.

КУРС МАТЕМАТИЧЕСКОГО АНАЛИЗА

Семестр 2

для экономического факультета СПбГУ

А.Е.Барабанов

Санкт–Петербург
2003

Функции нескольких переменных

Определение.

Функцией нескольких переменных называется отображение $f : D \to \mathsf{R}$, где D — некоторое подмножество R^n. □

Область определения D функции f состоит из векторов $x \in \mathsf{R}^n$, которые имеют n компонент: $x = (x_1, x_2, \ldots, x_n)$. Поэтому значение функции f иногда записывают как $f(x)$, где x — вектор, а иногда как $f(x_1, x_2, \ldots, x_n)$.

По определению графика произвольного отображения, графиком функции f является подмножество G пространства R^{n+1}, состоящее из всех точек следующего вида:

$$G = \{(x_1, x_2, \ldots, x_n, y) \mid (x_1, x_2, \ldots, x_n) \in D, y = f(x_1, x_2, \ldots, x_n)\}.$$

Обычно график функции двух переменных изображается в виде поверхности в трехмерном пространстве. Примером такой поверхности является горная местность, где высота над уровнем моря является функцией от положения точки на карте.

Определение *некоторых топологических свойств подмножеств в* R^n.

1. *Окрестностью* точки x в R^n называется любое подмножество в R^n, содержащее некоторый шар с центром в точке x.

2. Пусть $x \in \mathsf{R}^n$ и $D \subseteq \mathsf{R}^n$. Точка x называется *внутренней точкой* D, если существует окрестность точки x, полностью содержащаяся в D.

3. Пусть $x \in \mathsf{R}^n$ и $D \subseteq \mathsf{R}^n$. Точка x называется *граничной точкой* D, если любая окрестность точки x содержит как точки из D, так и точки не из D. (При этом x может не принадлежать D, находясь на его границе.)

4. Подмножество D в R^n называется *открытым*, если оно состоит только из внутренних точек.

5. Подмножество D в R^n называется *замкнутым*, если его дополнение в R^n открыто (или, что то же самое, D содержит все свои граничные точки).

6. Подмножество D в R^n называется *выпуклым*, если для любой пары точек x_1, x_2 из D отрезок $[x_1, x_2]$ в R^n полностью принадлежит D.

7. Подмножество D в R^n называется *связным*, если любую пару точек x_1, x_2 в D можно соединить непрерывной линией с концами в этих точках, лежащей полностью в D.

8. Подмножество D в R^n называется *ограниченным*, если существует шар в R^n, целиком содержащий это множество.

9. Подмножество D в R^n называется *компактным*, если оно замкнуто и ограничено.

10. Подмножество D в R^n называется *областью*, если оно открыто и связно.

Определение *нормы вектора в* R^n.

Нормой элемента (или вектора) x в R^n называется любая функция $p(x)$, обладающая свойствами положительности ($p(x) > 0$ при $x \neq 0$), положительной однородности ($p(\lambda x) = |\lambda| p(x) \, \forall \lambda \in \mathsf{R}$) и удовлетворяющая неравенству треугольника ($p(x) + p(y) \geq p(x+y)$). □

Наиболее распространены следующие нормы:

- кубическая: $p(x_1, x_2, \ldots, x_n) = \max\{|x_1|, |x_2|, \ldots, |x_n|\}$;
- октаэдрическая: $p(x_1, x_2, \ldots, x_n) = |x_1| + |x_2| + \cdots + |x_n|$;
- евклидова: $p(x_1, x_2, \ldots, x_n) = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$.

В трехмерном пространстве евклидова норма x совпадает с длиной радиуса–вектора, соединяющего начало координат с точкой x. Нормы в R^n обычно обозначаются символом $\|x\|$ вместо $p(x)$.

Для любой нормы $p(x)$ и любого вектора $x_0 \in \mathsf{R}^n$ множество $\{x \in \mathsf{R}^n \mid p(x - x_0) < \delta\}$ является окрестностью точки x_0 (при $\delta > 0$).

Нетрудно доказать, что каждая из перечисленных норм задает одно и то же множество окрестностей точки x_0 (т.е. множеств, содержащих шары $\{x | p(x - x_0) < \delta\}$). В этом смысле все эти нормы эквивалентны. Наиболее распространенной является евклидова норма. В дальнейшем она будет всегда подразумеваться, если конкретная норма не указана.

Понятие нормы заменяет понятие абсолютной величины числа. В соответствии с этим понятием несколько изменяется определение предела.

Определение *предела и непрерывности функции в* R^n.

Пусть D — область в R^n и функция f задана на D. Число A называется *пределом функции* f *в точке* a, если

$$\forall \epsilon > 0 \; \exists \delta > 0 \;\; : \;\; \text{если } \|x - a\| < \delta, \; x \in D, \; x \neq x_0, \; \text{то } |f(x) - A| < \epsilon.$$

В этом случае пишут $\lim_{x \to a} f(x) = A$.

Функция f называется *непрерывной в точке* $a \in D$, если f имеет в a предел, равный $f(a)$.

Пусть $D_1 \subseteq D$. Функция f называется *непрерывной на множестве* D_1, если она непрерывна в каждой точке этого множества. □

Для непрерывных функций нескольких переменных остаются верными все теоремы о непрерывных функциях на отрезке, доказанные для функций одной переменной. Они сформулированы ниже без доказательств.

Теорема *об ограниченности непрерывной функции.*

Пусть функция f задана и непрерывна на компактном множестве D. Тогда эта функция ограничена на D, т.е. существует такое число $M > 0$, что $|f(x)| < M$ для любого $x \in D$.

Теорема *о достижении максимума и минимума.*

Пусть функция f задана и непрерывна на компактном множестве D. Тогда f достигает наибольшего и наименьшего значений на D, т.е. существуют такие точки $c \in D$ и $d \in D$, что $f(c) \leq f(x) \leq f(d)$ для любого $x \in D$.

Теорема *об образе связного компактного множества при непрерывном отображении.*

Пусть функция f задана и непрерывна на компактном связном множестве D. Тогда образ множества D есть промежуток, т.е. множество чисел $f(D) = \{y = f(x) \mid x \in D\}$ есть отрезок прямой $[m, M]$, в котором m есть наименьшее значение f на D, а M — наибольшее значение f на D.

Теорема *о равномерной непрерывности.*

Пусть функция f задана и непрерывна на компактном множестве D. Тогда f равномерно непрерывна на D, т.е.

$$\forall \epsilon > 0 \ \exists \delta > 0 \ : \ \forall x, y \in D, \text{ если } \|x - y\| < \delta, \text{ то } |f(x) - f(y)| < \epsilon.$$

Дифференцируемость функции. Частные производные

Линейной функцией в R^n называется функция вида

$$\ell(x_1, x_2, \ldots, x_n) = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n,$$

где a_1, a_2, …, a_n — некоторые числа.

Очевидно, областью определения линейной функции является все пространство R^n.

Пусть функция f задана в области D и точка $z = (z_1, z_2, \ldots, z_n)$ принадлежит D. Приращением функции f в точке z называется функция $\Delta_z f$, определяемая равенством

$$\Delta_z f(h) = f(z + h) - f(z).$$

Функция определена для тех $h \in \mathsf{R}^n$, для которых $z + h \in D$. В частности, при достаточно малых h функция $\Delta_z f$ определена, так как точка z — внутренняя для D по определению области в R^n.

Определение *дифференцируемой функции в точке.*

Пусть функция f задана в области D и точка $z = (z_1, z_2, \ldots, z_n)$ принадлежит D. Пусть существует такая линейная функция ℓ, что

$$\lim_{h \to 0} \frac{\Delta_z f(h) - \ell(h)}{\|h\|} = 0.$$

Тогда функция f называется *дифференцируемой в точке* z, а функция $\ell(h)$ называется *дифференциалом функции* f *в точке* z. □

Обозначения для дифференциала в научной и прикладной литературе стандартизованы на протяжении столетий. А именно, приращение аргумента h всегда обозначается dx, чем подчеркивается, что это дифференциал (d) по переменной x. Это, конечно, вектор: $dx = (dx_1, dx_2, \ldots, dx_n)$. Значение линейной функции записывается в виде df, так что

$$df = a_1 dx_1 + a_2 dx_2 + \cdots + a_n dx_n.$$

Определение дифференциала можно записать также при помощи понятия бесконечно малой функции $o(1)$, как функции, стремящейся к нулю в указанной точке (т.е. при $h \to 0$). Следующая запись условия дифференцируемости функции является наиболее употребительной:

$$f(z+h) = f(z) + a_1h_1 + a_2h_2 + \cdots + a_nh_n + \|h\|o(1)$$

при $h \to 0$.

Иногда записью подчеркивают, что приращение функции примерно равно ее дифференциалу с точностью до бесконечно малой более высокого порядка:

$$f(z+h) - f(z) = df(h) + \|h\|o(1)$$

при $h \to 0$.

Не менее важным понятием, чем дифференциал функции в точке, является понятие частных производных, заменяющее понятие производной функции одной переменной.

Определение *частной производной.*

Пусть функция f задана в области D и точка $z = (z_1, z_2, \ldots, z_n)$ принадлежит D. Выберем индекс k между 1 и n. Определим функцию одной переменной, фиксировав в f все аргументы, кроме k-го:

$$\zeta(x_k) = f(z_1, \cdots, z_{k-1}, x_k, z_{k+1}, \cdots, z_n).$$

Пусть функция ζ дифференцируема в точке z_k. Тогда ее производная $\zeta'(z_k)$ называется *частной производной функции f в точке z по k-ой переменной.* Она обозначается

$$\frac{\partial f}{\partial x_k}(z_1, z_2, \cdots, z_n).$$

□

Другими словами, частная производная получается как обычная производная функции f, в которой все аргументы фиксированы, кроме k-го, по которому производная и вычисляется.

У функции $f(x_1, x_2, \ldots, x_n)$ может быть всего n частных производных в одной точке. Однако, возможны случаи, когда часть этих производных существует, а часть — нет.

Дифференциал и частные производные тесно связаны друг с другом. Пусть функция f дифференцируема в точке $z = (z_1, z_2, \ldots, z_n)$ и дифференциалом функции f в точке z является функция

$$\ell(h) = a_1h_1 + a_2h_2 + \cdots + a_nh_n.$$

Выберем индекс k между 1 и n. Приращение аргумента $h = (h_1, h_2, \ldots, h_n)$ определим следующим образом: $h_i = 0$ при $i \neq k$, а число $h_k \neq 0$ определяет величину приращения. Тогда по определению дифференциала,

$$f(z_1, \ldots, z_{k-1}, z_k + h_k, z_{k+1}, \ldots, z_n) - f(z_1, \ldots, z_{k-1}, z_k, z_{k+1}, \ldots, z_n) = a_kh_k + |h_k|o(1)$$

при $h_k \to 0$.

Но в левой части стоит приращение функции $\zeta(h_k)$ из определения частной производной. Разделив это приращение на h_k, получим:

$$\frac{\partial f}{\partial x_k}(z) = \lim_{h_k \to 0} \frac{\zeta(z_k + h_k) - \zeta(z_k)}{h_k} = \lim_{h_k \to 0} \left(a_k + \frac{|h_k|}{h_k} o(1) \right) = a_k.$$

Таким образом, коэффициенты линейной функции — дифференциала — равны частным производным функции f.

Теорема. *Формула полного дифференциала.*

Пусть функция f задана в области D, $z = (z_1, z_2, \ldots, z_n) \in D$ и функция f дифференцируема в точке z.

Тогда в точке z существуют все частные производные функции f и дифференциал может быть записан в виде

$$df = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \cdots + \frac{\partial f}{\partial x_n} dx_n,$$

где все частные производные вычисляются в точке z. □

Таким образом, у дифференцируемой функции есть все частные производные. Обратное, как правило, тоже верно, но не всегда. Существуют примеры функций, у которых есть все частные производные в точке, но которые в этой точке недифференцируемы, т.е. дифференциала нет.

Теорема. *Достаточные условия дифференцируемости функции.*

Пусть функция f определена в некоторой окрестности точки z, имеет там все частные производные и эти производные непрерывны. Тогда существует дифференциал функции f в точке z. □

(Без доказательства.)

Определение *класса функций $C^1(D)$.*

Множество всех функций, заданных и имеющих все непрерывные частные производные в области D, обозначается $C^1(D)$. □

Функции из $C^1(D)$ дифференцируемы в каждой точке области D по предыдущей теореме.

Производная сложной функции

Пусть функция $f(x, y)$ двух переменных определена в области $D \subseteq \mathsf{R}^2$. Пусть, кроме того, заданы две функции $\phi(t)$ и $\psi(t)$ на одном промежутке (t_1, t_2) и так, что при всех t из этого промежутка вектор $z = (\phi(t), \psi(t))$ находится в области D.

Тогда можно определить сложную функцию $\xi : (t_1, t_2) \to \mathsf{R}$ по правилу $\xi(t) = f(\phi(t), \psi(t))$. Иногда это отображение ξ определяют в менее точных терминах: рассмотрим функцию $f(x, y)$, в которой $x = \phi(t)$, а $y = \psi(t)$, так что получается функция от t. Задача состоит в вычислении производной этой сложной функции.

Рассмотрим некоторую точку t на промежутке (t_1, t_2) и соответствующие точки $x = \phi(t)$, $y = \psi(t)$. Приращение аргумента t обозначим Δt, а соответствующие приращения остальных функций —

$$\begin{aligned}
\Delta x &= \phi(t+\Delta t) - \phi(t) = \phi(t+\Delta t) - x,\\
\Delta y &= \psi(t+\Delta t) - \psi(t) = \psi(t+\Delta t) - y,\\
\Delta f &= f(x+\Delta x, y+\Delta y) - f(x,y),\\
\Delta \xi &= \xi(t+\Delta t) - \xi(t).
\end{aligned}$$

Из обозначений следует, что $\Delta f = \Delta\xi$.

Предположим, что функции f, ϕ, ψ непрерывно дифференцируемы в области определения. Применим теорему Лагранжа о конечных приращениях:

$$\begin{aligned}
\Delta\xi &= f(x+\Delta x, y+\Delta y) - f(x,y)\\
&= [f(x+\Delta x, y+\Delta y) - f(x, y+\Delta y)] + [f(x, y+\Delta y) - f(x,y)]\\
&= \frac{\partial f}{\partial x}(x+\vartheta_1\Delta x, y+\Delta y)\Delta x + \frac{\partial f}{\partial y}(x, y+\vartheta_2\Delta y)\Delta y,
\end{aligned}$$

где $0 < \vartheta_1 < 1$, $0 < \vartheta_2 < 1$.

Разделим последнее равенство на Δt и вычислим предел при $\Delta t \to 0$:

$$\xi'(t) = \frac{\partial f}{\partial x}(x,y)\,\frac{d\phi}{dt}(t) + \frac{\partial f}{\partial y}(x,y)\,\frac{d\psi}{dt}(t).$$

Полученное равенство легко обобщить для функций n переменных.

Теорема. *Производная сложной функции.*

Пусть функция $y = f(x_1, x_2, \ldots, x_n)$ непрерывно дифференцируема в области $D \subseteq \mathsf{R}^n$. Пусть на промежутке (t_1, t_2) заданы непрерывно дифференцируемые функции $\widetilde{x}_1$, $\widetilde{x}_2$, …, $\widetilde{x}_n$, причем вектор $\widetilde{x}(t) = (\widetilde{x}_1(t), \widetilde{x}_2(t), \ldots, \widetilde{x}_n(t))$ принадлежит D при всех $t \in (t_1, t_2)$.

Тогда сложная функция $\widetilde{y}(t) = f(\widetilde{x}(t))$ дифференцируема на промежутке (t_1, t_2), а ее производную можно записать в виде

$$\frac{d\widetilde{y}}{dt} = \frac{\partial f}{\partial x_1}\,\frac{d\widetilde{x}_1}{dt} + \frac{\partial f}{\partial x_2}\,\frac{d\widetilde{x}_2}{dt} + \cdots + \frac{\partial f}{\partial x_n}\,\frac{d\widetilde{x}_n}{dt}.$$

□

Касательная плоскость к графику функции

Говорят, что подмножество L пространства R^n задано уравнением

$$f(x_1, x_2, \ldots, x_n) = 0,$$

если координаты точек из L удовлетворяют этому уравнению и наоборот, координаты остальных точек не удовлетворяют ему.

Гиперплоскость в пространстве R^n — это аналог обычной плоскости в трехмерном пространстве. Гиперплоскость разбивает пространство на две части, два полупространства, "по разную сторону от гиперплоскости". Любая гиперплоскость задается уравнением

$$A_1x_1 + A_2x_2 + \cdots + A_nx_n = C,$$

где A_1, A_2, ..., A_n и C — некоторые числа. Естественно, исключается уравнение, в котором все эти числа равны нулю. Одно полупространство получается заменой в уравнении гиперплоскости знака "=" на знак "<", а второе — на знак ">".

Рассмотрим вектор $A = (A_1, A_2, \ldots, A_n)$. Пользуясь определением скалярного произведения векторов

$$\langle A, x\rangle = A_1x_1 + A_2x_2 + \cdots + A_nx_n,$$

запишем уравнение гиперплоскости в виде $\langle A, x\rangle = C$.

Если две точки x, y принадлежат плоскости, то разность векторов $x - y$ лежит в плоскости. В то же время,

$$\langle A, x - y\rangle = \langle A, x\rangle - \langle A, y\rangle = C - C = 0.$$

Таким образом, вектор A перпендикулярен любому вектору, лежащему в гиперплоскости L. По этой причине он называется нормалью к L.

