrow 0$, то для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что если $|h| < \delta$ то $|\psi(h)| < \varepsilon|h|$. Тогда и для любых $h \in B_\delta(0)$ и $\theta \in (0,1)$ (возможно, зависящего от h) выполнено $|\theta h| < \delta$, следовательно $|\psi(h)| < \varepsilon|h|$.

(Другое доказательство (индукцией по m) можно найти, например, в методичке М.В. Балашова.) Для разложения функции f по формуле Тейлора на практике бывает удобнее не считать её частные производные, а использовать известные разложения функций одного переменного, входящих в определение функции f . Правомерность этого метода обосновывается единственностью разложения по формуле Тейлора, которая вытекает из следующего простого утверждения:

Теорема 11: Пусть $P = P(x)$ - многочлен степени не выше $k \in \mathbb{N}$, $x \in \mathbb{R}^n$. Если $P(x) = o(|x|^k)$ при $x \rightarrow 0$, то $P = 0$

Доказательство:

Пусть

$$P(x) = \sum_{l=0}^k \sum_{k_1+\dots+k_n=l} a_{k_1\dots k_n} \cdot x_1^{k_1} \cdot \dots \cdot x_n^{k_n}$$

Зафиксируем произвольное $y \in \mathbb{R}^n$ и подставим $x = t \cdot y$, где $t \in \mathbb{R}$. Так как многочлен $p(t) = P(t \cdot y)$ имеет степень k и удовлетворяет равенству $p(t) = o(t^k)$ при $t \rightarrow 0$, то все его коэффициенты равны нулю. Следовательно, для каждого l от 0 до k имеем

$$\sum_{k_1+\dots+k_n=l} a_{k_1\dots k_n} \cdot y_1^{k_1} \cdot \dots \cdot y_n^{k_n} = 0$$

Так как это равенство выполнено для всех $y \in \mathbb{R}^n$, то дифференцируя его j_1 раз по y_1, \dots, j_n раз по y_n и затем подставляя $y_1 = \dots = y_n = 0$ получаем

$$a_{j_1\dots j_n} = 0$$

Так как это верно для любых j_1, \dots, j_n удовлетворяющих $j_1 + \dots + j_n = l$, то $P = 0$.

МА-48. (Безусловные) экстремумы функций многих переменных

Локальные экстремумы функций многих переменных Пусть $\Omega \subset \mathbb{R}^n$ открыто и задана функция $f : \Omega \rightarrow \mathbb{R}$.

Определение 1: Точка $x \in \Omega$ называется стационарной точкой функции f если f дифференцируема в точке x и в этой точке для каждого $i \in \overline{1, n}$ выполнено $\frac{\partial f}{\partial x_i} = 0$.

Так как $df(x, h) = \sum_{i=1}^n \frac{\partial f(x)}{\partial x_i} h_i$, то $x \in \Omega$ является стационарной точкой тогда и только тогда, когда дифференциал f в точке x тождественно равен нулю (то есть $df(x, h) = 0$ для всех $h \in \mathbb{R}^n$.)

Теорема 2: (необходимые условия локального экстремума). Пусть функция f имеет локальный экстремум в точке $z \in \Omega$ и дифференцируема в этой точке. Тогда z является стационарной точкой функции f .

Если f имеет локальный минимум и дважды дифференцируема в точке z , то её второй дифференциал определён неотрицательно в точке z . В случае локального максимума второй дифференциал определён неположительно.

Доказательство:

Если e_i — i -ый базисный вектор, то функция $t \mapsto f(z + te_i)$ имеет локальный экстремум в точке $t = 0$. По теореме Ферма её производная в этой точке равна нулю, следовательно $\frac{\partial f(z)}{\partial x_i} = 0$ для всех $i \in \overline{1, n}$.

Допустим, что z — локальный минимум f и нашёлся вектор $h \in \mathbb{R}^n$, для которого $d^2 f(z, h) < 0$. Тогда по формуле Тейлора с остаточным членом в форме Пеано

$$f(z + th) = f(z) + df(z, th) + \frac{1}{2} d^2 f(z, th) + o(|th|^2) = \frac{t^2}{2} d^2 f(z, h) + o(t^2)$$

при $t \rightarrow 0$. Следовательно, $f(z + th) < f(z)$ при достаточно малых t .

Лемма 3: Если квадратичная форма $k = k(h)$, $h \in \mathbb{R}^n$, положительно определена, то существует $\varkappa > 0$ такое, что $k(h) \geq \varkappa |h|^2$ для всех $h \in \mathbb{R}^n$. Пусть $\varkappa := \inf_{h \in \mathbb{S}^{n-1}} k(h)$, где $\mathbb{S}^{n-1} = \{h \in \mathbb{R}^n : |h| = 1\}$ — единичная сфера. Так как функция $k = k(h)$ непрерывна, а \mathbb{S}^{n-1} является компактом, то по теореме Больцано-Вейерштрасса функция k принимает значение \varkappa в некоторой точке $h_* \in \mathbb{S}^{n-1}$ и, следовательно, $\varkappa > 0$. Тогда для всех $h \in \mathbb{R}^n$ выполнено $k(h) = |h|^2 k\left(\frac{h}{|h|}\right) \geq \varkappa |h|^2$. Доказательство 2

Перейдём в ОНБ из собственных векторов матрицы квадратичной формы k . В этом базисе $k(w) = \lambda_1 w_1^2 + \dots + \lambda_n w_n^2$. Так как k определена положительно, то для каждого $i \in \overline{1, n}$ выполнено $\lambda_i > 0$. Следовательно взяв $\varkappa := \min(\lambda_1, \dots, \lambda_n)$ получим $k(w) \geq \varkappa (w_1^2 + \dots + w_n^2) = \varkappa |w|^2$. Так как $w = Sh$ для некоторой ортогональной матрицы S (не зависящей от h и w), то получаем $k(h) \geq \varkappa |h|^2$.