Рассмотрим следующую задачу: написать уравнение плоскости, проходящей через заданную точку $x^0 = (x_1^0, x_2^0, \ldots, x_n^0)$ и с заданной нормалью $A = (A_!, A_2, \ldots, A_n)$. Ответом является уравнение

$$A_1(x_1 - x_1^0) + A_2(x_2 - x_2^0) + \cdots + A_n(x_n - x_n^0) = 0.$$

Оно называется уравнением плоскости, проходящей через заданную точку.

Проверим правильность ответа. Если раскрыть скобки, то уравнение преобразуется к виду

$$A_1x_1 + A_2x_2 + \cdots + A_nx_n = A_1x_1^0 + A_2x_2^0 + \cdots + A_nx_n^0.$$

Правая часть не содержит переменных x_i и есть некоторое число, которое можно обозначить C. Получится уравнение гиперплоскости в стандартной форме. Следовательно, уравнение задает гиперплоскость, и эта гиперплоскость проходит через заданную точку x^0, что и требовалось доказать.

Расстоянием от точки x до произвольного множества L в R^n называется наименьшее из расстояний между x и точками y из L. Если L — гиперплоскость, то расстояние равно длине перпендикуляра, опущенного из точки x на L.

Гладкой поверхностью в R^n называется множество, задаваемое некоторым уравнением

$$f(x_1, x_2, \ldots, x_n) = 0,$$

в котором функция f непрерывно дифференцируема. Наример, уравнение $x_1^2 + x_2^2 + x_3^2 = r^2$ задает в трехмерном пространстве сферу радиуса r. Это гладкая поверхность в R^3.

Определение *касательной плоскости к поверхности.*

Пусть G — гладкая поверхность в R^n, $z \in G$ и гиперлоскость L проходит через точку z. Для каждой точки поверхности $g \in G$ обозначим расстояние от g до плоскости L через $\rho(g, L)$, а расстояние от g до точки z — через $\rho(g, z) = \|g - z\|$.

Гиперплоскость L называется *касательной плоскостью к поверхности* G, если

$$\lim_{g\to z,\, g\in G} \frac{\rho(g, L)}{\rho(g, z)} = 0.$$

Касательная плоскость ближе всего примыкает к гладкой поверхности в окрестности точки касания z, так что расстояние от точки поверхности g до этой гиперплоскости намного меньше, чем расстояние до точки z. Другими словами, угол между прямой, соединяющей точки g и z, и гиперплоскостью L стремится к нулю, когда g приближается к z, оставаясь на поверхности G.

Теорема *о касательной плоскости к графику дифференцируемой функции.*

Пусть функция f дифференцируема в точке $x^0 = (x_1^0, x_2^0, \ldots, x_n^0)$, $y^0 = f(x^0)$ — значение f в точке x^0 и хотя бы одно из чисел $\frac{\partial f}{\partial x_1}(x^0)$, $\frac{\partial f}{\partial x_2}(x^0)$, …, $\frac{\partial f}{\partial x_n}$ отлично от нуля. Тогда график функции f в пространстве R^{n+1} имеет касательную плоскость в точке (x^0, y^0), которая задается уравнением

$$y - y^0 = \frac{\partial f}{\partial x_1}(x^0) \cdot (x_1 - x_1^0) + \frac{\partial f}{\partial x_2}(x^0) \cdot (x_2 - x_2^0) + \cdots + \frac{\partial f}{\partial x_n}(x^0) \cdot (x_n - x_n^0).$$

Доказательство.

По условию, функция f дифференцируема в точке x^0. Поэтому она имеет все частные производные в этой точке. Обозначим их $A_1 = \partial f/\partial x_1(x^0)$, …, $A_n = \partial f/\partial x_n(x^0)$. Уравнение

$$y - y^0 = A_1(x_1 - x_1^0) + A_2(x_2 - x_2^0) + \ldots + A_n(x_n - x_n^0)$$

относительно переменных $z = (x_1, x_2, \ldots, x_n, y)$ задает гиперплоскость в пространстве R^{n+1}. Обозначим ее L. Гиперплоскость L содержит точку $z^0 = (x^0, y^0)$, так как уравнение после подстановки компонент этой точки вместо x_i и y обращается в верное равенство $0 = 0$.

Докажем, что построенная гиперплоскость — касательная к графику функции f в точке z^0. Выберем произвольную точку $z = (x, y)$ на графике функции f. По определению графика, $y = f(x)$. Определим вспомогательную точку $z_L = (x, y_L)$ на гиперплоскости L с теми же компонентами вектора x и с компонентой

$$y_L = y^0 + A_1(x_1 - x_1^0) + A_2(x_2 - x_2^0) + \ldots + A_n(x_n - x_n^0).$$

Расстояние $\rho(z, L)$ от точки z до ближайшей точки на гиперплоскости L, очевидно, не больше, чем расстояние от z до z_L, так как точка z_L лежит в L. Поэтому

$$\rho(z, L) \leq \|z - z_L\| = |y - y_L|.$$

По формуле полного дифференциала

$$df(h) = A_1 h_1 + A_2 h_2 + \cdots + A_n h_n,$$

где $h_i = x_i - x_i^0$, $i = 1, 2, \ldots, n$.

Следовательно,

$$\frac{\rho(z, L)}{\|z - z^0\|} \leq \frac{|y - y_L|}{\|x - x^0\|} = \frac{|f(x) - f(x^0) - df(x - x^0)|}{\|x - x^0\|}.$$

Последняя величина стремится к нулю при $x \to x^0$ по определению дифференцируемой функции в точке x^0. Таким образом, $\rho(z, L)/\|z - z^0\| \to 0$ при $z \to z^0$. По определению, L — касательная плоскость к графику функции f. □

Градиент и линии уровня

Определение *множества уровня.*

Пусть функция f задана на множестве $D \subseteq \mathsf{R}^n$ и C — некоторое число. *Множеством уровня C функции f* называется прообраз $\{C\}$, т.е. множество

$$U_C = \{x \in D \mid f(x) = C\}.$$

$\square$

Если функция f дифференцируема, то ее множества уровня обычно являются гладкими поверхностями в R^n. В частности, если $n = 3$, то множество уровня обычно является поверхностью в трехмерном пространстве. Тогда его называют поверхностью уровня. При $n = 2$ множество уровня обычно есть линия на плоскости, и тогда его называют линией уровня.

Линии уровня, определяемые разными числами C, не пересекаются. В совокупности они составляют все множество D, т.е. множество определения функции f. Таким образом, это множество оказывается похожим на "слоеный пирог", в котором на каждом слое значение функции f постоянно.

Определение *градиента.*

Пусть функция f дифференцируема в точке $x = (x_1, x_2, \ldots, x_n)$. Тогда вектор

$$\operatorname{grad} f(x) = \left(\frac{\partial f}{\partial x_1}(x),\ \frac{\partial f}{\partial x_2}(x),\ \ldots,\ \frac{\partial f}{\partial x_n}(x)\right)$$

называется *градиентом функции f в точке x*. $\square$

Для дифференцируемой функции в каждой точке области определения D существует вектор градиента. Этот вектор может меняться по направлению и величине от точке к точке. Получается поле векторов, заданное на множестве D.

Градиенты часто применяются не только в теоретических исследованиях, но и в вычислительных и расчетных работах. Значительную часть современных методов вычислений составляют так называемые градиентные методы. Они основаны на важном свойстве градиента: он перпендикулярен множествам уровня и он указывает направление наискорейшего возрастания функции. Эти свойства обсуждаются в этом и следующем разделах.

Теорема. *Градиент и линии уровня.*

Пусть функция f задана в области $D \subseteq \mathsf{R}^n$ и $x^0 \in D$. Множество уровня $C = f(x^0)$ функции f обозначим G. Пусть функция f дифференцируема в точке x^0 и $\operatorname{grad} f(x^0) \neq 0$.

Тогда G имеет касательную плоскость в точке x^0 и нормалью к этой гиперплоскости является вектор $\operatorname{grad} f(x^0)$.

Доказательство.

Пусть G — множество уровня функции f, проходящее через x^0. Тогда $f(x) = f(x^0)$ при всех $x \in G$. Введем обозначения для компонент градиента функции f:

$$\operatorname{grad} f(x^0) = \left(\frac{\partial f}{\partial x_1}(x^0), \frac{\partial f}{\partial x_2}(x^0), \ldots, \frac{\partial f}{\partial x_n}(x^0)\right) = (A_1, A_2, \ldots, A_n) = A.$$

Тогда уравнение гиперплоскости L, проходящей через точку x^0 и с нормалью $\operatorname{grad} f(x^0)$ можно записать в виде

$$A_1(x_1 - x_1^0) + A_2(x_2 - x_2^0) + \cdots + A_n(x_n - x_n^0) = 0,$$

или в терминах скалярного произведения $\langle A, x - x^0 \rangle = 0$, если x — точка на гиперплоскости.

Если же точка x не принадлежит гиперплоскости, то левая часть уравнения, с точностью до множителя $\pm\|A\|$, равна расстоянию от $\rho(x, L)$ точки x до плоскости L:

$$\rho(x, L) = \frac{1}{\|A\|}|A_1(x_1 - x_1^0) + A_2(x_2 - x_2^0) + \cdots + A_n(x_n - x_n^0)| = \frac{1}{\|A\|}|\langle A, x - x^0 \rangle|.$$

Действительно, перпендикуляр, опущенный из точки x на плоскость L, параллелен нормали A к гиперплоскости. Следовательно, произведение их длин равно абсолютной величине их скалярного произведения. И оно же равно скалярному произведению нормали A на произвольный вектор с концом в точке x и началом на гиперплоскости L, например, на вектор $x - x^0$.

Докажем, что гиперплоскость L является касательной плоскостью к G в точке x^0.

Пусть $x = (x_1, \ldots, x_n) \in G$ — точка на множестве уровня $f(x) = C$. Запишем условие дифференцируемости функции f в точке x^0:

$$f(x) = f(x^0) + \langle \operatorname{grad} f(x^0), x - x^0 \rangle + \|x - x^0\|\alpha(x),$$

где $\alpha(x) \to 0$ при $x \to x^0$. Но $f(x) = f(x^0) = C$, следовательно,

$$\langle \operatorname{grad} f(x^0), x - x^0 \rangle = -\|x - x^0\|\alpha(x).$$

Запишем отношение двух расстояний из определения касательной плоскости:

$$\frac{\rho(x, L)}{\rho(x, x^0)} = \frac{1}{\|A\|} \frac{|\langle A, x - x^0 \rangle|}{\|x - x^0\|} = \frac{1}{\|A\|}|\alpha(x)| \to 0$$

при $x \to x^0$. Это условие означает, что L — касательная плоскость по ее определению. $\square$

Градиентные методы численной оптимизации

Предположим, что требуется максимизировать функцию f, т.е. найти такой аргумент x, при котором значение $f(x)$ максимально. Обычный вычислительный подход к этой задаче состоит в следующем: вычислим значение f в какой-нибудь точке x^0, а затем сдвинем эту точку x^0 в том направлении, где f быстрее всего возрастает. Получим следующее приближение x^1, для которого направление возрастания f можно подкорректировать. Снова сдвинем аргумент в направлении возрастания функции f и т.д. Получается итеративный процесс. Направление возрастания функции вблизи точки, как следует из следующего утверждения, есть градиент.

Теорема *о направлении наибольшего возрастания функции.*

Пусть функция f задана в области D, $x^0 \in D$ и функция f дифференцируема в точке x^0. Окружим точку x^0 шаром радиуса δ: $B_\delta = \{x \in D \mid \|x - x^0\| \le \delta\}$. В каждом таком шаре выберем точку x_δ, в которой значение f наибольшее:

$$f(x_\delta) = \max_{x \in B_\delta} f(x).$$

Приращение обозначим $h_\delta = x_\delta - x^0$.

Пусть в точке x^0 не все частные производные равны нулю: $\operatorname{grad} f(x^0) \ne 0$. Тогда при $\delta \to 0$ вектор h_δ стремится к вектору $\operatorname{grad} f(x^0)$ по направлению (т.е. угол между этими векторами стремится к нулю). □

Предварительно докажем вспомогательное утверждение.

Лемма.

Пусть v — вектор в R^n и $\mathsf{D}_1 = \{x \in \mathsf{R}^n : \|x\| \le 1\}$ — единичный шар в R^n. Тогда

$$\max_{x \in \mathsf{D}_1} \langle x, v\rangle = \|v\|,$$

причем максимум достигается на векторе x, параллельном v. (Здесь $\langle\cdot,\cdot\rangle$ обозначает скалярное произведение векторов.)

Доказательство.

Скалярное произведение векторов равно произведению их длин на косинус угла между ними. Длина вектора равна его евклидовой норме. Поэтому

$$\langle x, v\rangle = \|x\| \, \|v\| \cos\phi,$$

где ϕ — угол между x и v. Наибольшее значение этого произведения достигается при $\|x\| = 1$ и $\phi = 0$, т.е. когда v и x параллельны. □

Доказательство теоремы о направлении наибольшего возрастания функции.

По определению дифференцируемой функции,

$$f(x^0 + h) = f(x^0) + \langle \operatorname{grad} f(x^0), h\rangle + \|h\|\alpha(h),$$

где $\alpha(h) \to 0$ при $h \to 0$. Условие $x = x^0 + h \in B_\delta$ равносильно условию $\|h\| \le \delta$. Поэтому оптимальный вектор h_δ максимизирует следуюшее выражение:

$$f(x^0 + h_\delta) = f(x^0) + \max_{\|h\| \le \delta} (\langle \operatorname{grad} f(x^0), h\rangle + \|h\|\alpha(h)).$$

Вектор $e = h/\|h\|$ имеет единичную длину: $\|e\| = 1$. Оптимальное значение $e_\delta = h_\delta/\|h_\delta\|$ определяется условием

$$\max_{\|e\| \le 1} (\langle \operatorname{grad} f(x^0), e\rangle + \alpha(\delta e)).$$

По лемме в задаче

$$\max_{\|e\| \le 1} \langle \operatorname{grad} f(x^0), e\rangle$$

оптимальный вектор e_0 параллелен вектору $\operatorname{grad} f(x^0)$ и имеет единичную длину. Дополнительное слагаемое $\alpha(\delta e)$ стремится к нулю при $\delta \to 0$. Поэтому вектор e_δ

стремится к вектору e_0. Но направление e_δ совпадает с направлением h_δ, что доказывает утверждение теоремы. □

Метод градиентного спуска

Теорема о направлении наибольшего возрастания дифференцируемой функции подсказывает итеративный метод нахождения экстремумов. Рассмотрим задачу минимизации некоторой функции f.

Направление быстрейшего убывания функции противоположно направлению быстрейшего возрастания, и определяется вектором $-\operatorname{grad} f(x)$. Пусть x_1 — некоторая точка в области определения функции f. Вычислим вектор частных производных $v_1 = \operatorname{grad} f(x_1)$. Теорема говорит о том, что при малом отклонении от вектора x_1 функция уменьшится сильнее, если направление этого отклонения будет равно $-v_1$. Поэтому определим $x_2 = x_1 - hv_1$, где h — некоторое небольшое число, называемое шагом алгоритма. По теореме, если h достаточно мало, то $f(x_2) < f(x_1)$. Таким образом, значение функции уменьшилось. Далее вычисляем градиент в новой точке x_2, смещаем вектор x_2 в направлении, противоположном этому градиенту, и получаем x_3 и т.д.

Описанный итеративный процесс называется градиентным спуском. Важнейшим параметром этого процесса является величина шага h. С одной стороны, теорема гарантирует уменьшение функции лишь "при достаточно малых h". С другой стороны, при малых h движение векторов x_k будет очень медленным и потребуется много вычислительных ресурсов для достижения, наконец, искомой точки минимума. Вообще говоря, правильный выбор шага h является искусством, но некоторые полезные рекомендации выработаны в современных разделах науки, называемой вычислительной математикой.

Итеративный процесс останавливается, если градиент в точке очередного приближения x_N равен нулю. Но именно этим свойством обладают точки экстремума.

Определение *ординарной и особой точек дифференцируемой функции.*

Точка x называется *ординарной* точкой функции f, если градиент f в этой точке не равен нулю, и *особой*, если равен нулю или не существует. □

Таким образом, искомая точка минимума или максимума — особая, а в ординарных точках всегда есть возможность применить процедуру градиентного спуска и уменьшить значение функции.

Неявная функция

Уравнение $y = \phi(x)$ явно задает способ определения значения функции y по заданному значению аргумента x. Однако, правило, по которому каждому значению x ставится в соответствие некоторое значение y может быть сформулировано более сложно.

Определение. *Функция, заданная неявно.*

Пусть $f(x, y)$ — отображение из области $D \subseteq \mathsf{R}^{n+m}$ в пространство R^m, причем $x{=}(x_1, x_2, \dots, x_n)$, $y = (y_1, y_2, \dots, y_m)$.

Говорят, что уравнение $f(x, y) = 0$ задает *неявно* функцию (или задает *неявную функцию*) $y = \phi(x)$ в области $G \subseteq \mathsf{R}^n$, если для любого вектора $x \in G$ это уравнение однозначно разрешимо относительно чисел y_1, y_2, …, y_m. Получающийся вектор $y = (y_1, y_2, \ldots, y_m)$ является значением функции ϕ для аргумента x. $\square$

Рассмотрим в качестве примера простейший случай линейной функции f и $m = 1$. Пусть
$$f(x_1, x_2, \ldots, x_n, y) = a_1x_1 + a_2x_2 + \ldots + a_nx_n + b\,y.$$
Уравнение рассматривается во всем пространстве, то есть $D = \mathsf{R}^{n+1}$ и $G = \mathsf{R}^n$. Введем вектор
$x = (x_1, x_2, \ldots, x_n)$ и будем писать $f(x, y)$ вместо $f(x_1, x_2, \ldots, x_n, y)$.