Теорема 4: Пусть функция f дважды дифференцируема в точке $z \in \Omega$ и первый дифференциал f в этой точке равен нулю.

1. Если $d^2 f(z, h) > 0$ для всех $h \in \mathbb{R}^n \setminus \{0\}$, то z является точкой строгого локального минимума функции f .

2. Если $d^2f(z, h) < 0$ для всех $h \in \mathbb{R}^n \setminus \{0\}$, то z является точкой строгого локального максимума функции f .
3. Если $d^2f(z, h)$ принимает значения разных знаков при некоторых $h \in \mathbb{R}^n$, то z не является точкой локального экстремума.

Доказательство:

Рассмотрим случай, когда второй дифференциал определён положительно. По доказанной ранее лемме найдётся $\varkappa > 0$ такое, что $d^2f(z, h) \geq \varkappa|h|^2$.

По формуле Тейлора с остаточным членом в форме Пеано имеем

$$f(z + h) = f(z) + \frac{1}{2}d^2f(z, h) + r(h)$$

где $r(h) = o(|h|^2)$ при $h \rightarrow 0$. По определению для любого $\varepsilon > 0$ существует $\delta > 0$ такое, что при $0 < |h| < \delta$ выполнено $|r(h)| \leq \varepsilon|h|^2$. При $\varepsilon = \frac{\varkappa}{4}$ из формулы Тейлора получаем

$$f(z + h) \geq \frac{1}{2}d^2f(z, h) - |r(h)| \geq \frac{1}{4}\varkappa|h|^2$$

откуда видно что z является точкой локального минимума функции f .

В случае, когда второй дифференциал определён отрицательно, рассуждения аналогичны. В случае, когда второй дифференциал неопределён, достаточно

МА-49. Теорема об обратном отображении. Криволинейные координаты

Пусть $\Omega \subset \mathbb{R}^n$ открыто. Если отображение $\varphi : \Omega \rightarrow \mathbb{R}^n$ непрерывно дифференцируемо в точке $x \in \Omega$, то под

$$J\varphi(x) := \det D\varphi(x)$$

будем понимать якобиан отображения φ в точке x .

Теорема 1: (об обратном отображении). Пусть отображение $\varphi : \Omega \rightarrow \mathbb{R}^n$ непрерывно дифференцируемо и $J\varphi \neq 0$ в точке $x^0 \in \Omega$ и $y^0 = \varphi(x^0)$. Тогда существуют $\varepsilon > 0$ и $\delta > 0$ и непрерывно дифференцируемое отображение

$$\psi : U_\varepsilon(y^0) \rightarrow U_\delta(x^0)$$

такое, что для любых $x \in U_\delta(x^0)$ и $y \in U_\varepsilon(y^0)$ равенство $\varphi(x) = y$ равносильно равенству $x = \psi(y)$. При этом

$$D\psi(y^0) = [D\varphi(x^0)]^{-1}$$

В частности, $J\psi(y^0) = 1/J\varphi(x^0)$

Доказательство:

Рассматриваемую систему уравнений $\varphi(x) = y$ можно записать в виде системы $F_i(x, y) = 0$, где $F_i(x, y) = \varphi_i(x) - y_i, i \in \overline{1, n}$. Так как $\frac{\partial(F_1, \dots, F_n)}{\partial(x_1, \dots, x_n)} \neq 0$, то по теореме о системе функций, заданных неявно, мы можем разрешить эту систему относительно x_1, \dots, x_n с помощью непрерывно дифференцируемого отображения $\psi : Q_{\varepsilon_1}(y^0) \rightarrow Q_{\delta_1}(x^0)$. Сузив прямоугольные окрестности $Q_{\varepsilon_1}(y^0)$ и $Q_{\delta_1}(x^0)$ (что возможно благодаря непрерывности ψ) мы получим искомые шары $U_\varepsilon(y^0)$ и $U_\delta(x^0)$.

Так как $\psi(\varphi(x)) = x$ в некоторой окрестности точки x_0 , то по теореме о дифференцировании композиции отображений

$$D\psi(\varphi(x)) \cdot D\varphi(x) = E$$

где E - единичная матрица размера $n \times n$. Отсюда получаем (1).

В теореме 1 не утверждается что ψ является биекцией, но это легко исправить: $J\varphi \neq 0$ в точке $x^0 \in \Omega$ то существуют открытые множества $U \subset \Omega, V \subset \mathbb{R}^n$ такие, что $x^0 \in U, \varphi(x^0) \in V, \varphi$ взаимно однозначно отображает U на V и обратное отображение φ^{-1} непрерывно дифференцируемо на V .

Доказательство:

Возьмём $V = U_\varepsilon(y^0)$ и $U = \varphi^{-1}(V) \cap U_\delta(x^0)$, где под $\varphi^{-1}(V)$ понимается полный прообраз V под действием φ . Так как φ непрерывно, то $\varphi^{-1}(V)$ открыто. Значит V открыто как пересечение двух открытых множеств.

По определению множеств U и V имеем $\varphi(U) \subset V$ и $\psi(V) \subset U$, то есть φ отображает U в V , а ψ отображает V в U . При этом для любых $x \in U$ и $y \in V$ по построению множеств U и V равенства $\varphi(x) = y$ и $x = \psi(y)$ равносильны. Тогда рассматриваемое отображение $\psi : V \rightarrow U$ является обратным по отношению к $\varphi : U \rightarrow V$.

Как известно, полный прообраз открытого множества под действием любого непрерывного отображения открыт. Образы открытых множеств не обладают аналогичным свойством: не всякое непрерывное (и даже непрерывно дифференцируемое) отображение переводит открытые множества в открытые (даже в одномерном случае; в качестве контрпримера можно рассмотреть \sin). Однако оказывается, что непрерывно дифференцируемые отображения с ненулевым якобианом всегда переводят открытые множества в открытые:

Теорема 3: Если отображение $\varphi : \Omega \rightarrow \mathbb{R}^n$ непрерывно дифференцируемо и $J\varphi \neq 0$ в каждой точке $x \in \Omega$, то множество $\varphi(\Omega)$ открыто.

Доказательство:

Пусть $x^0 \in \Omega$. По теореме 1 существует $\varepsilon > 0$ такое, что для любого $y \in U_\varepsilon(y^0)$ существует $x \in \Omega$ такое, что $\varphi(x) = y$. Значит $U_\varepsilon(y^0) \subset \varphi(\Omega)$, то есть y^0 является внутренней точкой множества Ω .

Определение 4: Если отображение $g : \Omega \rightarrow \mathbb{R}^n$ непрерывно дифференцируемо, инъективно и множество $g(\Omega)$ открыто, то говорят, что функции g_1, \dots, g_n образуют криволинейную систему координат на множестве Ω . При этом для любой точки $x \in \Omega$ числа $\tilde{x}_i = g_i(x)$ называются криволинейными координатами.

Как известно, в декартовой системе координат координатная сетка состоит из $(n - 1)$ -мерных плоскостей вида $x_i = \text{const}$. При введении криволинейных координат аналогичным образом возникает «криволинейная координатная сетка», которая в декартовой системе координат состоит из $(n - 1)$ -мерных поверхностей вида $g_i(x_1, \dots, x_n) = \text{const}$. В «новой» системе координат уравнения этих поверхностей имеют вид $\tilde{x}_i = \text{const}$.

МА-50. Условные экстремумы функций многих переменных

Пусть на открытом множестве $\Omega \subset \mathbb{R}^n$ заданы функция $f : \Omega \rightarrow \mathbb{R}$ и непрерывно дифференцируемое отображение $g : \Omega \rightarrow \mathbb{R}^m$. Рассмотрим множество

$$S = \{x \in \Omega : g(x) = 0\}$$

заданное системой ограничений

$$g_1(x) = \dots = g_m(x) = 0$$

(называемых также условиями связи).

Определение 1: Точка $z \in S$ называется точкой условного локального строгого минимума функции f при выполнении ограничений $g_1(x) = \dots = g_m(x) = 0$, если существует $\delta > 0$ такое, что для всех $x \in \overset{\circ}{U}_\delta(z) \cap S$ выполнено $f(x) > f(z)$. При этом значение функции f в точке z называется условным локальным минимумом.

Аналогично определяются условный локальный нестрогий минимум, а также строгий и нестрогий условный локальный максимум.

Локальные условные максимумы и минимумы (строгие или нестрогие) называются локальными условными экстремумами.

На практике локальными экстремумами часто называются именно точки локального экстремума (а не значения функции f в этих точках).

Если уравнения (1) можно разрешить относительно каких-то m переменных, то выражения для этих переменных можно подставить в функцию f и задача о нахождении условных экстремумов f сведётся к задаче о нахождении обычного (безусловного) экстремума некоторой функции от $m - n$ переменных.

Во многих приложениях возникают задачи, в которых уравнения (1) не удаётся в явном виде разрешить относительно каких-либо m переменных (либо это затруднительно). В таких случаях условные экстремумы функции f ищут с помощью метода множителей Лагранжа.

Несмотря на то, что в этом методе не требуется разрешать уравнения связи в явном виде, сам факт разрешимости этих уравнений важен для корректной работы данного метода. Как мы знаем из теоремы о системе функций, заданных неявно, для разрешимости уравнений (1) в некоторой окрестности точки $z \in S$ достаточно, чтобы якобиан g по каким-то m переменным был отличен от нуля, то есть $\text{rg } Dg = m$. В качестве примера (при $n = 2$ и $m = 1$) можно рассмотреть функцию $f(x, y) = x + y$ на множестве S , заданном уравнением $g(x, y) = 0$, где $g(x, y) = x^2/16 + y^2/9 - 1$. Так как линии уровня функции f представляют собой прямые вида $x + y = \text{const}$, то с геометрической точки зрения условные экстремумы f легко найти как точки эллипса S , в которых касательные к нему параллельны прямой $x + y = \text{const}$. Заметим, что в этих точках градиенты функций f и g параллельны.