Очевидно, что уравнение $f(x, y) = 0$ легко разрешимо относительно y при $b \neq 0$:
$$y = -\frac{1}{b}(a_1x_1 + a_2x_2 + \ldots + a_nx_n).$$
Поэтому уравнение $f(x, y) = 0$ задает неявную функцию $y = \phi(x)$, которая в данном случае может быть также явно определена последней формулой. К сожалению, для других уравнений часто бывает так, что решение y уравнения $f(x, y) = 0$ существует и единственно, но его невозможно задать формулой, содержащей элементарные функции. По этой причине понятие неявной функции и было введено.

Заметим, что в данном примере число b есть частная производная функции f по y: $b = \partial f/\partial y$. Если $b = 0$, то уравнение уже неразрешимо относительно y при произвольном x, и поэтому неявной функции нет. Оказывается, что условие $\partial f/\partial y \neq 0$ достаточно для существования неявной функции и в общем случае.

Теорема *о неявной функции.*

Пусть функция $f(x_1, \ldots, x_n, y)$ задана и непрерывно дифференцируема в области $D \subseteq \mathsf{R}^{n+1}$ (т.е. $f \in C^1(D)$). Рассмотрим точку $(\tilde{x}, \tilde{y}) \in D$, где $\tilde{x} = (\tilde{x}_1, \ldots, \tilde{x}_n)$. Пусть $f(\tilde{x}, \tilde{y}) = 0$ и $\partial f/\partial y(\tilde{x}, \tilde{y}) \neq 0$.

Тогда существуют такая окрестность B точки $\tilde{x}$ в R^n и такой интервал (c, d) вокруг точки $\tilde{y}$ ($\tilde{x} \in B$, $\tilde{y} \in (c, d)$), что для любого $x \in B$ уравнение $f(x, y) = 0$ однозначно разрешимо относительно $y \in (c, d)$. Кроме того, получающаяся функция $y(x)$ из B в R^1 непрерывно дифференцируема по x.

Доказательство.

Пусть для определенности, $\partial f/\partial y(\tilde{x}, \tilde{y}) > 0$. Функция $\partial f/\partial y$ непрерывна на D. Поэтому в некоторой окрестности $D_1 \subseteq D$ точки $(\tilde{x}, \tilde{y})$ эта функция сохраняет знак: $\partial f/\partial y(x, y) > 0$ при всех $(x, y) \in D_1$.

Поскольку множество D_1 есть окрестность точки $(\tilde{x}, \tilde{y})$, то существует такая окрестность B_1 точки $\tilde{x}$ и такой интервал (c_1, d_1), содержащий $\tilde{y}$, что множество $B_1 \times (c, d)$ содержится в D.

Пусть x — произвольная точка в B_1. Рассмотрим функцию $\psi(y) = f(x, y)$. По определению частной производной, $\psi'(y) = \partial f/\partial y(x, y) > 0$. Таким образом, $\psi'(y) > 0$ при всех $y \in (c, d)$. Следовательно, функция ψ монотонно возрастает на (c, d). Таким образом, доказано, что для любого $x \in B_1$ и любого $y \in (c, d)$
$$f(x, c) < f(x, y) < f(x, d).$$
Но $f(\tilde{x}, \tilde{y}) = 0$ по условию. Следовательно, $f(\tilde{x}, c) < 0$ и $f(\tilde{x}, d) > 0$.

По непрерывности функции f существует такая окрестность B точки $\tilde{x}$, что условия $f(x, c) < 0$, $f(x, d) > 0$ выполняются при всех $x \in B$. Докажем, что найденный интервал (c, d) и окрестность B точки $\tilde{x}$ удовлетворяют утверждению теоремы.

Пусть $x \in B$. Тогда функция $\psi(y) = f(x, y)$ удовлетворяет условиям:

1. $\psi(c) < 0$, $\psi(d) > 0$,
2. функция ψ монотонно возрастает на (c, d).

Следовательно, функция $\psi(y)$ имеет на (c, d) единственный корень y. Это число y зависит от x, и получающаяся функция $y = y(x)$ удовлетворяет равенству $f(x, y(x)) = 0$. Следовательно, неявная функция существует. Докажем ее непрерывность.

Пусть $x_0 \in B$ и $\epsilon > 0$. Требуется найти такое $\delta > 0$, что при всех x из δ–окрестности точки x_0, т.е. при $|x - x_0| < \delta$, выполнено

$$y(x_0) - \epsilon < y(x) < y(x_0) + \epsilon.$$

Рассмотрим функции

$$\begin{aligned} \eta_+(x) &= f(x, y(x_0) + \epsilon), \\ \eta_-(x) &= f(x, y(x_0) - \epsilon). \end{aligned}$$

Для них

$$\begin{aligned} \eta_+(x_0) &= f(x_0, y(x_0) + \epsilon) > f(x_0, y(x_0)) = 0. \\ \eta_-(x_0) &= f(x_0, y(x_0) - \epsilon) < f(x_0, y(x_0)) = 0. \end{aligned}$$

Эти функции непрерывны по x, следовательно,

$$\exists \delta > 0 \,:\, |x - x_0| < \delta \Rightarrow \eta_+(x) > 0,\ \eta_-(x) < 0.$$

Таким образом, если $|x - x_0| < \delta$, то

$$\begin{aligned} f(x, y(x_0) - \epsilon) &< 0, \\ f(x, y(x_0) + \epsilon) &> 0. \end{aligned}$$

Следовательно, между числами $y(x_0)-\epsilon$ и $y(x_0)+\epsilon$ есть такое число y, что $f(x, y) = 0$. Но это и значит, что $y(x) \in (y(x_0) - \epsilon,\ y(x_0) + \epsilon)$, что совпадает с определением непрерывности функции $y(x)$ в точке x_0.

Докажем, что функция $y(x)$ дифференцируема на B и все частные производные непрерывны. Поскольку при вычислении частных производных нужно в векторе $x = (x_1, \ldots, x_n)$ фиксировать все компоненты, кроме одной, то достаточно рассмотреть случай, когда $n = 1$.

Пусть $n = 1$. Вычислим производную функции $y(x)$ как предел. Пусть Δx — приращение аргумента, $\Delta y = y(x+\Delta x) - y(x)$ — приращение функции. Для краткости будем писать y вместо $y(x)$. Тогда $f(x, y) = 0$ и $f(x+\Delta x, y+\Delta y) = 0$ по определению функции $y(x)$. Проведем вычисления:

$$\begin{aligned} 0 &= f(x + \Delta x, y + \Delta y) - f(x, y) \\ &= (f(x + \Delta x, y + \Delta y) - f(x, y + \Delta y)) + (f(x, y + \Delta y) - f(x, y)) \\ &= \frac{\partial f}{\partial x}(x + \vartheta_1 \Delta x, y + \Delta y)\Delta x + \frac{\partial f}{\partial y}(x, y + \vartheta_2 \Delta y)\Delta y. \end{aligned}$$

В последнем равенстве дважды применена теорема Лагранжа о конечных приращениях. Из этой теоремы ϑ_1 и ϑ_2 — некоторые числа из промежутка $(0,1)$. Таким образом,

$$\frac{\Delta y}{\Delta x} = -\frac{\frac{\partial f}{\partial x}(x+\vartheta_1\Delta x, y+\Delta y)}{\frac{\partial f}{\partial y}(x, y+\vartheta_2\Delta y)}.$$

Пусть $\Delta x \to 0$. По доказанной непрерывности функции $\Delta y \to 0$. Функции $\frac{\partial f}{\partial x}$ и $\frac{\partial f}{\partial y}$ непрерывны. Следовательно, предел существует и

$$y'(x) = \lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x} = -\frac{\frac{\partial f}{\partial x}(x, y(x))}{\frac{\partial f}{\partial y}(x, y(x))}.$$

Полученная функция правильно определена, так как знаменатель отличен от нуля, и непрерывна как композиция непрерывных функций. Таким образом, функция $y'(x)$ непрерывно дифференцируема на B, что завершает доказательство теоремы. $\square$

Следствие.

В доказательстве теоремы была получена следующая формула для производной неявной функции:

$$y'(x) = -\frac{\frac{\partial f}{\partial x}(x, y(x))}{\frac{\partial f}{\partial y}(x, y(x))}.$$

Неявные функции, определяемые системой уравнений.

Определение. *Вектор–функция, заданная неявно.*

Рассмотрим систему уравнений:

$$\begin{cases} f_1(x_1,\ldots,x_n,y_1,\ldots,y_m) &= 0,\\ f_2(x_1,\ldots,x_n,y_1,\ldots,y_m) &= 0,\\ \qquad\ldots & \\ f_m(x_1,\ldots,x_n,y_1,\ldots,y_m) &= 0.\end{cases}$$

Здесь функции $f_1,\ldots,f_m$ непрерывно дифференцируемы в области $D\subseteq \mathsf{R}^n\times\mathsf{R}^m$, а количество переменных y_j равно количеству уравнений. Если при всех $x=(x_1,\ldots,x_n)$ из некоторого множества B эта система однозначно разрешима относительно $y = (y_1,\ldots,y_m)$, то говорят, что получающаяся функция $y=y(x)$ задана *неявно* этой системой. $\square$

Рассмотрим частный случай. Пусть функции f_i линейны по y_j, т.е.

$$f_i(x_1,\ldots,x_n,y_1,\ldots,y_m) = g_i(x_1,\ldots,x_n) + a_{i1}y_1+\ldots+a_{im}y_m.$$

Тогда систему уравнений можно записать в следующем виде:

$$\begin{cases} a_{11}y_1+\ldots+a_{1m}y_m &= -g_1(x_1,\ldots,x_n),\\ a_{21}y_1+\ldots+a_{2m}y_m &= -g_2(x_1,\ldots,x_n),\\ \qquad\ldots &= \qquad\ldots\\ a_{m1}y_1+\ldots+a_{mm}y_m &= -g_m(x_1,\ldots,x_n).\end{cases}$$

При фиксированных x_k это система линейных уравнений относительно неизвестных $y_1,\ldots,y_m$. Количество уравнений равно количеству неизвестных. Следовательно, по теореме о решении систем линейных уравнений решение $y(x)$ всегда существует, если определитель матрицы коэффициентов отличен от нуля. Таким образом, достаточным условием существования неявной функции является неравенство

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mm} \end{vmatrix} \neq 0.$$

С другой стороны, нетрудно видеть, что частные производные функции f_i по y_j равны коэффициентам этой матрицы:

$$a_{ij} = \frac{\partial f_i}{\partial y_j},$$

и они не зависят ни от x, ни от y.

Определение *якобиана системы функций.*

Пусть функции $f_1, \dots, f_m$ зависят от переменных $x_1, \dots, x_n$ и $y_1, \dots, y_m$ и каждая функция f_i имеет частную производную по переменной y_j, $1 \le i, j \le m$, в некоторой точке (x, y), где $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_m)$. Тогда определитель

$$\begin{vmatrix} \frac{\partial f_1}{\partial y_1}(x,y) & \frac{\partial f_1}{\partial y_2}(x,y) & \dots & \frac{\partial f_1}{\partial y_m}(x,y) \\ \frac{\partial f_2}{\partial y_1}(x,y) & \frac{\partial f_2}{\partial y_2}(x,y) & \dots & \frac{\partial f_2}{\partial y_m}(x,y) \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_m}{\partial y_1}(x,y) & \frac{\partial f_m}{\partial y_2}(x,y) & \dots & \frac{\partial f_m}{\partial y_m}(x,y) \end{vmatrix}$$

называется *якобианом* функций $f_1, \dots, f_m$ по переменным $y_1, \dots, y_m$ в точке (x, y). Якобиан обозначается

$$\frac{\partial(f_1, \dots, f_m)}{\partial(y_1, \dots, y_m)}(x, y).$$

□

Выше доказано, что для линейных по переменным y_j функций f_i если якобиан отличен от нуля, то неявная функция существует. Это же верно и для общего случая. Действительно, в окрестности любой точки $(\tilde{x}, \tilde{y})$ приращение функции $f_i(x, y)$ приближенно равно дифференциалу, а дифференциал — линейная функция, для которой утверждение доказано.

Теорема *о неявной функции для систем уравнений*

Пусть функции $f_i(x_1, \dots, x_n, y_1, \dots, y_m)$, $1 \le i \le m$, заданы и непрерывно дифференцируемы в области $D \subseteq \mathsf{R}^{n+m}$ (т.е. $f \in C^1(D)$). Рассмотрим точку $(\tilde{x}, \tilde{y}) \in D$, где $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_n)$, $\tilde{y} = (\tilde{y}_1, \dots, \tilde{y}_m)$. Пусть $f(\tilde{x}, \tilde{y}) = 0$ и якобиан

$$\frac{\partial(f_1, \dots, f_m)}{\partial(y_1, \dots, y_m)}(\tilde{x}, \tilde{y}) \neq 0.$$

Тогда существуют такая окрестность B точки $\tilde{x}$ в R^n и такая окрестность C точки $\tilde{y}$ ($\tilde{x} \in B$, $\tilde{y} \in C$), что для любого $x \in B$ уравнение $f(x, y) = 0$ однозначно разрешимо относительно $y = (y_1, \dots, y_m) \in C$. Кроме того, получающиеся функции $y_j(x)$, $1 \le$

$j \le m$, из B в R^1 непрерывно дифференцируемы по x и частные производные имеют вид

$$\frac{\partial y_j}{\partial x_k} = -\frac{\frac{\partial(f_1,\dots,f_m)}{\partial(y_1,\dots,y_{j-1},x_k,y_{j+1},\dots,y_m)}}{\frac{\partial(f_1,\dots,f_m)}{\partial(y_1,\dots,y_m)}}.$$

(Без доказательства.)

Частные производные высших порядков

Определение *частных производных высших порядков.*

Пусть D — область в R^n и $f : D \to \mathsf{R}$. Пусть функция f непрерывно дифференцируема в области D. Функции $g_1(x) = \partial f/\partial x_1, \dots, g_n(x) = \partial f/\partial x_n$ заданы в области D. Они могут быть также дифференцируемы в D. Их частные производные (если существуют) называются частными производными функции f второго порядка. По индукции определяются частные производные функции f любого целого положительного порядка. □

О терминологии и обозначениях. Частная производная второго порядка по x_i от функции $\partial f/\partial x_j$ есть по определению $\partial(\partial f/\partial x_j)/\partial x_i$. Можно доказать, что если эти частные производные существуют и непрерывны в некоторой области, то порядок дифференцирования несущественен, т.е.

$$\frac{\partial\left(\frac{\partial f}{\partial x_j}\right)}{\partial x_i} = \frac{\partial\left(\frac{\partial f}{\partial x_i}\right)}{\partial x_j}.$$

Поскольку порядок ∂x_i и ∂x_i несущественен, то эта перекрестная производная обозначается одним из двух способов:

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

Теорема *о перестановке частных производных.*

Частные производные высших порядков не зависят от последовательности дифференцирований, а зависят только от того, по каким переменным и сколько раз по каждой из них производилось дифференцирование, при условии, что все встречающиеся частные производные непрерывны в рассматриваемой области. □

(Без доказательства.)

Если $i \ne j$, то такая частная производная второго порядка называется смешанной или перекрестной, а если $i = j$, то чистой. В последнем случае она обозначается

$$\frac{\partial^2 f}{\partial x_i^2}(x_1, \dots, x_n).$$

Формула Тейлора для двух переменных

Пусть $n = 2$, и функция $f(x, y)$ задана в некоторой области в R^2. Любая ее частная производная (если существует) состоит в дифференцировании несколько раз по x и несколько раз по y, т.е. имеет вид

$$\frac{\partial^m f}{\partial x^i \partial y^j} = f^{(m)}_{x^i y^j}, \qquad m = i + j.$$

Пусть $(x, y) \in D$. Выберем направление изменения аргумента (h, k) и предположим, что весь отрезок $(x + th, y + tk)$ при $0 \le t \le 1$ находится в D. Построим формулу Тейлора для приращения функции

$$\Delta f = f(x + h, y + k) - f(x, y).$$

Для этого определим функцию одного аргумента t:

$$\zeta(t) = f(x + th, y + tk).$$

Для этой функции можно вывести формулу Тейлора. Сначала вычислим производные функции ζ через частные производные функции f. Формула для производной сложной функции дает

$$\begin{aligned} \zeta'(t) &= f'_x h + f'_y k, \\ \zeta''(t) &= f''_{x^2} h^2 + 2 f''_{xy} hk + f''_{y^2} k^2. \end{aligned}$$

Здесь для краткости аргумент $(x+th, y+tk)$ всех функций опускается. По индукции можно доказать, что

$$\zeta^{(m)}(t) = \sum_{i=0}^{m} C^i_m h^i k^{m-i} \frac{\partial^m f}{\partial x^i \partial y^{m-i}},$$

где C^i_m — число сочетаний из m по i. Полученное выражение напоминает бином Ньютона. По правилу раскрытия бинома Ньютона оно совпадает со следующим выражением:

$$\zeta^{(m)}(t) = \left(h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y} \right)^m f(x + th, y + tk).$$

Теперь можно непосредственно применить формулу Тейлора в форме Лагранжа.

Теорема. *Формула Тейлора в форме Лагранжа.*

Пусть функция f задана в области D и имеет там $m + 1$ непрерывную частную производную. Пусть отрезок между точками (x, y) и $(x + h, y + k)$ лежит в D. Тогда существует такое число $\vartheta \in (0, 1)$, что

$$\begin{aligned} f(x + h, y + k) &= f(x, y) + \left(h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y} \right) f(x, y) + \frac{1}{2} \left(h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y} \right)^2 f(x, y) + \ldots \\ &+ \frac{1}{m!} \left(h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y} \right)^m f(x, y) + \\ &+ \frac{1}{(m+1)!} \left(h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y} \right)^{m+1} f(x + \vartheta h, y + \vartheta k). \end{aligned}$$

Доказательство.

Утверждение теоремы следует из приведенной выше формулы вычисления производных функции $\zeta(t)$ и из формулы Тейлора для функции одной переменной $\zeta(t)$ с оценкой остаточного члена по Лагранжу. □

В частности, при $m = 1$ получаем

$$f(x+h, y+k) = f(x, y) + f'_x(x, y)h + f'_y(x, y)k +$$
$$+\frac{1}{2}[f''_{x^2}(x+\vartheta h, y+\vartheta k)h^2 + 2f''_{xy}(x+\vartheta h, y+\vartheta k)hk + f''_{y^2}(x+\vartheta h, y+\vartheta k)k^2].$$

Экстремумы функций нескольких переменных

Пусть D — область в R^n и $f : D \to \mathsf{R}$ — функция n аргументов. Пусть $x^0 = (x_1^0, \dots, x_n^0)$ — некоторая точка в D.

Определение *локальных и глобальных максимумов и минимумов.*

1. Точка x^0 называется *точкой максимума* (или *глобального максимума*) *функции f в D*, если в этой точке f принимает наибольшее в D значение, т.е.

$$f(x) \leq f(x^0) \qquad \forall x \in D.$$

Аналогично определяется *точка минимума* (или *глобального минимума*) *функции f в D*.

2. Точка x^0 называется *точкой локального максимума функции f в D*, если существует окрестность $\mathcal{O}$ точки x^0 в D, в которой точка x^0 является точкой глобального максимума f, т.е.

$$f(x) \leq f(x^0) \qquad \forall x \in \mathcal{O}.$$

Аналогично определяется *точка локального минимума функции f в D.* □

О терминологии. Точки максимума и точки минимума функции f называются точками экстремума, а значения f в них — экстремумами функции f.

Соответственно, точки локального минимума и локального максимума называются точками локального экстремума, а значения в них — локальными экстремумами.

В литературе по математическим методам в разных науках, в том числе в экономике, часто термин "экстремум" употребляют вместо термина "локальный экстремум". При этом максимум и минимум f называют наибольшим и наименьшим значениями, соответственно.

Пример. На карте горной местности локальными максимумами функции высоты над уровнем моря являются все горные вершины. Но только наивысшая вершина есть точка максимума.

Замечания.

1. Локальные и глобальные максимумы функции f в области D могут не существовать.

2. Если же x^0 — точка локального экстремума, то это свойство функции f в точках, близких к x^0. Оно никак не связано со значениями функции f за пределами любой окрестности точки x^0, в частности, от изменения области определения функции f (сужение или продолжение).

Теорема. *Необходимое условие экстремума.*

Пусть x^0 — точка локального экстремума функции f в области D и функция f

дифференцируема в x^0. Тогда все частные производные в этой точке равны нулю:

$$\frac{\partial f}{\partial x_i}(x^0) = 0, \qquad 1 \le i \le n.$$

Доказательство.

По определению частной производной $\partial f/\partial x_i$, это производная функции от x_i при фиксированных всех остальных переменных x_k, равных x_k^0, $1 \le k \le n$, $k \ne i$. Но такая функция $\kappa(x_i)$ одной переменной также имеет экстремум в точке x_i^0. По теореме Ферма производная в этой точке равна нулю. Теорема применима, так как точка x_i^0 оказывается внутренней точкой промежутка определения функции κ. □

Определение *точек стационарности.*

Точки в области определения дифференцируемой функции f, в которых все частные производные равны нулю, называются *точками стационарности* (а также *критическими точками*). □

Так же, как и для функции одной переменной, нахождение всех экстремумов дифференцируемой функции в области состоит из следующих этапов:

1. находим все n частных производных f и приравниваем их к нулю;
2. решаем полученную систему из n уравнений с n неизвестными и находим все точки стационарности;
3. для каждой точки стационарности индивидуально решаем вопрос о наличии в ней экстремума и какого именно.

Теорема. *Достаточное условие экстремума функции двух переменных.*

Пусть функция $f(x, y)$ имеет вторые непрерывные производные в области определения $D \subseteq \mathsf{R}^2$ (т.е. $f \in C^2(D)$). Пусть (x_0, y_0) — точка стационарности функции f, т.е.

$$\frac{\partial f}{\partial x}(x_0, y_0) = 0, \qquad \frac{\partial f}{\partial y}(x_0, y_0) = 0.$$

Тогда если в точке (x_0, y_0)

$$\frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 > 0,$$

то это точка экстремума. Если при этом

$$\frac{\partial^2 f}{\partial x^2}(x_0, y_0) > 0,$$

то это точка минимума, а если

$$\frac{\partial^2 f}{\partial x^2}(x_0, y_0) < 0,$$

то точка максимума.

Если же в точке (x_0, y_0)

$$\frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 < 0,$$

то функция f не имеет экстремума в точке (x_0, y_0).

Доказательство.

Пусть (h, k) — приращения аргумента, $x = x_0 + h$, $y = y_0 + k$. По формуле Тейлора

$$\begin{aligned} f(x,y) &= f(x_0,y_0) + \frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}(x_0+\vartheta h, y_0+\vartheta k)h^2\right. \\ &\left. +2\frac{\partial^2 f}{\partial x \partial y}(x_0+\vartheta h, y_0+\vartheta k)hk + \frac{\partial^2 f}{\partial y^2}(x_0+\vartheta h, y_0+\vartheta k)k^2\right), \end{aligned}$$

где $\vartheta \in (0, 1)$.

Введем обозначения:

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2}(x_0+\vartheta h, y_0+\vartheta k) &= A_1, \\ \frac{\partial^2 f}{\partial y^2}(x_0+\vartheta h, y_0+\vartheta k) &= B_1, \\ \frac{\partial^2 f}{\partial x \partial y}(x_0+\vartheta h, y_0+\vartheta k) &= C_1. \end{aligned}$$

Тогда приращение функции

$$f(x,y) - f(x_0,y_0) = \begin{pmatrix} h \\ k \end{pmatrix}^T \begin{pmatrix} A_1 & B_1 \\ B_1 & C_1 \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix} = X^T G X$$

есть квадратичная форма от вектора приращений $X = (h, k)^T$ с указанной матрицей коэффициентов G. При фиксированных коэффициентах эта квадратичная форма принимает только положительные значения (при ненулевых h и k), если матрица G положительно определена. По критерию Сильвестра эта матрица положительно определена, если $A_1 > 0$ и определитель матрицы G положителен.

Таким образом, приращение $f(x, y) - f(x_0, y_0)$ положительно, а функция имеет локальный минимум в точке (x_0, y_0), если выполнены неравенства

$$\frac{\partial^2 f}{\partial x^2} > 0, \qquad \frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 > 0$$

в точке $(x_0 + \vartheta h, y_0 + \vartheta k)$. Но при малых приращениях k и h эти неравенства остаются выполненными, если они выполнены в точке (x_0, y_0) по условию непрерывности вторых производных. Следовательно, доказано утверждение теоремы о локальном минимуме.

Для доказательства утверждения о локальном максимуме воспользуемся снова критерием Сильвестра. Квадратичная форма $X^T G X$ принимает только отрицательные значения (при $X \neq 0$), если $X^T(-G)X > 0$, т.е. матрица $(-G)$ положительно определена. По критерию Сильвестра $-A_1 > 0$ и $det(-G) > 0$. Эти условия выполнены вблизи точки (x_0, y_0), если в точке (x_0, y_0)

$$\frac{\partial^2 f}{\partial x^2} < 0, \qquad \frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 > 0,$$

что совпадает с условием теоремы.

Наконец, квадратичная форма принимает разные по знаку значения, если ее матрица не является ни положительно определенной, ни отрицательно определенной. Это может быть только если $\det(G) \le 0$. В этом случае экстремума в точке (x_0, y_0) нет, а неравенство $\det(G) \le 0$ выполнено, если в точке (x_0, y_0)

$$\frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 < 0,$$

что завершает доказательство теоремы. □

Теорема дает наиболее распространенный алгоритм определения, является ли точка стационарности локальным экстремумом и каким. Достаточно вычислить первые и вторые производные и проверить выполнение неравенств. Однако, если в точке (x_0, y_0)

$$\frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 = 0,$$

то вопрос об экстремуме остается открытым.

Условный экстремум. Множители Лагранжа

Пусть функция f определена в области $D \subseteq \mathsf{R}^n$. Если в области аргументов D провести линию ℓ, то можно определить сужение f на ℓ, т.е. функцию $g : \ell \to \mathsf{R}$, удовлетворяющую равенству $g(x) = f(x)$ при всех $x \in \ell$.

Функция g может иметь в точке x^0 локальный экстремум, даже если функция f его не имела. Пример: $f(x_1, x_2) = x_1^2 - x_2^2$ в точке $x^0 = (0, 0)$. На оси абсцисс (при $x_2 = 0$) сужение функции f имеет минимум в точке x^0, а на оси ординат (при $x_1 = 0$) имеет максимум в точке x^0. Сама же функция f в этой точке не имеет локального экстремума.

Линия ℓ или другое подмножество D, на котором требуется найти экстремумы f, обычно задается множеством уровня некоторой функции ϕ или пересечением множеств уровня нескольких функций $\phi_i(x)$.

Определение *условного экстремума.*

Пусть функция f и функции ϕ_k, $1 \le k \le m$, заданы в области $D \subseteq \mathsf{R}^n$ и $x^0 \in D$. Точка x^0 называется *точкой условного экстремума f при условии*

$$\phi_1(x) = 0, \dots, \phi_m(x) = 0,$$

если сужение f на подмножество D, определяемое последними уравнениями, имеет экстремум в точке x^0. □

Говорят также, что переменные $x = (x_1, \dots, x_n)$ связаны зависимостями $\phi_k(x) = 0$, $1 \le k \le m$, и требуется найти экстремум f при учете этих зависимостей. Уравнения $\phi_k(x) = 0$, $1 \le k \le m$, называют также ограничениями или уравнениями связи.

Теорема. *Необходимое условие условного экстремума.*

Пусть функция f имеет в точке $x^0 \in D$ условный экстремум при условии $\phi_1(x) = 0$, ..., $\phi_m(x) = 0$. Пусть все указанные функции дифференцируемы в области D. Тогда существуют такие числа λ_0, λ_1, ..., λ_m, что для функции

$$F(x) = \lambda_0 f(x) + \lambda_1\phi_1(x) + \ldots + \lambda_m\phi_m(x)$$

точка x^0 является стационарной точкой, причем среди чисел λ_k есть ненулевые. □

Функция $F(x)$ задана с точностью до множителя. Обычно коэффициент при функции f ненулевой, и его выбирают равным 1.

Определение *функции Лагранжа в задаче условной оптимизации.*

Функция

$$L(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_m) = f(x) + \lambda_1\phi_1(x) + \ldots + \lambda_m\phi_m(x)$$

называется *функцией Лагранжа* в задаче на условный экстремум, а числа λ_1, ..., λ_m называются *множителями Лагранжа.* □

Доказательство теоремы для случая $n = 3$, $m = 2$.

Изменим обозначения, чтобы уменьшить количество индексов в формулах. Пусть f — функция трех аргументов, которые обозначим x, y, z, а вектор аргументов обозначим $X = (x, y, z)$. Пусть функция $f(x, y, z)$ имеет условный экстремум в точке $X^0{=}(x^0, y^0, z^0)$ при ограничениях $\phi(x, y, z){=}0$, $\psi(x, y, z){=}0$. Рассмотри два случая.

1. Пусть один из трех якобианов, вычисленных в точке X^0,

$$\frac{\partial(\phi,\psi)}{\partial(x,y)}(X^0), \qquad \frac{\partial(\phi,\psi)}{\partial(x,z)}(X^0), \qquad \frac{\partial(\phi,\psi)}{\partial(y,z)}(X^0),$$

не равен нулю. Переменные x, y, z равноправны, поэтому для определенности будем считать, что это третий якобиан. Тогда можно применить теорему о неявной функции, заданной ограничениями $\phi(x, y, z) = 0$ и $\psi(x, y, z) = 0$. По этой теореме в окрестности точки x^0 переменные y и z являются функциями от x, $y = y(x)$, $z = z(x)$. Функция

$$g(x) = f(x, y(x), z(x))$$

имеет локальный экстремум в точке x^0 по определению условного экстремума. По теореме Ферма в этой точке производная равна нулю. Применим формулу производной сложной функции:

$$\frac{dg}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx} + \frac{\partial f}{\partial z}\frac{dz}{dx} = 0$$

в точке x^0.

Функции $\phi(x, y(x), z(x))$ и $\psi(x, y(x), z(x))$ тождественно равны нулю по определению неявных функций. Следовательно, равны нулю и их производные по x в точке x^0:

$$\begin{aligned} \frac{\partial\phi}{\partial x} + \frac{\partial\phi}{\partial y}\frac{dy}{dx} + \frac{\partial\phi}{\partial z}\frac{dz}{dx} &= 0, \\ \frac{\partial\psi}{\partial x} + \frac{\partial\psi}{\partial y}\frac{dy}{dx} + \frac{\partial\psi}{\partial z}\frac{dz}{dx} &= 0. \end{aligned}$$

Исключим из полученных трех уравнений полные производные dy/dx и dz/dx. Для этого умножим уравнение для ϕ на λ, уравнение для ψ на μ и сложим с уравнением для f:

$$\frac{\partial f}{\partial x}+\lambda\frac{\partial\phi}{\partial x}+\mu\frac{\partial\psi}{\partial x}+\left(\frac{\partial f}{\partial y}+\lambda\frac{\partial\phi}{\partial y}+\mu\frac{\partial\psi}{\partial y}\right)\frac{dy}{dx}+\left(\frac{\partial f}{\partial z}+\lambda\frac{\partial\phi}{\partial z}+\mu\frac{\partial\psi}{\partial z}\right)\frac{dz}{dx}=0.$$

Выберем числа λ и μ так, чтобы выражения в последних двух скобках были равны нулю. Это можно сделать, так как определитель системы линейных уравнений относительно λ и μ как раз равен якобиану $\partial(\phi,\psi)/\partial(y,z)$, который отличен от нуля.

Выбрав числа λ и μ так, чтобы коэффициенты при dy/dx и dz/dx были равны нулю, получим

$$\frac{\partial f}{\partial x}+\lambda\frac{\partial\phi}{\partial x}+\mu\frac{\partial\psi}{\partial x}=0,$$

что и требовалось доказать.

2. Пусть все три якобиана равны нулю в точке $X^0=(x^0,y^0,z^0)$. Рассмотрим матрицу

$$A=\begin{pmatrix}\frac{\partial\phi}{\partial x}(X^0) & \frac{\partial\phi}{\partial y}(X^0) & \frac{\partial\phi}{\partial z}(X^0)\\ \frac{\partial\psi}{\partial x}(X^0) & \frac{\partial\psi}{\partial y}(X^0) & \frac{\partial\psi}{\partial z}(X^0)\end{pmatrix}$$

размерности 2×3. Миноры второго порядка этой матрицы как раз совпадают с якобианами, вычисленными в точке X^0, и поэтому они равны нулю. По определению ранга матрицы, ранг матрицы A меньше 2. Следовательно, ее строки линейно зависимы. По определению линейной зависимости существуют такие числа λ и μ, не равные одновременно нулю, что строка $(\lambda,\mu)A$ равна нулю. Эта строка состоит из трех элементов:

$$\begin{aligned}\lambda\frac{\partial\phi}{\partial x}(X^0)+\mu\frac{\partial\psi}{\partial x}(X^0)&=0,\\ \lambda\frac{\partial\phi}{\partial y}(X^0)+\mu\frac{\partial\psi}{\partial y}(X^0)&=0,\\ \lambda\frac{\partial\phi}{\partial z}(X^0)+\mu\frac{\partial\psi}{\partial z}(X^0)&=0.\end{aligned}$$

Отсюда следует, что X^0 является стационарной точкой функции

$$F(X)=\lambda\phi(X)+\mu\psi(x),$$

т.е. в обозначениях условия теоремы получается, что $\lambda_0=0$, $\lambda_1=\lambda$ и $\lambda_2=\mu$. □

Метод множителей Лагранжа в задаче нахождения условного экстремума

Пусть задана оптимизируемая функция $f(x_1,\ldots,x_n)$ и ограничения $\phi_k(x_1,\ldots,x_n)$=0, $1\le k\le m$. Требуется найти точки условного экстремума.

1. Составляется функция Лагранжа:

$$L(x_1,\ldots,x_n,\lambda_1,\ldots,\lambda_m)=f(x)+\lambda_1\phi_1(x)+\ldots+\lambda_m\phi_m(x).$$

2. Вычисляются частные производные $\partial L/\partial x$.

3. Определяются точки стационарности функции L по всем переменным. Для этого приравниваются к нулю все частные производные функции L. Отметим, что $\partial L/\partial\lambda_k = \phi_k$.

4. Решается система из $n+m$ уравнений

$$\begin{aligned} \frac{\partial L}{\partial x_i}(x_1,\ldots,x_n,\lambda_1,\ldots,\lambda_m) &= 0, \qquad 1 \le i \le n, \\ \phi_k(x_1,\ldots,x_n) &= 0, \qquad 1 \le k \le m, \end{aligned}$$

   с $n+m$ неизвестными $x_1,\ldots,x_n,\lambda_1,\ldots,\lambda_m$.

5. Для полученных точек стационарности отдельно решается вопрос о наличии условного экстремума.