При $n = 3$ и $m = 2$ множество S задано системой из двух уравнений: $g_1(x, y, z) = 0$ и $g_2(x, y, z) = 0$. С геометрической точки зрения это означает, что множество S представляет собой пересечение двух

поверхностей, то есть некоторую кривую. Если точка $M = (x^0, y^0, z^0) \in S$ является локальным условным экстремумом функции f , то производная f по направлению касательной к этой кривой в точке (x^0, y^0, z^0) должна быть равна нулю (если задать S параметрически в виду $x = x(t)$, то это следует из теоремы Ферма для функции $t \mapsto f(x(t))$ и теоремы о дифференцировании композиции отображений). Значит вектор $\nabla f(x^0, y^0, z^0)$ должен быть перпендикулярен этой касательной. С другой стороны, касательная к кривой S в точке M должна лежать в пересечении касательных плоскостей к поверхностям $g_1(x, y, z) = 0$ и $g_2(x, y, z) = 0$. Нормали к этим плоскостям параллельны векторам ∇g_1 и ∇g_2 (вычисленным в точке M). Таким образом, в точке (x^0, y^0, z^0) вектор ∇f должен быть линейной комбинацией векторов ∇g_1 и ∇g_2 с некоторыми коэффициентами λ_1 и λ_2 . Это равносильно тому, что в точке (x^0, y^0, z^0)

$$df = \lambda_1 dg_1 + \lambda_2 dg_2$$

что также равносильно тому, что точка (x^0, y^0, z^0) является стационарной точкой функции Лагранжа

$$L(x) = f(x) - \lambda_1 g_1(x) - \lambda_2 g_2(x)$$

При этом x не обязана быть стационарной точкой функции f . Отметим, что M также будет стационарной точкой L как функции от x и λ , так как $\frac{\partial L}{\partial \lambda_i} = g_i(x) = 0$ в силу условий связи.

Второе ключевое наблюдение, лежащее в основе метода множителей Лагранжа, состоит в том, что на множестве S функции g_i равны нулю и поэтому $f = L$ на множестве S . Следовательно на множестве S функции L и f имеют одни и те же условные экстремумы. При этом функция L лучше подходит для исследования на экстремум, так как для неё M является стационарной точкой (в отличие от функции f).

Третье ключевое наблюдение состоит в том, что в силу условий связи дифференциалы dx_1, \dots, dx_n , от которых зависит второй дифференциал функции Лагранжа, удовлетворяют системе линейных уравнений, которая получается дифференцированием равенств (1). Множество решений этой системы называется касательным пространством и исследовать d^2L на положительную/отрицательную определённости нужно лишь на этом пространстве. Для получения необходимых, а также достаточных условий условного экстремума удобно перейти в некоторой окрестности исследуемой точки к криволинейным координатам $y = (y_1, \dots, y_n)$, в которых пересечение S с рассматриваемой окрестностью задаётся уравнениями $y_1 = y_2 = \dots = y_m = 0$.

Лемма 2: (о выпрямляющей замене координат). Пусть $z \in \text{Surg } Dg(z) = m$. Тогда существуют окрестность U точки z , непрерывно дифференцируемое отображение $\varphi : U \rightarrow \mathbb{R}^n$

окрестность $V \subset \mathbb{R}^n$ точки $\varphi(z)$ такие, что

1. $\varphi_i(x) = g_i(x)$ для всех $x \in U$ и всех $i \in \overline{1, m}$
2. $\varphi : U \rightarrow V$ является биекцией и обратное отображение $\psi : V \rightarrow U$ непрерывно дифференцируемо.

Доказательство:

Так как $\text{rg } Dg(z) = m$, то существуют индексы $i_1, \dots, i_m \in \overline{1, n}$ такие, что $\frac{\partial(g_1, \dots, g_m)}{\partial(x_{i_1}, \dots, x_{i_m})} \neq 0$. Не уменьшая общности можно считать, что $i_1 = 1, \dots, i_m = m$. Положим

$$\varphi_i(x) = \begin{cases} g_i(x), & i \in \overline{1, m} \\ x_i, & i \in \overline{m+1, n} \end{cases}$$

для всех $x \in \Omega$. Тогда якобиан отображения φ в точке z равен

$$\frac{\partial(g_1, \dots, g_m, x_{m+1}, \dots, x_n)}{\partial(x_1, \dots, x_m, x_{m+1}, \dots, x_n)} = \det \frac{D(g_1, \dots, g_m)}{D(x_1, \dots, x_m)} \neq 0$$

По теореме об обратном отображении существуют окрестности $U, V \subset \mathbb{R}^n$ точек z и $\varphi(z)$ такие, что сужение φ на U взаимно однозначно отображает U на V и обратное отображение $\psi : V \rightarrow U$ непрерывно дифференцируемо.

Определение 3: Для любой точки $x \in S$ подпространство

$$T_x S := \{h \in \mathbb{R}^n : dg(x, h) = 0\}$$

будем называть касательным подпространством ко множеству S в точке x .

Легко видеть, что $T_x S = \ker Dg(x) = \bigcap_{i=1}^m \ker Dg_i$.

Лемма 4: Пусть $z \in S$ является точкой локального условного экстремума функции f при выполнении уравнений связи $g(x) = 0$, причём $\text{rg } Dg(z) = m$. Если f дифференцируема в точке z , то существуют единственный набор чисел $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ такой, что в точке z

дифференциал функции f представляется линейной комбинацией дифференциалов функций g_i с коэффициентами λ_i :

$$df(z, h) = \sum_{i=1}^m \lambda_i \cdot dg_i(z, h) \quad \forall h \in \mathbb{R}^n$$

Перейдём в окрестности U точки z к криволинейным координатам $y_i = \varphi_i(x)$, удовлетворяющим лемме 2. (Отображение, обратное к φ будем также обозначать как ψ .)

Очевидно, что z является локальным условным экстремумом функции f тогда и только тогда, когда $\tilde{z} = \varphi(z)$ является локальным условным экстремумом функции $\tilde{f}(y) = f(\psi(y))$ при выполнении условий $y_1 = y_2 = \dots = y_m = 0$. Последнее равносильно тому, что точка $(\varphi_{m+1}(z), \dots, \varphi_n(z))$ является локальным безусловным экстремумом функции $(y_{m+1}, \dots, y_n) \mapsto \tilde{f}(0, 0, \dots, 0, y_{m+1}, \dots, y_n)$. Тогда в силу необходимого условия безусловного экстремума получаем

$$\frac{\partial \tilde{f}(\tilde{z})}{\partial y_i} = 0, \quad i = m+1, \dots, n$$

Так как $f(x) = \tilde{f}(\varphi(x))$, то пользуясь инвариантностью дифференциала первого порядка получаем

$$df(z, h) = \sum_{i=1}^n \frac{\partial \tilde{f}(y)}{\partial y_i} \bigg|_{y=\tilde{z}} d\varphi_i(x, h) = \sum_{i=1}^m \lambda_i dg_i(x, h)$$

где $\lambda_i = \frac{\partial \tilde{f}(y)}{\partial y_i} \big|_{y=\tilde{z}}$. Единственность разложения следует из линейной независимости дифференциалов dg_i

Определение 5: Числа $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ называются множителями Лагранжа (в точке z), а функция

$$L(x) = f(x) - \sum_{i=1}^m \lambda_i g_i(x)$$

называется функцией Лагранжа (в точке z).

Заметим, что на множестве S значения функций f и L совпадают в силу условий связи. Поэтому функция f имеет локальный условный минимум в точке $z \in S$ тогда и только тогда, когда L имеет локальный условный минимум в точке z .

Лемма 6: (условная инвариантность дифференциала второго порядка). Пусть отображение v дважды дифференцируемо в точке $x \in \mathbb{R}^n$, а отображение u дважды дифференцируемо в точке $v(x) \in \mathbb{R}^m$. Если дифференциал в точке $v(x)$ равен нулю, то

$$d^2(u \circ v)(x, h) = d^2u(v(x), dv(x, h))$$

Доказательство:

Так как $v(x+h) - v(x) = O(|h|)$ при $h \rightarrow 0$, то по формуле Тейлора для функции u в точке $v(x)$ имеем

$$u(v(x+h)) = u(v(x)) + du(v(x), v(x+h) - v(x)) + \frac{1}{2} d^2u(v(x), v(x+h) - v(x)) + o(|h|^2)$$

при $h \rightarrow 0$. Второе слагаемое в правой части равно нулю по условию, а кроме того $v(x+h) - v(x) = dv(x, h) + o(|h|) = O(|h|)$ при $h \rightarrow 0$. Так как $d^2u(y, \xi)$ является линейной комбинацией слагаемых вида $\xi_i \cdot \xi_j$, то записанное выше равенство принимает вид

$$u(v(x+h)) = u(v(x)) + \frac{1}{2} d^2u(v(x), dv(x, h)) + o(|h|^2)$$

что и требовалось доказать.

Теорема 7: (необходимое условие локального условного экстремума). Пусть $z \in S$ является точкой локального минимума (максимума) функции f при выполнении уравнений связи $g(x) = 0$, причём $\text{rg } Dg(z) = m$. Если f дифференцируема в точке z , то существуют единственный набор чисел $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ такой, что точка z является стационарной точкой функции Лагранжа L .

Если f дважды дифференцируема в точке z , то сужение второго дифференциала $d^2L(z, \cdot)$ на касательное подпространство $\ker Dg(z)$ представляет собой квадратичную форму, которая положительно полуопределена в случае локального минимума и отрицательно полуопределена в случае локального максимума.

При нарушении условия $\text{rg } Dg(z) = m$ точка локального условного экстремума может не быть стационарной точкой функции Лагранжа. Например, рассмотрим функции $f(x, y, z) = x+y$, $g_1(x, y, z) = z$ и $g_2(x, y, z) = (x+y)^2 + z$. Точка $M = (0, 0, 0)$ является локальным условным экстремумом функции f , но при этом в точке M дифференциал $df = dx + dy$ не является линейной комбинацией дифференциалов $dg_1 = dz$ и $dg_2 = dz$.

Доказательство:

Продолжим рассуждение из доказательства леммы 4. Условие стационарности точки z для функции

Лагранжа равносильно равенству (2).

Обозначим $\tilde{L}(y) = L(\psi(y))$

В силу необходимого условия локального минимума квадратичная форма $\nu \mapsto d^2\tilde{L}(\varphi(z), \nu)$ определена неотрицательно на подпространстве $\{\nu \in \mathbb{R}^n : \nu_1 = \dots = \nu_m = 0\}$. (Действительно, при подстановке $\nu_1 = \dots = \nu_m = 0$ в эту квадратичную форму получим второй дифференциал функции $(y_{m+1}, \dots, y_n) \mapsto \tilde{L}(0, \dots, 0, y_{m+1}, \dots, y_n)$, который представляет собой неотрицательно определённую квадратичную форму по необходимому условию безусловного локального минимума.)