Задача математического программирования

Пусть $x_1,\ldots,x_n \ge 0$ – неотрицательные переменные. Пусть $f(x_1,\ldots,x_n)$ – некоторая функция, подлежащая максимизации. Например, это доход фирмы, а x_i – затраченные ресурсы.

Пусть по смыслу задачи переменные x_i не могут быть произвольными, и удовлетворяют ограничениям $\phi_k(x_1,\ldots,x_n) \le 0$, $1 \le k \le m$. Требуется оптимизировать функцию при этих ограничениях.

Формально задача записывается следующим образом:

$$f(x_1,\ldots,x_n) \to \max$$

при условии

$$\begin{gathered} \phi_1(x_1,\ldots,x_n) \le 0, \\ \cdots \\ \phi_m(x_1,\ldots,x_n) \le 0, \\ x_1 \ge 0,\ldots,x_n \ge 0 \end{gathered}$$

Эта задача на условный экстремум обычно называется задачей математического программирования. Термин не является общепринятым, в математике это просто задача на условный экстремум. Аналогично ставится задача на условный минимум.

Набор аргументов (или план) $x = (x_1,\ldots,x_n)$ называется допустимым, если он удовлетворяет всем ограничениям. Набор аргументов $x^0 = (x_1^0,\ldots,x_n^0)$ называется оптимальным, если он допустим и на нем достигается условный глобальный максимум функции f. Далее будем предполагать, что множество допустимых планов имеет внутренние точки.

Теорема *о множителях Лагранжа для ограничений типа неравенство.*

Пусть функции f, $\phi_1,\ldots,\phi_m$ – непрерывно дифференцируемы во всей области определения. Пусть точка $x^0 = (x_1^0,\ldots,x_n^0)$ допустима и в ней достигается локальный условный максимум в поставленной задаче.

Тогда существуют такие неотрицательные числа $\lambda_1,\ldots,\lambda_m \ge 0$, что для функции Лагранжа

$$L(x_1,\ldots,x_n,\lambda_1,\ldots,\lambda_m) = f(x_1,\ldots,x_n) - \lambda_1\phi_1(x_1,\ldots,x_n) - \ldots - \lambda_m\phi_m(x_1,\ldots,x_n)$$

точка x^0 есть точка безусловного локального максимума в области определения $x_1 \geq 0, \ldots, x_n \geq 0$. При этом если $\phi_k(x_1^0, \ldots, x_n^0) < 0$ для какого-нибудь k, то $\lambda_k = 0$ $(1 \leq k \leq m)$. □

(Без доказательства.)

Последняя теорема указывает способ нахождения условных экстремумов на заданном подмножестве $D \subset \mathsf{R}^n$, ограниченном линиями уровня функция ϕ_k. Если экстремум находится строго внутри множества D, то он оказывается безусловным экстремумом функции f. В этом случае все множители Лагранжа могут быть выбраны нулевыми. Если же точка экстремума оказывается на границе D, то ограничения типа неравенство можно заменить на ограничения типа равенство. А для таких задач можно воспользоваться теоремой о стационарных точках функции Лагранжа.

Пример.

Пусть $x_1 = K$ – затраты капитала, $x_2 = L$ – затраты труда, а $f(K, L) = f(x_1, x_2)$ – производственная функция, т.е. объем производства некоторого продукта при заданной паре (K, L). Часто используют модель Кобба–Дугласа, в которой

$$f(x_1, x_2) = a_0 x_1^{\alpha_1} x_2^{\alpha_2},$$

и параметры a_0, α_1, α_2 – неотрицательные и удовлетворяющие равенству $\alpha_1 + \alpha_2 = 1$. Эти параметры обычно настраиваются эмпирически по статистическим данным.

Производственная функция выпукла вверх, равна нулю при $x_1 = 0$ и при $x_2 = 0$ и положительно однородна, т.е. $f(px_1, px_2) = pf(x_1, x_2)$ при всех положительных p.

Пусть p_0 – цена выпускаемой продукции, а p_1 и p_2 – соответственно, цены первого ресурса и второго ресурса. Тогда общие затраты есть

$$C = p_1 x_1 + p_2 x_2,$$

а доход равен

$$R = p_0 f(x_1, x_2).$$

Величина R подлежит максимизации при условии ограниченности объема затрат. Получается следующая задача математического программирования:

$$\begin{gathered} p_0 f(x_1, x_2) \to \max \\ p_1 x_1 + p_2 x_2 \leq C_0, \\ x_1 \geq 0, \ x_2 \geq 0 \end{gathered}$$

Линия уровня $C = C_0$ на плоскости (x_1, x_2) для функции затрат называется изокостой, а линия уровня $f = f_0$ производственной функции – изоквантой. Можно показать на рисунке, что в оптимальной точке эти линии касаются.

Найдем необходимые условия оптимальности плана. Безусловных экстремумов в данной задаче нет, так как положительно однородная функция f их не может иметь. Поэтому надо искать стационарные точки функции Лагранжа. Составим эту функцию:

$$L(x_1, x_2, \lambda) = p_0 f(x_1, x_2) - \lambda(p_1 x_1 + p_2 x_2 - C_0).$$

По условию, функция L дифференцируема в каждой точке области определения. Поэтому необходимым условием точки максимума будет система

$$\frac{\partial L}{\partial x_1} = 0, \quad \frac{\partial L}{\partial x_2} = 0, \quad \frac{\partial L}{\partial \lambda} = 0,$$

которая записывается в виде

$$\begin{aligned}
\frac{\partial f}{\partial x_1}(x_1, x_2) &= \lambda \frac{p_1}{p_0}, \\
\frac{\partial f}{\partial x_2}(x_1, x_2) &= \lambda \frac{p_2}{p_0}, \\
p_1 x_1 + p_2 x_2 &= C_0.
\end{aligned}$$

Это система трех уравнений с тремя неизвестными: x_1, x_2, λ. Для модели Кобба–Дугласа ее удается решить аналитически.

Частные производные в модели Кобба–Дугласа вычисляются непосредственно:

$$\begin{aligned}
\frac{\partial f}{\partial x_1}(x_1, x_2) &= a_0 \alpha_1 x_1^{\alpha_1 - 1} x_2^{\alpha_2} = a_0 \alpha_1 \left(\frac{x_2}{x_1}\right)^{\alpha_2}, \\
\frac{\partial f}{\partial x_2}(x_1, x_2) &= a_0 \alpha_2 x_1^{\alpha_1} x_2^{\alpha_2 - 1} = a_0 \alpha_2 \left(\frac{x_1}{x_2}\right)^{\alpha_1}.
\end{aligned}$$

Поэтому первые два уравнения в системе записываются в виде

$$\begin{aligned}
\frac{x_2}{x_1} &= \left(\frac{\lambda p_1}{a_0 p_0 \alpha_1}\right)^{\frac{1}{\alpha_2}}, \\
\frac{x_1}{x_2} &= \left(\frac{\lambda p_2}{a_0 p_0 \alpha_2}\right)^{\frac{1}{\alpha_1}}.
\end{aligned}$$

Умножим почленно эти уравнения и получим одно уравнение относительно одной неизвестной λ:

$$1 = \left(\frac{\lambda p_1}{a_0 p_0 \alpha_1}\right)^{\frac{1}{\alpha_2}} \left(\frac{\lambda p_2}{a_0 p_0 \alpha_2}\right)^{\frac{1}{\alpha_1}}.$$

Возведем это равенство в степень $\alpha_1 \alpha_2$ и сгруппируем однородные члены:

$$1 = \left(\frac{\lambda}{a_0 p_0}\right)^{\alpha_1 + \alpha_2} \left(\frac{p_2}{\alpha_2}\right)^{\alpha_2} \left(\frac{p_1}{\alpha_1}\right)^{\alpha_1} = \frac{\lambda}{a_0 p_0} \left(\frac{p_2}{\alpha_2}\right)^{\alpha_2} \left(\frac{p_1}{\alpha_1}\right)^{\alpha_1}.$$

Отсюда находим оптимальный множитель Лагранжа:

$$\lambda^0 = a_0 p_0 \left(\frac{\alpha_2}{p_2}\right)^{\alpha_2} \left(\frac{\alpha_1}{p_1}\right)^{\alpha_1}.$$

Здесь и далее стационарные точки функции Лагранжа обозначаются λ^0, x_1^0, x_2^0. Подставим найденный множитель в остальные уравнения:

$$\begin{aligned}
\frac{x_2^0}{x_1^0} &= \left[\frac{p_1}{\alpha_1} \left(\frac{\alpha_2}{p_2}\right)^{\alpha_2} \left(\frac{\alpha_1}{p_1}\right)^{\alpha_1}\right]^{\frac{1}{\alpha_2}} \\
&= \left[\left(\frac{p_1}{\alpha_1}\right)^{1 - \alpha_1} \left(\frac{\alpha_2}{p_2}\right)^{\alpha_2}\right]^{\frac{1}{\alpha_2}} = \frac{p_1}{p_2} \frac{\alpha_2}{\alpha_1}.
\end{aligned}$$

Обозначим последнее выражение буквой k. Оно не зависит от ограничения C_0 на объем капитала. На плоскости (x_1, x_2) оптимальная пара (x_1^0, x_2^0) лежит на прямой

$x_2 = kx_1$. Остается найти пересечение этой прямой с изокостой $p_1x_1 + p_2x_2 = C_0$. Нетрудно видеть, что это точка

$$\begin{aligned} x_1^0 &= \frac{C_0}{p_1 + kp_2} = \frac{C_0\alpha_1}{p_1}, \\ x_2^0 &= \frac{kC_0}{p_1 + kp_2} = \frac{C_0\alpha_2}{p_2}. \end{aligned}$$

Максимальное значение дохода равно

$$R_{\max} = C_0 \left(\frac{\alpha_1}{p_1}\right)^{\alpha_1} \left(\frac{\alpha_2}{p_2}\right)^{\alpha_2}.$$

Полученное аналитическое решение задачи о максимизации дохода имеет наглядное геометрическое истолкование. По данным пропорциям p_2/p_1 и α_1/α_2 определяется угол наклона k и луч $x_2 = kx_1$. Оптимальный план всегда лежит на этом луче.

Если объем капитала C_0 фирмы будет увеличиваться с течением времени, то оптимальные решения будут двигаться по этому лучу. По этой причине такой луч иногда называется долговременной линией развития фирмы. Решение можно получить геометрически, построив пересечение двух прямых: долговременной линии развития $x_2 = kx_1$ и изокосты $p_1x_1 + p_2x_2 = C_0$.

Линейное программирование

Частный случай задачи математического программирования, в котором все функции линейны или аффинны, называется задачей линейного программирования. Эти задачи играют огромную роль в математических методах экономики, для их решения написаны специальные пакеты программ, имеющие коммерческое значение.

Задача линейного программирования имеет следующую стандартную запись:

$$\begin{gathered} \textstyle\sum_{j=1}^n c_jx_j \to \max \\ \textstyle\sum_{j=1}^n a_{1j}x_j \le b_1, \\ \dots \\ \textstyle\sum_{j=1}^n a_{mj}x_j \le b_m, \\ x_1 \ge 0, \ \dots \ x_n \ge 0. \end{gathered}$$

Последняя строчка неравенств называется общими ограничениями, остальные называются специальными. Здесь матрица (a_{ij}), $1{\le}i{\le}m$, $1{\le}j{\le}n$, и векторы $(b_i)_{i=1}^m$, $(c_j)_{j=1}^n$ заданы, и требуется найти вектор $(x_j)_{j=1}^n$.

Отметим, что если операция max заменена на min, или если в некоторых специальных неравенствах знак $\le$ заменен на $\ge$, то задача также относится к классу линейного программирования. Эти задачи сводятся друг к другу заменой знаков у коэффициентов. Действительно, эквивалентны следующие операции:

$$\begin{aligned} \sum_{j=1}^n d_jx_j \to \min \quad &\Leftrightarrow \quad \sum_{j=1}^n (-d_j)x_j \to \max \\ \sum_{j=1}^n a_{ij}x_j \ge b_i, \quad &\Leftrightarrow \quad \sum_{j=1}^n (-a_{1j})x_j \le -b_i. \end{aligned}$$

Важнейшим приемом анализа задач линейного программирования является переход к другой задаче, называемой двойственной и относящейся также к линейному программированию.

Поскольку линейная функция, подлежащая максимизации, не имеет стационарных точек, то применима теорема о множителях Лагранжа. Составим формально функцию Лагранжа только по специальным ограничениям:

$$L(x_1,\dots,x_n,\lambda_1,\dots,\lambda_m)=\sum_{j=1}^{n}c_jx_j-\sum_{i=1}^{m}\lambda_i\left(\sum_{j=1}^{n}a_{ij}x_j-b_i\right),$$

где $\lambda_i\ge 0$, $1\le i\le m$, — множителя Лагранжа.

Рассмотрим две задачи:

$$\min_{\lambda_i\ge 0,1\le i\le m}\ \max_{x_j\ge 0,1\le j\le n} L(x,\lambda),\qquad \max_{x_j\ge 0,1\le j\le n}\ \min_{\lambda_i\ge 0,1\le i\le m} L(x,\lambda),$$

первая из которых называется минимаксной, а вторая — максиминной. В этих задачах отсутствуют специальные ограничения задач линейного программирования, а остаются только общие — неотрицательность всех переменных.

Лемма.

Точная нижняя грань линейной функции

$$\inf_{y\ge 0} cy$$

равна $-\infty$, если $c<0$, и равна 0, если $c\ge 0$. В последнем случае оптимальным является значение $y=0$.

Аналогично, точная верхняя грань линейной функции

$$\sup_{y\ge 0} cy$$

равна ∞, если $c>0$, и равна 0, если $c\le 0$. В последнем случае оптимальным является значение $y=0$. □

Доказательство очевидно.

Рассмотрим максиминную задачу:

$$\begin{aligned}L_{\min\max} &= \max_{x_j\ge 0,1\le j\le n}\ \min_{\lambda_i\ge 0,1\le i\le m} L(x,\lambda)\\ &= \max_{x_j\ge 0}\left\{\sum_{j=1}^{n}c_jx_j+\min_{\lambda_i\ge 0}\left[\sum_{i=1}^{m}\lambda_i\left(b_i-\sum_{j=1}^{n}a_{ij}x_j\right)\right]\right\}.\end{aligned}$$

По лемме, если выражение в круглых скобках хотя бы для одного индекса i отрицательно, то минимум по λ_i даст $-\infty$. Это значение говорит о том, что числа $(x_j)_{j=1}^{n}$ выбраны неудачно, ведь по ним максимизируется функция. Следовательно, достаточно рассматривать только такие числа $(x_j)_{j=1}^{n}$, для которых все выражения в круглых скобках неотрицательны. По той же лемме в этом случае оптимальные значения λ_i равны 0 при всех $1\le i\le m$. Таким образом,

$$L_{\min\max}=\max_{x_j\ge 0,1\le j\le n}\left\{\sum_{j=1}^{n}c_jx_j\ \Big|\ \sum_{j=1}^{n}a_{ij}x_j\le b_i,\ 1\le i\le m\right\}.$$

Но последняя задача в точности совпадает с исходной задачей линейного программирования.

Важнейшим свойством функций, линейных отдельно по аргументам x_j и отдельно по аргументам λ_i, является совпадание решений минимаксной и максиминной задачи. Это следует из так называемой теоремы о седловой точке (или теоремы о минимаксе) в математической теории игр. В седловой точке изменение одной группы переменных $(x = (x_j))$ приводит к уменьшению целевой функции $(L(x, \lambda))$, а изменение другой группы переменных $(\lambda = (\lambda_i))$ — к ее увеличению. Поскольку цели выбора этих переменных как раз противоположные, то в седловой точке любое изменение переменных "невыгодно". Эта точка дает решение одновременно минимаксной и максиминной задачи, а значение целевой функции в этой точке называется ценой игры. Полная формулировка теоремы о седловой точке здесь не приводится, так как она содержит новые понятия и относится к теории выпуклого анализа, не входящей в данный курс.

Итак, максимин равен минимаксу, и для минимаксной задачи выполним те же операции:

$$
\begin{aligned}
L_{\max\min} &= \min_{\lambda_i \geq 0, 1 \leq i \leq m} \max_{x_j \geq 0, 1 \leq j \leq n} L(x, \lambda) \\
&= \min_{\lambda_i \geq 0} \left\{ \sum_{i=1}^{m} \lambda_i b_i + \max_{x_j \geq 0} \left[\sum_{j=1}^{n} x_j \left(c_j - \sum_{i=1}^{m} a_{ij} \lambda_i \right) \right] \right\} \\
&= \min_{\lambda_i \geq 0, 1 \leq i \leq m} \left\{ \sum_{i=1}^{m} b_i \lambda_i \mid \sum_{i=1}^{m} a_{ij} \lambda_i \geq c_j,\ 1 \leq j \leq n \right\}.
\end{aligned}
$$

В итоге получилась новая задача линейного программирования. Она называется двойственной к исходной задаче линейного программирования.

Легко показать, что если двойственную задачу на условный максимум привести методом замены знаков у коэффициентов к максиминной задаче, для нее снова можно написать (вторую) двойственную задачу, которая совпадет с исходной. Таким образом, задачи линейного программирования объединяются по парам. В каждой паре любая из задач является двойственной ко второй.

Двойственные задачи линейного программирования обычно записываются рядом:

$$
\begin{array}{c|c}
\sum_{j=1}^{n} c_j x_j \to \max & \sum_{i=1}^{m} b_i y_i \to \min \\
\sum_{j=1}^{n} a_{1j} x_j \leq b_1, & \sum_{i=1}^{m} a_{i1} y_i \geq c_1, \\
\dots & \dots \\
\sum_{j=1}^{n} a_{mj} x_j \leq b_m, & \sum_{i=1}^{m} a_{in} y_i \geq c_n, \\
x_1 \geq 0, \dots, x_n \geq 0. & y_1 \geq 0, \dots, y_m \geq 0.
\end{array}
$$

Для правильного написания двойственной задачи (чему необходимо научиться каждому студенту) полезно запомнить следующие правила:

- Каждому ограничению одной задачи соответствует переменная другой задачи.
- Переменные b_i и c_j меняются местами.
- В паре двойственных задач: одна – минимизации, другая – максимизации.

- В специальных неравенствах знак $\leq$ меняется на $\geq$.
- Общие неравенства выглядят одинаково: неотрицательность свободных переменных.
- Матрица специальных неравенств a_{ij} транспонируется при переходе к двойственной задаче: если в одной задаче в каждом ограничении эта матрица суммируется вдоль строки, то в двойственной задаче она суммируется вдоль столбца.

Кроме того, оптимальные векторы решений двойственных задач $x^0 = (x_j^0)_{j=1}^n$ и $y^0 = (y_i^0)_{i=1}^m$ обладают следующим свойством трансверсальности:
если $\sum_{j=1}^n a_{ij}x_j^0 < b_i$ для некоторого i, то $y_i^0 = 0$;
если $\sum_{i=1}^m a_{ij}y_i^0 > c_j$ для некоторого j, то $x_j^0 = 0$.

Наиболее распространенные пакеты программ по решению задач линейного программирования основаны на так называемом симплекс–методе. В настоящее время программы успешно справляются с задачами, в которых количество переменных и ограничений измеряется сотнями. Особое внимание уделено эффективности алгоритмов по быстродействию и памяти в задачах, содержащих матрицы a_{ij} с большим количеством нулей.

В заключение следует отметить, что теория линейного программирования зарождалась в Ленинграде в 30–е годы и связана с именем лауреата Нобелевской премии Л.В.Канторовича.

Дифференциальные уравнения

Определение *дифференциального уравнения.*

Дифференциальным уравнением n-го порядка называется уравнение вида

$$\phi(x, y, y', \ldots, y^{(n)}) = 0,$$

где x — аргумент, а $y = y(x)$ — функция, которую требуется найти. □

В частности, дифференциальное уравнение первого порядка записывается в виде

$$\phi(x, y, y') = 0.$$

Отметим, что в отличие от алгебраических уравнений неизвестным в дифференциальном уравнении является не число и не набор чисел, а функция, заданная на промежутке.

Определение *решение дифференциального уравнения.*

Решением дифференциального уравнения называется функция, при подстановке которой в уравнение получается верное тождество. График решения дифференциального уравнения называется *интегральной кривой. Решением в квадратурах* называется решение дифференциального уравнения, записанное при помощи конечного числа интегралов. □

Определение. *Виды уравнения первого порядка.*

Уравнение n-го порядка называется разрешенным относительно старшей (или высшей) производной, если оно записано в виде

$$y^{(n)} = f(x, y, y', \ldots, y^{(n-1)}).$$

Уравнение первого порядка, разрешенное относительно производной, имеет вид

$$y' = f(x, y) \qquad \text{или} \qquad \frac{dy}{dx} = f(x, y).$$

Оно может быть также записано $dy = f(x, y)dx$.

Дифференциальное уравнение

$$M(x, y)\, dx + N(x, y)\, dy = 0$$

называется уравнением первого порядка, записанным в дифференциальной форме. Если считать, что y есть функция от x, то это уравнение равносильно следующему:

$$\frac{dy}{dx} = -\frac{M(x, y)}{N(x, y)} \qquad \text{или} \qquad y' = -\frac{M(x, y)}{N(x, y)}.$$

Однако, можно считать, что x есть функция от y, и тогда

$$x' = -\frac{N(x, y)}{M(x, y)}.$$

□

Геометрический смысл дифференциального уравнения первого порядка

Пусть функция f определена в области $G \subseteq \mathsf{R}^2$ и задано дифференциальное уравнение $y' = f(x, y)$. Выберем произвольную точку (x_0, y_0) в области G и предположим, что некоторая некоторая интегральная кривая проходит через эту точку. Тогда тангенс угла наклона касательной в этой точке равен $f(x, y)$. Но через заданную точку можно провести только одну прямую с заданным углом наклона.

Таким образом, дифференциальное уравнение $y' = f(x, y)$ определяет в каждой точке (x, y) области G некоторую прямую, проходящую через эту точку. График любого решения $y = y(x)$ дифференциального уравнения касается заданных прямых в каждой точке. Поэтому задачу нахождения решения дифференциального уравнения можно сформулировать на геометрическом языке так:

на плоскости каждой точке поставлена в соответствие некоторая прямая, проходящая через нее; требуется найти линию в области G, которая в каждой своей точке касается соответствующей прямой.

Одна гладкая интегральная кривая не может заполнить всей области G на плоскости. Через те точки, которые не находятся на такой кривой, можно попытаться провести новые интегральные кривые, касающиеся соответствующих прямых в каждой своей точке. Разные кривые являются графиками разных функций. Поэтому решение дифференциального уравнения, как правило, неединственно. Однако, если фиксировать хотя бы одну точку интегральной кривой, то таких решений оказывается значительно меньше.

Задача Коши. Ломаные Эйлера

Определение *задачи Коши для дифференциального уравнения.*

Пусть функция f задана в области G на плоскости и выбрана точка $(x_0, y_0) \in G$. *Задача Коши* состоит в нахождении такого решения $y = y(x)$ дифференциального уравнения $y' = f(x, y)$, что $y(x_0) = y_0$, т.е. график этого решения проходит через точку (x_0, y_0).

Заданная точка (x_0, y_0) или условие $y(x_0) = y_0$ называется *начальным данным* в задаче Коши. □

Некоторым приближением к решению задачи Коши являются кусочно–аффинные функции, графики которых называются *ломаными Эйлера.* Они определяются следующим образом.

Пусть задана задача Коши: $y' = f(x, y)$, $y(x_0) = y_0$. Пусть $a = x_0$ и $b > a$. Промежуток $[a, b]$ разобьем на n частей точками

$$a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b.$$

На промежутке $[x_0, x_1]$ определим аффинную функцию, график которой является частью касательной прямой, связанной с точкой (x_0, y_0):

$$\widetilde{y}(x) = y_0 + f(x_0, y_0)(x - x_0).$$

Эта функция в точке x_1 принимает некоторое значение

$$y_1 = \widetilde{y}(x_1) = y_0 + f(x_0, y_0)(x_1 - x_0).$$

Далее, на промежутке $[x_1, x_2]$ определим аффинную в соответствии с касательной в точке x_1 и т. д. По индукции функция $\widetilde{y}$ опредяется на всем промежутке $[a, b]$ равенствами

$$\widetilde{y}(x) = y_k + f(x_k, y_k)(x - x_k), \qquad x_k \le x \le x_{k+1},$$

при $0 \le k \le n-1$. График этой функции называется ломаной Эйлера.

Построенная функция $\widetilde{y}(x)$ аффинна на каждом промежутке $[x_k, x_{k+1}]$, а в точках x_k имеет изломы. За это она называется кусочно–аффинной.

Рассмотрим последовательность разбиений промежутка $[a, b]$, для каждого из которых строится ломаная Эйлера. Получается последовательность ломаных Эйлера. Пусть $n \to \infty$ и мелкость дробления $\max_{0 \le k \le n-1}(x_{k+1} - x_k)$ стремится к нулю. Предположим также, что функция $f(x, y)$ непрерывно дифференцируема.

Тогда можно доказать, что последовательность ломаных Эйлера сходится, и даже равномерно, к некоторой интегральной кривой, которая дает решение исходной задаче Коши. По этой причине ломаные Эйлера являются аппроксимациями точного решения, и на этом основаны многие численные методы решения дифференциальных уравнений.

Теоремы существования и единственности для уравнения первого порядка

Важный теоретический вопрос при изучении дифференциальных уравнений состоит в следующем: при каких условиях задача Коши *имеет* какое–либо решение, и при каких условиях такое решение *единственно*?

Теорема *существования решения задачи Коши.*

Пусть функция $f(x, y)$ непрерывна в области G и $(x_0, y_0) \in G$. Тогда существует промежуток (α, β), содержащий точку x_0, и функция $y(x)$ на (α, β), которая решает задачу Коши: $y' = f(x, y)$, $y(x_0) = y_0$. □

Доказательство основано на свойствах сходимости ломаных Эйлера и здесь не приводится. Рассмотрим пример, иллюстрирующий это свовство сходимости.

Пример. Требуется построить решение задачи Коши для уравнения $y' = 2y$ с начальным данным $y(0) = 1$ и на промежутке $[0, 1]$.

Разделим промежуток $[0, 1]$ на n частей точками $x_{n,k} = k/n$, $0 \le k \le n$. Ломаная Эйлера для этого разбиения есть график кусочно–аффинной функции $\widetilde{y}_n(x)$. Здесь индекс n указывает на номер функции, совпадающий с количеством узлов разбиения, а не на номер узла.

Обозначим ординаты ломаной Эйлера в узлах разбиения через $\widetilde{y}_n(x_k) = y_{n,k}$. По определению ломаной Эйлера,

$$\widetilde{y}_n(x) = y_{n,k} + f(x_{n,k}, y_{n,k})(x - x_{n,k}), \qquad x_{n,k} \le x \le x_{n,k+1},$$

при $f(x, y) = 2y$ и $0 \le k \le n-1$. Значения в узлах вычисляются непосредственно:

$$y_{n,k+1} = y_{n,k} + 2y_{n,k}(x_{n,k+1} - x_{n,k}) = \left(1 + \frac{2}{n}\right) y_{n,k}.$$

Отсюда

$$y_{n,k} = \left(1 + \frac{2}{n}\right)^k, \qquad 0 \le k \le n.$$

Выберем точку $x \in [0,1]$. Пусть $k_n(x)$ — номер ближайшей к x точки в разбиении на n частей. Тогда $k_n(x)/n \approx x$ или $k_n(x) \approx nx$.

Легко доказать, что ломаные Эйлера имеют предел:

$$\lim_{n\to\infty} \widetilde{y}_n(x) = \lim_{n\to\infty} \left(1+\frac{2}{n}\right)^{k_n(x)} = \lim_{n\to\infty} \left(1+\frac{2}{n}\right)^{nx} = e^{2x}.$$

Нетрудно видеть, что функций $y(x) = e^{2x}$ удовлетворяет как дифференциальному уравнению, так и начальному данному, и поэтому решает поставленную задачу Коши. □

Теорема *единственности решения задачи Коши.*

Пусть функция $f(x,y)$ непрерывна по x и непрерывно дифференцируема по y в области G и $(x_0, y_0) \in G$. Тогда решение задачи Коши $y' = f(x,y)$, $y(x_0) = y_0$ в области G единственно. □

(Без доказательства.)

Контрпример. Рассмотрим задачу Коши: $y' = \sqrt{|y|}$, $y(0) = 0$. Она имеет по крайней мере два решения: $y(x) \equiv 0$ и $y(x) = x|x|/4$. Решение неединственно и заключение последней теоремы не выполнено, потому что призводная правой части

$$\left(\sqrt{|y|}\right)'_y = \frac{\operatorname{sign} y}{2\sqrt{|y|}}$$

не является непрерывной в точке $y = 0$.

Системы уравнений первого порядка

По аналогии с системами алгебраических уравнений несколько дифференциальных уравнений относительно нескольких неизвестных составляют систему дифференциальных уравнений. Только неизвестными являются функции, а не числа.

Определение *системы дифференциальных уравнений.*

Системой n дифференциальных уравнений первого порядка с n неизвестными, разрешенной относительно производных, называется система

$$\begin{cases} y_1' &= f_1(x, y_1, y_2, \ldots, y_n), \\ y_2' &= f_2(x, y_1, y_2, \ldots, y_n), \\ \ldots & \ldots \ \ldots \\ y_n' &= f_n(x, y_1, y_2, \ldots, y_n). \end{cases}$$

Решением системы называется совокупность функций $y_1(x), y_2(x), \ldots, y_n(x)$, заданных на одном и то же промежутке, которые после подстановки превращают систему в верное тождество. □

Систему дифференциальных уравнений принято записывать в векторной форме. Обычно вводят вектор $y = (y_1, \ldots, y_n)$, состоящий из неизвестных функций, и вектор $f(x,y) = (f_1(x,y), \ldots, f_n(x,y))$, состоящий из правых частей системы дифференциальных уравнений. В этих обозначениях система записывается в виде

$$y' = f(x,y),$$

который не отличается внешне от одного дифференциального уравнения. Однако здесь y и f — векторы.

Любое дифференциальное уравнение n-го порядка, разрешенное относительно старшей производной, можно преобразовать к системе дифференциальных уравнений первого порядка, разрешенных относительно производной. Стандартное преобразование состоит в следующем. Пусть задано уравнение

$$y^{(n)} = f(x, y, y', \ldots, y^{(n-1)}).$$

Обозначим производные функции y новыми переменными: $y' = y_1$, …, $y^{(n-1)} = y_{n-1}$. Тогда заданное уравнение равносильно системе

$$\left\{ \begin{array}{rcl} y' & = & y_1, \\ y_1' & = & y_2, \\ \ldots & \ldots & \ldots \\ y_{n-1}' & = & f(x, y, y_1, \ldots, y_{n-1}). \end{array} \right.$$

Таким образом, для уравнений, разрешенных относительно старших производных, теория уравнений высших порядков является частью теории систем уравнений первого порядка.

Рассмотрим систему дифференциальных уравнений первого порядка, разрешенных относительно производной:

$$\left\{ \begin{array}{rcl} y_1' & = & f_1(x, y_1, y_2, \ldots, y_n), \\ y_2' & = & f_2(x, y_1, y_2, \ldots, y_n), \\ \ldots & \ldots & \ldots \\ y_n' & = & f_n(x, y_1, y_2, \ldots, y_n) \end{array} \right.$$

или в векторной форме

$$y' = f(x, y).$$

Пусть все функции f_i определены в некоторой области $G \subseteq \mathsf{R} \times \mathsf{R}^n$, т.е. значения $f_i(x, y_1, \ldots, y_n)$ определены, если $(x, y_1, \ldots, y_n) \in G$. Выберем какую–нибудь точку $(x_0, y_0) \in G$, где $y_0 = (y_{10}, \ldots, y_{n0})$.

Определение. *Задача Коши для системы уравнений первого порядка.*

Для заданной системы дифференциальных уравнений и заданной точки $(x_0, y_0) \in G$ найти решение $y(x) = (y_1(x), \ldots, y_n(x))$, удовлетворяющее (векторным) начальным данным (или начальным условиям) $y(x_0) = y_0$. □

Теорема *существования для систем дифференциальных уравнений.*

Пусть $(x_0, y_0) \in G$ и задана задача Коши: $y' = f(x, y)$, $y(x_0) = y_0$. Если функция f непрерывна в области G, то решение существует. (Т.е. существует промежуток (α, β), содержащий x_0, и вектор–функция $y(x)$ на нем, которая удовлетворяет системе уравнений и начальному условию.) □

Теорема *единственности для систем дифференциальных уравнений.*

Пусть функция $f(x, y_1, \ldots, y_n)$ непрерывна по x в области G и имеет непрерывные частные производные по $y_1, \ldots, y_n$ в области G. Пусть $(x_0, y_0) \in G$ и задана задача

Коши: $y' = f(x, y)$, $y(x_0) = y_0$. Тогда решение этой задачи в области G единственно. □

Замечание о теоремах существования и единственности для дифференциальных уравнений n-го порядка, разрешенных относительно старшей производной. Рассмотрим такое уравнение:

$$y^{(n)} = f(x, y, y', \dots, y^{(n-1)}).$$

Здесь f — функция от $n+1$ переменной. Стандартный способ сведения этого уравнения к системе первого порядка включает обозначения $y_0 = y$, $y_1 = y'$, …, $y^{(n-1)} = y_{n-1}$. В правой части системы дифференциальных уравнений появляется функция $f(x, y_0, y_1, \dots, y_{n-1})$.

Остается применить теоремы существования и единственности для систем дифференциальных уравнений. Для существования решения задачи Коши достаточно, чтобы функция f была непрерывна по всем своим аргументам. Для единственности решения достаточно, чтобы частные производные переменным y_0, y_1, …, y_{n-1} существовали и были непрерывны.

Общий интеграл и общее решение системы дифференциальных уравнений

Пусть в области $G \subseteq \mathsf{R} \times \mathsf{R}^n$ задана вектор–функция f и система дифференциальных уравнений $y' = f(x, y)$.

Определение *общего решения системы дифференциальных уравнений.*

Общим решением системы дифференциальных уравнений называется функция от $n+1$ аргумента $Y(x, C_1, \dots, C_n)$, обладающая следующими свойствами:

1. При любых фиксированных числах $C_1 \in \mathsf{R}$, …, $C_n \in \mathsf{R}$ функция одного аргумента $y(x) = Y(x, C_1, \dots, C_n)$ является решением системы уравнений.

2. Для любого решения системы уравнений $y(x)$ в области G существуют такие числа $C_1 \in \mathsf{R}$, …, $C_n \in \mathsf{R}$, что $y(x) = Y(x, C_1, \dots, C_n)$. □

Другими словами, общее решение параметризует все интегральные кривые в области G: любое решение системы уравнений получается выбором некоторых параметров $C_1, \dots, C_n$ и наоборот, при любом выборе параметров получается некоторое решение.

Если выполнены условия теорем существования и единственности, то каждая задача Коши в области G имеет единственное решение, или через каждую точку в области G можно провести ровно одну интегральную кривую. Каждая интегральная кривая характеризуется набором из n чисел $C_1, \dots, C_n$. В этом смысле общее решение аккумулирует в себе все интегральные кривые. Каждая интегральная кривая связана с набором из n констант $C_1, \dots, C_n$.