Тогда по лемме 6

$$d^2L(z, h) = d^2(\tilde{L} \circ \varphi)(z, h) = d^2\tilde{L}(\varphi(z), d\varphi(z, h))$$

Если вектор $h \in \mathbb{R}^n$ таков, что первые m компонент вектора $d\varphi(z, h)$ равны нулю, то $d^2\tilde{L}(\varphi(z), d\varphi(z, h)) \geq 0$ и, следовательно, $d^2L(z, h) \geq 0$

Теорема 8: (достаточное условие локального экстремума). Пусть f дважды дифференцируема в точке $z \in S$ и эта точка является стационарной точкой функции Лагранжа $L(x) = f(x) - \sum_{i=1}^m \lambda_i g_i(x)$

1. Если квадратичная форма $h \mapsto d^2L(z, h)$ определена положительно на касательном пространстве $\ker Dg(z)$, то точка z является точкой локального условного минимума;
2. Если квадратичная форма $h \mapsto d^2L(z, h)$ определена отрицательно на касательном пространстве $\ker Dg(z)$, то точка z является точкой локального условного максимума;
3. Если квадратичная форма $h \mapsto d^2L(z, h)$ не является знакоопределённой на касательном пространстве $\ker Dg(z)$, то точка z не является точкой локального условного экстремума.

Доказательство:

Будем рассуждать так же, как в доказательстве теоремы 7.

Пусть отображения φ и ψ удовлетворяют условию леммы 2, а также $\tilde{L}(y) = L(\psi(y))$.

Предположим, что выполнено первое предположение и докажем, что z является точкой локального минимума функции Лагранжа.

По условию с учётом леммы 6 для всех векторов $h \in \ker Dg(z) \setminus \{0\}$ выполнено

$$0 < d^2L(z, h) = d^2(\tilde{L} \circ \varphi)(z, h) = d^2\tilde{L}(\varphi(z), d\varphi(z, h))$$

Так как матрица Якоби $D\varphi(z)$ невырождена, то любой вектор $\nu \in \mathbb{R}^n \setminus \{0\}$, удовлетворяющий условию $\nu_1 = \dots = \nu_m = 0$, можно представить в виде $\nu = d\varphi(z, h)$, где $h \in \ker Dg(z) \setminus \{0\}$. Следовательно,

$$d^2\tilde{L}(\varphi(z), \nu) = d^2\tilde{L}(\varphi(z), d\varphi(z, h)) > 0$$

Таким образом второй дифференциал функции $(y_{m+1}, \dots, y_n) \mapsto \tilde{L}(0, \dots, 0, y_{m+1}, \dots, y_n)$ в точке $\varphi(z)$ является положительно определённой квадратичной формой. Значит $\varphi(z)$ является точкой локального минимума функции \tilde{L} .

В случае локального минимума рассуждения аналогичны. В случае когда квадратичная форма $d^2L(z, h)$ принимает на подпространстве $\ker Dg(z)$ значения разных знаков такие же рассуждения приводят к противоречию с необходимым условием условного экстремума.

Пример применения метода Лагранжа

Задача 9. С помощью метода множителей Лагранжа исследовать функцию

$$f(x, y, z) = x^2 + y^2 + 2z$$

на условный экстремум при выполнении условий связи

$$x + y + z = \frac{1}{2}, \quad x^2 + y^2 + 2z^2 = 1$$

Прежде всего заметим, что в силу второго условия связи $x^2 + y^2 = 1 - 2z^2$, поэтому исследуемую на условный экстремум функцию f можно заменить на

$$f(x, y, z) := 1 - 2z^2 + 2z$$

(Такая функция зависит только от переменной z и это упростит систему уравнений, которую нам предстоит решать.)

Введём обозначения для функций, задающих условия связи:

$$g(x, y, z) = x + y + z - \frac{1}{2}, \quad h(x, y, z) = x^2 + y^2 + 2z^2 - 1$$

Рассмотрим функцию Лагранжа:

$$L(x, y, z) := f(x, y, z) - \lambda g(x, y, z) - \mu h(x, y, z)$$

Шаг 1. Запишем матрицу Якоби для отображения G с компонентами (g, h) :

$$DG = \frac{D(g, h)}{D(x, y, z)} = \begin{bmatrix} \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} & \frac{\partial g}{\partial z} \\ \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} & \frac{\partial h}{\partial z} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 2x & 2y & 4z \end{bmatrix}$$

В необходимых (и достаточных) условиях условного экстремума требуется, чтобы $\text{rg} DG = 2$ в исследуемых точках (то есть на всём множестве S , заданном условиями связи). Легко видеть, что $\text{rg} DG < 2$ тогда и только тогда, когда

$$x = y = 2z$$

Если при этом $g(x, y, z) = 0$, то $z = \frac{1}{10}$. Тогда $g(x, y, z) = 10z^2 - 1 \neq 0$. Значит $\text{rg} DG = 2$ на всём множестве S .