Определение *общего интеграла системы дифференциальных уравнений.*

Общим интегралом системы дифференциальных уравнений называется уравнение (или система уравнений) $\Phi(x, y, C_1, \dots, C_n) = 0$ (или сама функция Φ от $2n+1$ аргумента), обладающее следующими свойствами:

1. При любых фиксированных числах $C_1 \in \mathsf{R}$, …, $C_n \in \mathsf{R}$ это уравнение задает интегральную кривую в области G (т.е. множество всех точек области G, удовлетворяющее этому уравнению, есть интегральная кривая для системы дифференциальных уравнений).

2. Для любой интегральной кривой в области G существуют такие числа $C_1 \in \mathsf{R}$, …, $C_n \in \mathsf{R}$, что кривая задается уравнением $\Phi(x, y, C_1, \dots, C_n) = 0$. □

Понятия общего решения и общего интеграла тесно связаны между собой. Если $Y(x, C_1, \dots, C_n)$ есть общее решение, то уравнение

$$y - Y(x, C_1, \dots, C_n) = 0$$

есть общий интеграл системы дифференциальных уравнений.

Наоборот, если уравнение общего интеграла $\Phi(x, y, C_1, \dots, C_n) = 0$ удается решить относительно y (либо если это уравнение задает неявную функцию Y), то это решение $y = Y(x, C_1, \dots, C_n)$ есть общее решение системы дифференциальных уравнений.

Обычно нахождение общего решения является целью исследования системы дифференциальных уравнений. Однако, в процессе вычислений сначала появляется общий интеграл, который затем надо решить относительно переменных $y = (y_1, \dots, y_n)$.

Алгоритм решения задачи Коши. Пусть задана задача Коши: $y' = f(x, y)$, $y(x_0) = y_0$. Обычно сначала стараются избавиться от производных в дифференциальном уравнении, или проинтегрировать его. Если это удается, то после интегрирования n уравнений появляются n констант интегрирования $C_1, \dots, C_n$, а дифференциальное уравнение преобразуется к общему интегралу $\Phi(x, y, C_1, \dots, C_n) = 0$.

Затем определяются неизвестные константы. Поскольку нас интересуют решение, проходящее через точку (x_0, y_0), то $\Phi(x_0, y_0, C_1, \dots, C_n) = 0$. Это система алгебраических уравнений относительно чисел $C_1, \dots, C_n$. Предположим, что ее удается решить и получаются значения $C_{10}, \dots, C_{n0}$. Тогда искомое решение задачи Коши $y = y(x)$ удовлетворяет уравнению

$$\Phi(x, y, C_{10}, \dots, C_{n0}) = 0.$$

В этом уравнении переменными являются только x и $y = (y_1, \dots, y_n)$. Если это уравнение удается решить относительно переменных y, то получается функция $y(x)$, являющаяся решением поставленной задачи Коши.

Методы интегрирования уравнений первого порядка.

Далеко не все дифференциальные уравнения интегрируются в квадратурах, т.е. их общие интегралы представимы в виде выражения, содержащего конечный набор интегралов. Это удается сделать скорее в исключительных случаях. Существуют справочники по дифференциальным уравнениям, содержащие общие интегралы и общие решения конкретных дифференциальных уравнений. Однако имеются классы уравнений, которые всегда интегрируются в квадратурах, и их общие решения стандартным образом вычисляются.

Общее решение и общий интеграл одного дифференциального уравнения первого порядка содержат только одну константу C, так как $n = 1$.

1. Уравнения с разделяющимися переменными

Дифференциальное уравнение вида

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

называется уравнением с разделяющимися переменными.

Это уравнение интегрируется в квадратурах. Пусть F — некоторая первообразная функции f, а G — некоторая первообразная функции $1/g$. Тогда уравнение

$$G(y) - F(x) - C = 0$$

является общим интегралом дифференциального уравнения.

Действительно, если $y(x)$ — некоторое решение дифференциального уравнения, то

$$(G(y(x)) - F(x))' = G'(y(x))y'(x) - F'(x) = \frac{y'(x)}{g(y(x))} - f(x) = 0.$$

Следовательно, на каждой интегральной кривой функция $G(y) - F(x)$ есть величина постоянная.

На практике удобно сначала преобразовать исходное уравнение к дифференциальному виду:

$$dy = f(x)g(y)\,dx.$$

Затем нужно переменные "разделить" и проинтегрировать:

$$\frac{dy}{g(y)} = f(x)\,dx, \qquad \int \frac{dy}{g(y)} = \int f(x)\,dx, \qquad G(y) = F(x) + C.$$

Общее решение получится, если уравнение $G(y) - F(x) - C = 0$ решить относительно y.

Пример 1.

$$(1 + y^2)\,dx + (1 + x^2)\,dy = 0.$$

Это уравнение уже записано в дифференциальной форме. Поэтому остается разделить переменные и проинтегрировать:

$$\begin{aligned} \frac{dy}{1+y^2} &= -\frac{dx}{1+x^2}, \\ \int \frac{dy}{1+y^2} &= -\int \frac{dx}{1+x^2}, \\ \operatorname{arctg} y &= -\operatorname{arctg} x + C. \end{aligned}$$

Последнее выражение есть общий интеграл дифференциального уравнения. Если его решить относительно y по формуле тангенса разности, то получится общее решение:

$$y = \frac{\operatorname{tg} C - x}{1 + x \operatorname{tg} C}.$$

В последнем выражении $\operatorname{tg} C$ есть произвольная константа, обозначим ее C_1. Общее решение имеет вид

$$Y(x, C_1) = \frac{C_1 - x}{1 + C_1 x},$$

где C_1 — произвольное число из $[-\infty, \infty]$. По установившейся традиции эту константу также обозначают C (общий символ для произвольных констант), хотя это C отличается от числа из предыдущих строчек. Итак,

$$Y(x, C) = \frac{C - x}{1 + Cx}, \qquad C \in [-\infty, \infty].$$

Предположим, что задана задача Коши для того же уравнения и начального условия $y(0) = 3$. Подставим в общее решение $x = 0$ и $y = 3$. Получится $3 = C$. Следовательно, решение поставленной задачи Коши, которое единственно по теореме единственности, имеет вид

$$y(x) = \frac{3 - x}{1 + 3x}.$$

2. Однородные уравнения

Функция $f(x, y)$ называется однородной нулевой степени, если для любых x, y и t

$$f(tx, ty) = f(x, y).$$

Легко доказать, что любая такая функция зависит от отношения y/x (для этого достаточно выбрать $t = 1/x$). Существует такая функция ϕ от одной переменной, что

$$f(x, y) = \phi\left(\frac{y}{x}\right).$$

Дифференциальное уравнение $y' = f(x, y)$ называется однородным, если функция f — однородная нулевой степени.

Однородные дифференциальные уравнения интегрируются в квадратурах. Сделаем замену переменной: $z = y/x$. Здесь z есть функция от x так же, как и y. Следовательно, $y(x) = xz(x)$ и $y' = z + xz'$. Подставим это равенство в дифференциальное уравнение:

$$z + xz' = f(x, y) = \phi(y/x) = \phi(z).$$

Получилось дифференциальное уравнение, в котором можно разделить переменные:

$$\begin{aligned} xz' &= \phi(z) - z, \\ \frac{dz}{\phi(z) - z} &= \frac{dx}{x}, \\ \int \frac{dz}{\phi(z) - z} &= \int \frac{dx}{x} = \ln|x| + C. \end{aligned}$$

Последнее выражение есть общий интеграл преобразованного дифференциального уравнения. Если его решить относительно z, то получится функция $Z(x, C)$. Общее решение исходного однородного уравнения имеет вид

$$Y(x, C) = xZ(x, C).$$

Однородные уравнения часто записываются в дифференциальной форме. Функция f называется однородной степени k, если для любых x, y и t

$$f(tx, ty) = t^k f(x, y).$$

Пусть $M(x,y)$ и $N(x,y)$ — однородные функции одной и той же степени k. Тогда уравнение в дифференциальной форме

$$M(x,y)\,dx + N(x,y)\,dy = 0$$

является однородным дифференциальным уравнением. Действительно, оно равносильно уравнению

$$\frac{dy}{dx} = -\frac{M(x,y)}{N(x,y)} = f(x,y).$$

Очевидно, что функция $f(x,y)$ удовлетворяет условию $f(tx,ty) = f(x,y)$. Поэтому последнее уравнение однородное.

Пример 2.

$$2xy\,dy = (y^2 - x^2)\,dx.$$

Обе функции: $2xy$ и $y^2 - x^2$ — однородные степени 2. Поэтому данное дифференциальное уравнение является однородным.

В соответствии с алгоритмом решения таких уравнений, сделаем замену переменной: $z = y/x$. Тогда $y = zx$, $dy = z\,dx + x\,dz$ и

$$2x(zx)(z\,dx + x\,dz) = ((zx)^2 - x^2)\,dx.$$

После сокращения на x^2 получим

$$2z(z\,dx + x\,dz) = (z^2 - 1)\,dx.$$

Это уравнение с разделяющимися переменными:

$$\begin{aligned}
2zx\,dz &= -(z^2+1)\,dx, \\
\frac{2z\,dz}{z^2+1} &= -\frac{dx}{x}, \\
\int \frac{2z\,dz}{z^2+1} &= -\int \frac{dx}{x}, \\
\ln(z^2+1) &= -\ln|x| + C, \\
z^2+1 &= \frac{C_1}{|x|}, \qquad C_1 = e^C > 0.
\end{aligned}$$

Условие $C_1 > 0$ можно отбросить, если заменить $|x|$ на x. Действительно, знак константы выбирается так, чтобы дробь получалась положительной. Кроме того, вместо C_1 по традиции пишется C. Заменим также z на y/x. После подстановки и элементарных преобразований получится

$$x^2 + y^2 = Cx.$$

Это общий интеграл исходного однородного дифференциального уравнения. Множество интегральных кривых, получаемых при различных C, есть множество всех окружностей с центрами на оси абсцисс и проходящих через начало координат.

Уравнения в полных дифференциалах. Интегрирующий множитель

Рассмотрим уравнение в дифференциальной форме

$$P(x,y)\,dx + Q(x,y)\,dy = 0.$$

Если существует такая функция $u(x, y)$, что

$$P(x, y) = \frac{\partial u}{\partial x}, \qquad Q(x, y) = \frac{\partial u}{\partial y},$$

то уравнение называется уравнением в полных дифференциалах.

В этом случае его можно записать в виде

$$\frac{\partial u}{\partial x}\, dx + \frac{\partial u}{\partial y}\, dy = 0, \qquad \text{или} \qquad d(u(x, y)) = 0.$$

Полный дифференциал равен нулю только у константы, поэтому последнее уравнение равносильно

$$u(x, y) = C.$$

Это общий интеграл уравнения в полных дифференциалах.

Если функции P и Q непрерывно дифференцируемы, то для того, чтобы уравнение было в полных дифференциалах, необходимо и достаточно, чтобы

$$\frac{\partial P(x, y)}{\partial y} = \frac{\partial Q(x, y)}{\partial x}.$$

Необходимость очевидна: каждое из этих выражений есть вторая перекрестная производная от функции $u(x, y)$ по x и по y. Достаточность доказывается методами теории поля, которые не входят в данный курс лекций.

Иногда оказывается, что функции P и Q не удовлетворяют последнему условию и дифференциальное уравнение не есть уравнение в полных дифференциалах, но оно становится таковым после умножения на подходящий множитель $\mu(x, y)$. Уравнение принимает вид

$$\mu(x, y)P(x, y)\, dx + \mu(x, y)Q(x, y)\, dy = 0,$$

а условие того, что в левой части стоит полный дифференциал некоторой функции, —

$$\frac{\partial[\mu(x, y)P(x, y)]}{\partial y} = \frac{\partial[\mu(x, y)Q(x, y)]}{\partial x}.$$

Функция $\mu(x, y)$ называется интегрирующим множителем.

Этот множитель может быть найден из дифференциального уравнения в частных производных относительно неизвестной функции $\mu(x, y)$:

$$Q\frac{\partial \mu}{\partial x} - P\frac{\partial \mu}{\partial y} + \mu\left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) = 0.$$

Пример 3.

$$y' = ay + h(x),$$

где $a \in R$ и h — заданная функция одной переменной. Уравнение можно записать также в форме

$$dy - (ay + h(x))\, dx = 0$$

и определить $P(x, y) = -ay - h(x)$, $Q(x, y) = 1$.

Это уравнение не в полных дифференциалах, так как $\partial P/\partial y \neq \partial Q/\partial x$. Напишем уравнение для интегрирующего множителя:

$$\frac{\partial \mu}{\partial x} + (ay + h(x))\frac{\partial \mu}{\partial y} + a\mu = 0.$$

Этому уравнению удовлетворяет функция $\mu(x, y) = e^{-ax}$. Умножим уравнение на эту функцию:

$$e^{-ax} dy - (ae^{-ax} y + e^{-ax} h(x))\, dx = 0.$$

В левой части уравнения стоит полный дифференциал от функции

$$u(x, y) = e^{-ax} y - \int_0^x e^{-at} h(t)\, dt.$$

Отсюда получаем общий интеграл дифференциального уравнения:

$$e^{-ax} y - \int_0^x e^{-at} h(t)\, dt = C.$$

Общее решение легко получается умножением на e^{ax}:

$$Y(x, C) = e^{ax} C + \int_0^x e^{a(x-t)} h(t)\, dt.$$

Если задана задача Коши с начальным данным $y(0) = y_0$, то подстановка $x = 0$ и $y = y_0$ дает $y_0 = C$. Поэтому единственное решение задачи Коши есть

$$y(x) = e^{ax} y_0 + \int_0^x e^{a(x-t)} h(t)\, dt.$$

Понижение порядка уравнения, содержащего только производные

Рассмотрим уравнение n-го порядка, не содержащего самой неизвестной функции и нескольких первых производных:

$$\phi(x, y^{(k)}, y^{(k+1)}, \ldots, y^{(n)}) = 0.$$

Тогда порядок уравнения можно понизить, введя новую функцию $z = y^{(k)}$. После подстановки получается уравнение

$$\phi(x, z, z', \ldots, z^{(n-k)}) = 0.$$

Если удалось найти общее решение этого уравнения $Z(x, C_1, \ldots, C_{n-k})$, то общее решение исходного уравнения примет вид

$$Y(x, C_1, \ldots, C_n) = \int_0^x \frac{(x-t)^{k-1}}{(k-1)!} Z(t, C_1, \ldots, C_{n-k})\, dt + C_{n-k+1} + C_{n-k+2} x + \ldots + C_n x^{k-1}.$$

Прямым дифференцированием получается, что $Y^{(k)} = Z$. Кроме того, Y содержит n независимых констант.

В частности, уравнение $y^{(n)} = f(x, y^{n-1})$ сводится заменой переменной $z = y^{(n-1)}$ к уравнению первого порядка $z' = f(x, z)$.

Пример 4.

$$y''' - y'' = 6x + 5.$$

Введем обозначение $z = y''$. Тогда

$$z' = z + 6x + 5.$$

В соответствии с результатом в примере 3, общее решение последнего уравнения есть

$$Z(x, C) = Ce^x + \int_0^x e^{x-t}(6t + 5)\, dt = -6x - 11 + (C + 11)e^x.$$

Общее решение исходного уравнения получается двойным интегрированием:

$$\begin{aligned} Y(x, C_1, C_2, C_3) &= \int_0^x [(x - t)(-6t - 11 + (C_1 + 11)e^t)]\, dt + C_2 + C_3 x \\ &= -x^3 - 5.5x^2 + (C_1 + 11)e^x - (C_1 + 11)(x + 1) + C_2 + C_3 x. \end{aligned}$$

Введя новые константы C_1, C_2, C_3, получим более короткую запись:

$$Y(x, C_1, C_2, C_3) = C_1 e^x - x^3 - 5.5x^2 + C_2 + C_3 x.$$

Количество произвольных констант совпадает с порядком исходного уравнения.

Понижение порядка уравнения, не содержащего независимую переменную

Пусть уравнение второго порядка не содержит переменной x:

$$\phi(y, y', y'') = 0.$$

Тогда оно сводится к уравнению первого порядка.

Введем обозначение $z(x) = y'(x)$. Тогда

$$y''(x) = \frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx} = z\frac{dz}{dy}.$$

Исходное уравнение записывается в виде

$$\phi(y, z, z\frac{dz}{dy}) = 0.$$

Это уравнение первого порядка, в котором независимой переменной является y, а z есть искомая функция от y. Предположим, удалось найти общее решение $Z(y, C)$ этого уравнения. Тогда уравнение $z = Z(y, C)$ можно также записать в виде

$$y' = Z(y, C),$$

где производная берется по x. Это уравнение с разделяющимися переменными, и его общий интеграл имеет вид

$$\int \frac{dy}{Z(y, C)} = x + C_2.$$

Пример 5. Решить задачу Коши

$$y'' = 8y^3, \qquad y(0) = 1,\ y'(0) = 2.$$

Замена переменной $z = y'$ сводит уравнение к

$$z\frac{dz}{dy} = 8y^3.$$

Это уравнение с разделяющимися переменными имеет общий интеграл

$$z^2/2 = 2y^4 + C.$$

Подставим начальные данные при $x = 0$: $y(0) = 1$, $z(0) = y'(0) = 2$. Вычисления показывают, что $C = 0$. Поэтому

$$z = \pm 2y^2.$$

При $x = 0$ левая и правая части положительны, следовательно, надо выбрать знак $+$. Подставим сюда $z = y'$. Получится уравнение с разделяющимися переменными

$$y' = 2y^2.$$

Оно имеет общий интеграл

$$-\frac{1}{y} = 2x + C.$$

Подстановка начального данного $x = 0$, $y = 1$ дает $C = -1$.