Шаг 2. Найдём все стационарные точки функции Лагранжа. Для этого составим систему уравнений

$$\begin{cases} \frac{\partial L}{\partial x} = 0 \\ \frac{\partial L}{\partial y} = 0 \\ \frac{\partial L}{\partial z} = 0 \\ g(x, y, z) = 0 \\ h(x, y, z) = 0 \end{cases}$$

(Последние два уравнения равносильны тому, что $\frac{\partial L}{\partial \lambda} = 0$ и $\frac{\partial L}{\partial \mu} = 0$.) Вычислив частные производные функции Лагранжа, получим нелинейную систему из 5 уравнений относительно неизвестных x, y, z, λ, μ :

$$\begin{cases} -\lambda - 2\mu x = 0 \\ -\lambda - 2\mu y = 0 \\ -\lambda - 4\mu z - 4z + 2 = 0 \\ x + y + z = \frac{1}{2} \\ x^2 + y^2 + 2z^2 = 1 \end{cases}$$

Основная трудность при решении задач данного типа с помощью метода множителей Лагранжа, как правило, состоит в решении этой системы.

Начнём с рассмотрения случая $\mu = 0$. Тогда $\lambda = 0$ и из третьего уравнения $z = \frac{1}{2}$. Из четвёртого уравнения $y = -x$, и тогда пятое уравнение принимает вид $2x^2 = \frac{1}{2}$, откуда $x = \pm \frac{1}{2}$. Итак, при $\lambda = \mu = 0$ стационарными точками являются $(\frac{1}{2}, -\frac{1}{2}, \frac{1}{2})$ и $(-\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$.

В рассматриваемой нами ситуации легко выразить x, y, z через λ и μ :

$$x = y = \frac{-\lambda}{2\mu}, \quad z = \frac{2 - \lambda}{4(1 + \mu)}$$

Теперь можно подставить эти выражения в условия связи и получить нелинейную систему из двух уравнений относительно λ и μ . Однако такая система получится довольно громоздкой, и мы этого делать не будем. Вместо этого подставим полученное равенство $x = y$ в условия связи:

$$\begin{cases} 2x + z = \frac{1}{2} \\ 2x^2 + 2z^2 = 1 \end{cases}$$

Подставив z из первого уравнения во второе получим квадратное уравнение относительно x :

$$10x^2 - 4x - \frac{1}{2} = 0$$

Его корни $x_1 = \frac{1}{2}, x_2 = -\frac{1}{10}$. Соответственно получаем $y_1 = x_1, z_1 = \frac{1}{2} - 2x_1 = -\frac{1}{2}$ и $y_2 = x_1, z_2 = \frac{1}{2} - 2x_1 = \frac{7}{10}$

Итак, мы нашли следующие стационарные точки функции Лагранжа:

1. $(\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}), \lambda = 0, \mu = 0$
2. $(-\frac{1}{2}, \frac{1}{2}, \frac{1}{2}), \lambda = 0, \mu = 0;$
3. $(\frac{1}{2}, \frac{1}{2}, -\frac{1}{2})$
4. $(-\frac{1}{10}, -\frac{1}{10}, \frac{7}{10})$

Для последних двух точек мы пока что не нашли λ и μ . Их можно найти разрешив (3) относительно λ и μ . Так как в данный момент эти значения нам не нужны, то отложим это действие до шага 5. Шаг 3. В каждой из найденных стационарных точек найдём касательное пространство. По определению

это пространство в точке (x, y, z) , удовлетворяющей уравнениям связи, состоит из всех векторов (dx, dy, dz) таких, что

$$DG \cdot \begin{bmatrix} dx \\ dy \\ dz \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 2x & 2y & 4z \end{bmatrix} \cdot \begin{bmatrix} dx \\ dy \\ dz \end{bmatrix} = 0$$

то есть dx, dy, dz должны удовлетворять системе

$$\begin{cases} dx + dy + dz = 0 \\ xdx + ydy + 2zdz = 0 \end{cases}$$

Шаг 4. Вычислим второй дифференциал функции Лагранжа:

$$d^2L = -2\mu dx^2 - 2\mu dy^2 - 4(\mu + 1)dz^2$$

Так как в рассматриваемой задаче d^2L не зависит от λ , то в последних двух найденных стационарных точках нам требуется найти только λ , а μ можно не вычислять.

Шаг 5. Исследуем сужение второго дифференциала функции Лагранжа на касательное пространство для каждой из найденных стационарных точек.

Шаг 5.1. Рассмотрим точку $(\frac{1}{2}, -\frac{1}{2}, \frac{1}{2})$, $\lambda = 0, \mu = 0$.

Касательное пространство (4) задаётся равенствами

$$\begin{aligned} dx + dy + dz &= 0, \\ dx - dy + 2dz &= 0 \end{aligned}$$

Например, мы можем выразить dx и dy через dz . Но так как в данном случае $\mu = 0$, то явные выражения роли не играют (важен лишь факт, что dx и dy выражаются через dz). Мы получаем

$$d^2L = -4dz^2$$

Эта (уже одномерная!) квадратичная форма определена отрицательно, значит имеем локальный условный максимум.

Шаг 5.2. Рассмотрим точку $(-\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$, $\lambda = 0, \mu = 0$.

Касательное пространство (4) задаётся равенствами

$$\begin{aligned} dx + dy + dz &= 0 \\ -dx + dy + 2dz &= 0 \end{aligned}$$

поэтому мы снова можем выразить dx и dy через dz . Аналогично предыдущему случаю получаем локальный условный максимум.

Шаг 5.3. Рассмотрим точку $(\frac{1}{2}, \frac{1}{2}, -\frac{1}{2})$.

$$\begin{aligned} dx + dy + dz &= 0 \\ dx + dy - 2dz &= 0 \end{aligned}$$

В данном случае мы уже не можем выразить dx и dy через dz . Но из уравнений касательного пространства $dz = 0$ и, следовательно, $dy = -dx$. Подставив найденные выражения для dz и dy во второй дифференциал функции Лагранжа, получим одномерную квадратичную форму

$$d^2L = -4\mu dx^2$$

Из уравнений (3) находим, что $\lambda = -\mu$ и $\mu = -\frac{4}{3}$. Значит d^2L (на касательном пространстве) определена положительно. Поэтому в рассматриваемой точке имеется локальный условный минимум.

Шаг 5.4. Рассмотрим точку $(-\frac{1}{10}, -\frac{1}{10}, \frac{7}{10})$.

Касательное пространство (4) задаётся равенствами

$$\begin{aligned} dx + dy + dz &= 0 \\ dx + dy - 14dz &= 0 \end{aligned}$$

Снова $dz = 0$ и $dy = -dx$, то есть на касательном пространстве d^2L становится одномерной квадратичной формой, зависящей только от dx :

$$d^2L = -4\mu dx^2$$

Из уравнений (3) находим, что $\lambda = -\mu/5$ и $\mu = -\frac{4}{15}$. Значит d^2L (на касательном пространстве) определена положительно. Поэтому в рассматриваемой точке имеется локальный условный минимум.

Шаг 6. Ответ.

1. $(\frac{1}{2}, -\frac{1}{2}, \frac{1}{2})$ - локальный условный максимум;
2. $(-\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ - локальный условный максимум;
3. $(\frac{1}{2}, \frac{1}{2}, -\frac{1}{2})$ - локальный условный минимум;
4. $(-\frac{1}{10}, -\frac{1}{10}, \frac{7}{10})$ - локальный условный минимум.

Ограничения, заданные неравенствами*

Рассмотрим случай, когда вместо условий связи вида $g_i(x) = 0$ для всех $i = 1, \dots, m$ рассматриваются неравенства вида $g_i(x) \leq 0$ для всех $i = 1, \dots, m$. Как и выше будем предполагать, что функции f и g_i дважды дифференцируемы. Также будем предполагать, что в

каждой точке x дифференциалы всех функций g_i , обращающихся в нуль в x , линейно независимы.

Теорема 10: (условия Каруша-Куна-Таккера). Если точка z является точкой локального условного минимума функции f при выполнении неравенств $g_i(x) \leq 0, i = \overline{1, m}$, то существуют числа $\lambda_i \leq 0 (i \in \overline{1, m})$ такие, что имеет место равенство (2), причём для всех $i \in \overline{1, m}$ выполнены условия дополняющей нежёсткости:

$$\lambda_i g_i(z) = 0$$

При этом для функции Лагранжа $L = f - \sum_{i=1}^m \lambda_i g_i$ второй дифференциал $d^2L(z, \cdot)$ определён неотрицательно на подпространстве

$$T = \bigcap_{i=1}^m \{h \in \mathbb{R}^n : \lambda_i \cdot dg_i(z, h) = 0\}.$$

Обратно, если для некоторого набора чисел $\lambda_i \leq 0$ имеют место равенства (2) и (5), а второй дифференциал $d^2L(z, \cdot)$ определён положительно на подпространстве T , то точка z является локальным условным минимумом функции f .

Доказательство:

Перейдём (как выше) к криволинейной системе координат $y = \varphi(x)$, в которой условия $g_i(x) \leq 0$ принимают вид $y_i \leq 0, i = 1, \dots, m$.

Предположим что в точке $z \in \Omega$ функция f имеет локальный минимум. Напомним, что $\tilde{f}(y) = f(\varphi(y))$ - это запись функции f в рассматриваемой криволинейной системе координат. Пусть $\tilde{z} = \varphi(z)$

Обозначим через A множество тех индексов $i \in \overline{1, m}$, для которых $g_i(z) = 0$ (такие ограничения g_i называют активными).

Для $i \notin A$ (включая случай $i \in \overline{m+1, n}$) функция $t \mapsto \tilde{f}(\tilde{z}_1, \dots, \tilde{z}_{i-1}, \tilde{z}_i + t, \tilde{z}_{i+1}, \dots)$ имеет локальный минимум в нуле и, следовательно,

$$\lambda_i := \frac{\partial \tilde{f}(\tilde{z})}{\partial y_i} = 0$$

Для $i \in A$ функция $t \mapsto \tilde{f}(\tilde{z}_1, \dots, \tilde{z}_{i-1}, \tilde{z}_i + t, \tilde{z}_{i+1}, \dots)$ должна нестрого возрастать при убывании $t \leq 0$, следовательно

$$\lambda_i := \frac{\partial \tilde{f}(\tilde{z})}{\partial y_i} \leq 0$$

Таким образом, нашлись числа $\lambda_1, \dots, \lambda_m \leq 0$ такие, что имеют место равенства (2) и (5). (В силу предположения о линейной независимости дифференциалов функций g_i в точке z числа λ_i определены однозначно.)

Заметим, что всякий локальный условный минимум функции f является локальным условным минимумом функции Лагранжа $L(x) = f(x) - \sum_{i=1}^m \lambda_i g_i(x)$. Действительно, $\lambda_i \leq 0$ и для всех x

Примечания