В итоге получается единственное решение поставленной задачи Коши:

$$y(x) = \frac{1}{1 - 2x}.$$

Линейные дифференциальные уравнения

Определение *линейного дифференциального уравнения и оператора.*

1. Линейным дифференциальным уравнением n-го порядка называется уравнение вида

   $$y^{(n)} + p_1(x)y^{(n-1)} + \dots + p_n(x)y = f(x).$$

   Функции $p_1(x)$, ..., $p_n(x)$ называются коэффициентами уравнения, а функция $f(x)$ — правой частью уравнения.

   Будем предполагать, что все коэффициенты и правая часть — непрерывные функции от x на промежутке (a, b), на котором требуется найти решение уравнения.

2. Линейное дифференциальное уравнение называется однородным, если правая часть равна нулю, и неоднородным, если она не равна нулю.

3. Левую часть линейного дифференциального уравнения можно символически записать в виде

$$\left(\frac{d^n}{dx^n} + p_1(x)\frac{d^{n-1}}{dx^{n-1}} + \cdots + p_n(x)\right) y = f(x).$$

Выражение

$$L = \frac{d^n}{dx^n} + p_1(x)\frac{d^{n-1}}{dx^{n-1}} + \cdots + p_n(x)$$

называется линейным дифференциальным оператором. □

Оператор — это отображение, которое преобразует функцию в функцию. В данном случае, если $y(x)$ — некоторая дифференцируемая n раз функция, то $L(y)$ — тоже функция. Однородное дифференциальное уравнение записывается символически в виде

$$L(y) = 0,$$

а неоднородное уравнение —

$$L(y) = f(x).$$

Замечание о существовании и единственности решение задачи Коши. Проверим условия теорем существования и единственности решения задачи Коши. В соответствии с общей теорией, запишем уравнение в виде

$$y^{(n)} = F(x, y, y', \ldots, y^{n-1}) = -p_1(x)y^{(n-1)} - \cdots - p_n(x)y + f(x).$$

Функция

$$F(x, y_0, \ldots, y_{n-1}) = -p_1(x)y_{n-1} - \cdots - p_n(x)y_0 + f(x)$$

непрерывна по всем аргументам, так как коэффициенты $p_k(x)$ и правая часть $f(x)$ предполагаются непрерывными. Следовательно, решение задачи Коши с любыми начальными данными существует. Далее, функции $\partial F/\partial y_k = p_{n-k}(x)$ непрерывны, $0 \le k \le n-1$. Следовательно, любая задача Коши имеет единственное решение.

Свойства линейности решений однородного и неоднородного уравнений

Решения однородного уравнения можно складывать и умножать на числа, а решения неоднородного уравнения получаются как сумма решений однородного уравнения и любого решения неоднородного уравнения.

Теорема *о линейности решений однородного линейного уравнения.*

Если $y_1(x)$ и $y_2(x)$ — решения линейного однородного уравнения, то функция $C_1y_1(x) + C_2y_2(x)$ есть решение того же уравнения для любых чисел C_1 и C_2.

Доказательство.

Дифференциальный оператор k-й производной d^k/dx^k обладает свойством

$$(C_1y_1 + C_2y_2)^{(k)} = C_1y_1^{(k)} + C_2y_2^{(k)}.$$

Сумма таких операторов обладает тем же свойством. В частности,

$$L(C_1y_1 + C_2y_2) = C_1L(y_1) + C_2L(y_2).$$

Это равенство называется свойством линейности оператора L. Оператор, обладающий этим свойством, обычно называется линейным оператором. Отсюда происхождение названия L — линейный дифференциальный оператор.

Если y_1 и y_2 есть решения уравнения $L(y) = 0$, то в правой части получается нуль, следовательно, $C_1y_1 + C_2y_2$ есть решение того же уравнения. □

Множество функций, удовлетворяющих однородному линейному дифференциальному уравнению, образует линейное пространство: их можно умножать на числа и складывать в любом количестве, и вновь получится решение того же уравнения.

Теорема *о свойствах линейности решений неоднородного линейного уравнения.*

Пусть L — линейный дифференциальный оператор, $f_1(x)$ и $f_2(x)$ — непрерывные функции и $y_1(x)$, $y_2(x)$ — решения уравнений $L(y_1) = f_1$, $L(y_2) = f_2$, соответственно. Тогда функция $y_1 + y_2$ есть решение уравнения $L(y) = f_1 + f_2$.

Доказательство. По свойству линейности дифференциального опрератора $L(y_1+y_2) = L(y_1) + L(y_2) = f_1 + f_2$, что и требовалось доказать. □

Линейная независимость решений. Вронскиан

Напомним, что линейной комбинацией системы функций $y_1(x)$, …, $y_m(x)$ называется любая функция вида $C_1y_1(x) + \cdots + C_my_m(x)$ с числовыми коэффициентами $C_1 \in \mathsf{R}$, …, $C_m \in \mathsf{R}$.

Определение *линейной независимости системы функций.*

Система функций $y_1(x)$, $y_2(x)$, …, $y_n(x)$, заданных на одном и том же промежутке (a, b), называется *линейно независимой*, если ни одна функция из этого набора не есть линейная комбинация остальных функций. В противном случае система функций называется *линейно зависимой.* □

Лемма *о линейно зависимых функциях.*

Для того, чтобы система функций $y_1(x)$, $y_2(x)$, …, $y_n(x)$, была линейно зависимой, необходимо и достаточно, чтобы некоторая линейная комбинация этих функций, среди коэффициентов в которой есть ненулевые, была тождественно равна нулю.

Другая формулировка того же утверждения: для того, чтобы система функций была линейно независима, необходимо и достаточно, чтобы из тождественного равенства $C_1y_1 + C_2y_2 + \ldots + C_ny_n \equiv 0$ следовало, что все коэффициенты C_1, C_2, …, C_n равны нулю. □

Доказательство.

Пусть $C_1y_1 + C_2y_2 + \ldots + C_ny_n = 0$ и среди коэффициентов есть ненулевые. Пусть $C_k \neq 0$. Тогда

$$y_k = -\frac{1}{C_k}(C_1y_1 + \cdots + C_{k-1}y_{k-1} + C_{k+1}y_{k+1} + C_ny_n).$$

По определению, система функций линейно зависима. Обратное утверждение очевидно. □

Определение *вронскиана.*

Пусть функции $y_1(x)$, $y_2(x)$, ..., $y_n(x)$ имеют $n-1$ производную. Тогда определитель

$$W = \begin{vmatrix} y_1 & y_2 & \cdots & y_n \\ y_1' & y_2' & \cdots & y_n' \\ \ldots & \ldots & \cdots & \ldots \\ y_1^{(n-1)} & y_2^{(n-1)} & \cdots & y_n^{(n-1)} \end{vmatrix}$$

называется *определителем Вронского* (или *вронскианом*).

Вронскиан есть функция от x и поэтому обычно обозначается $W(x)$. В частности, при $n = 2$ получаем

$$W(x) = y_1(x)y_2'(x) - y_2(x)y_1'(x).$$

Лемма.

Если несколько функций линейно зависимы, то их определитель Вронского тождественно равен нулю.

Доказательство.

Действительно, если одна из функций есть линейная комбинация остальных, то в определителе Вронского соответствующий столбец есть линейная комбинация остальных столбцов. А определитель, у которого один столбец есть линейная комбинация других столбцов, равен нулю. □

Существует много систем линейно независимых функций, у которых определитель Вронского всюду отличен от нуля, за исключением нескольких значений аргумента x (корней функции $W(x)$). Однако для решений линейного дифференциального уравнения, как показывает следующая теорема, такого быть не может: либо определитель Вронского тождественно равен нулю, либо у него вообще нет корней.

Теорема. *Вронскиан и линейная независимость решений однородного уравнения.*

Если n решений линейного однородного дифференциального уравнения n-го порядка линейно независимы, то их определитель Вронского $W(x)$ не имеет корней.

Доказательство.

Проведем его от противного. Пусть определитель Вронского $W(x)$ имеет корень в некоторой точке $x = x_0$, т.е. $W(x_0) = 0$. Рассмотрим систему алгебраических уравнений

$$\begin{cases} C_1y_1(x_0) & + \ldots + & C_ny_n(x_0) & = & 0, \\ C_1y_1'(x_0) & + \ldots + & C_ny_n'(x_0) & = & 0, \\ \cdots & & \cdots & & \cdots \\ C_1y_1^{(n-1)}(x_0) & + \ldots + & C_ny_n^{(n-1)}(x_0) & = & 0 \end{cases}$$

относительно неизвестных чисел C_1, C_2, ..., C_n. Нетрудно видеть, что определитель этой системы совпадает с определителем Вронского $W(x_0)$ и поэтому равен нулю.

По теореме из курса алгебры система из n однородных линейных уравнений относительно n неизвестных с нулевым определителем всегда имеет ненулевое решение. Пусть $\widetilde{C}_1$, $\widetilde{C}_2$, ..., $\widetilde{C}_n$ — это решение. Среди указанных чисел есть ненулевые. Рассмотрим функцию

$$\widetilde{y}(x) = \widetilde{C}_1y_1(x) + \widetilde{C}_2y_2(x) + \ldots + \widetilde{C}_ny_n(x).$$

Из системы алгебраических уравнений следует, что

$$\widetilde{y}(x_0) = \widetilde{y}'(x_0) = \dots = \widetilde{y}^{(n-1)}(x_0) = 0.$$

Таким образом, функция $\widetilde{y}(x)$ решает задачу Коши для исходного дифференциального уравнения с нулевыми начальными данными. Но эту же задачу решает и функция $y(x) \equiv 0$. По теореме единственности отсюда следует, что эти функции совпадают, т.е.

$$\widetilde{C}_1 y_1(x) + \widetilde{C}_2 y_2(x) + \ldots + \widetilde{C}_n y_n(x) \equiv 0.$$

По лемме о линейно зависимых функциях система функций $y_1(x)$, $y_2(x)$, …, $y_n(x)$ линейно зависима, что противоречит условию. □

Общее и частное решения линейного дифференциального уравнения

Пусть L – линейный дифференциальный оператор n-го порядка. Рассмотрим линейное однородное дифференциальное уравнение $L(y) = 0$.

Определение *фундаментальной системы решений.*

Система из n решений линейного однородного уравнения называется *фундаментальной*, если она линейно независима. □

Теорема *об общем решении линейного однородного уравнения.*

1. Фундаментальные системы решений уравнения $L(y) = 0$ существуют.

2. Для любой фундаментальной системы решений y_1, y_2, …, y_n функция

$$Y(x, C_1, C_2, \ldots, C_n) = C_1 y_1(x) + C_2 y_2(x) + \ldots + C_n y_n(x)$$

образует общее решение уравнения $L(y) = 0$.

Доказательство.

Докажем утверждение 1. Выберем произвольным образом n столбцов чисел $A_j = (a_{ij})_{i=1}^n$, $1 \le j \le n$, так, чтобы определитель матрицы $A = (A_1, A_2, \ldots, A_n)$ был ненулевым. (Достаточно выбрать любую матрицу с ненулевым определителем и взять n ее столбцов.) Для каждого столбца A_j поставим задачу Коши для уравнения $L(y) = 0$ с начальными данными

$$y(x_0) = a_{1j}, \quad y'(x_0) = a_{2j}, \quad \ldots, \quad y^{(n-1)}(x_0) = a_{nj},$$

где x_0 — какая-нибудь точка из рассматриваемого интервала. Эта задача Коши имеет единственное решение, которое обозначим $y_j(x)$.

Рассмотрим построенную систему функций $y_1(x), y_2(x), \ldots, y_n(x)$. По построению при $x = x_0$ определитель Вронского этой системы функций совпадает с определителем матрицы A и поэтому не равен нулю. По теореме о вронскиане данная система решений линейно независима и поэтому является фундаментальной, что доказывает утверждение 1.

Докажем утверждение 2. Пусть $(y_1, y_2, \ldots, y_n)$ — какая–либо фундаментальная система решений. В частности, в любой точке интервала, на котором она задана, определитель Вронского отличен от нуля. Пусть $\bar{y}(x)$ — некоторое другое решение уравнения $L(y) = 0$. Требуется доказать, что функция $\bar{y}(x)$ раскладывается в линейную комбинацию функций $(y_1, y_2, \ldots, y_n)$.

Выберем произвольную точку x_0 из области определения решений. Система из n алгебраических уравнений

$$\begin{cases} C_1 y_1(x_0) & +\ldots+ & C_n y_n(x_0) & = & \bar{y}(x_0), \\ C_1 y_1'(x_0) & +\ldots+ & C_n y_n'(x_0) & = & \bar{y}'(x_0), \\ \cdots & & \cdots & & \cdots \\ C_1 y_1^{(n-1)}(x_0) & +\ldots+ & C_n y_n^{(n-1)}(x_0) & = & \bar{y}^{(n-1)}(x_0) \end{cases}$$

с n неизвестными $C_1, C_2, \ldots, C_n$ имеет единственное решение, поскольку ее определитель есть определитель Вронского. Выберем это решение и определим новую функцию

$$\widetilde{y}(x) = C_1 y_1(x) + C_2 y_2(x) + \ldots + C_n y_n(x).$$

Дифференцируя эту функцию, получим, что

$$\begin{aligned} \widetilde{y}(x_0) &= \bar{y}(x_0), \\ \widetilde{y}'(x_0) &= \bar{y}'(x_0), \\ \cdots & \quad \cdots \\ \widetilde{y}^{(n-1)}(x_0) &= \bar{y}^{(n-1)}(x_0). \end{aligned}$$

Две функции — $\bar{y}(x)$ и $\widetilde{y}(x)$ решают одну и ту же задачу Коши с начальными данными в точке x_0. По теореме единственности эти функции равны. Следовательно,

$$\bar{y}(x) = C_1 y_1(x) + C_2 y_2(x) + \ldots + C_n y_n(x),$$

т.е. функция $\bar{y}$ раскладывается в линейную комбинацию функций из фундаментальной системы решений, что и требовалось доказать. □

Рассмотрим линейное неоднородное дифференциальное уравнение $L(y) = f(x)$. *Частным решением* этого уравнения называют любое конкретное решение этого уравнения, т.е. любую функцию, удовлетворяющую уравнению. Частное решение не содержит констант C_i. Одновременно с данным неоднородным уравнением будем рассматривать соответствующее ему линейное однородное уравнение $L(y) = 0$.

Теорема. *Общее решение линейного неоднородного уравнения.*

Пусть $y_{\text{ч}}(x)$ — некоторое частное решение неоднородного уравнения $L(y) = f(x)$ и пусть функция $Y_{\text{O}}(x, C_1, \ldots, C_n)$ есть общее решение однородного уравнения $L(y) = 0$. Тогда функция

$$Y_{\text{H}}(x, C_1, \ldots, C_n) = y_{\text{ч}}(x) + Y_{\text{O}}(x, C_1, \ldots, C_n)$$

есть общее решение линейного неоднородного уравнения $L(y) = f(x)$.

Доказательство.

Пусть $\bar{y}(x)$ — некоторое решение уравнения $L(y) = f(x)$. Тогда

$$L(\bar{y} - y_{\text{ч}}) = L(\bar{y}) - L(y_{\text{ч}}) = 0,$$

поэтому функция $\bar{y} - y_{\text{ч}}$ является решением однородного уравнения $L(y) = 0$. А любое решение последнего уравнения представляется его общим решением $Y_{\text{O}}(x, C_1, \ldots, C_n)$ с некоторым набором констант. □

Нахождение частного решения. Метод вариации произвольных постоянных

Рассмотрим линейное неоднородное дифференциальное уравнение $L(y) = f(x)$. Пусть найдена фундаментальная система решений однородного уравнения $y_1, y_2, \ldots, y_n$. Ей соответствует определитель Вронского $W(x)$, не равный нулю при всех x. Общее решение однородного уравнения есть

$$C_1 y_1(x) + C_2 y_2(x) + \cdots + C_n y_n(x).$$

Будем искать частное решение неоднородного уравнения в той же форме, но вместо констант C_i подставим пока неопределенные функции $u_i(x)$:

$$y(x) = u_1(x) y_1(x) + u_2(x) y_2(x) + \cdots + u_n(x) y_n(x).$$

Функции u_1, u_2, …, u_n определим из условий

$$\begin{cases} u_1' y_1 & + \ldots + & u_n' y_n & = & 0, \\ u_1' y_1' & + \ldots + & u_n' y_n' & = & 0, \\ \cdots & & \cdots & & \cdots \\ u_1' y_1^{(n-2)} & + \ldots + & u_n' y_n^{(n-2)} & = & 0, \\ u_1' y_1^{(n-1)} & + \ldots + & u_n' y_n^{(n-1)} & = & f(x). \end{cases}$$

Докажем, что так построенное решение удовлетворяет неоднородному уравнению. Действительно, по индукции дифференцированием можно доказать, что

$$\begin{cases} y & = & u_1 y_1 & + \ldots + & u_n y_n, \\ y' & = & u_1 y_1' & + \ldots + & u_n y_n', \\ \cdots & & \cdots & & \cdots \\ y^{(n-1)} & = & u_1 y_1^{(n-1)} & + \ldots + & u_n y_n^{(n-1)}, \\ y^{(n)} & = & u_1 y_1^{(n)} & + \ldots + & u_n y_n^{(n)} + f(x). \end{cases}$$

Складывая эти уравнения с коэффициентами $p_n(x)$, $p_{n-1}(x)$, …, $p_1(x)$, 1, получим уравнение

$$L(y) = u_1 L(y_1) + u_2 L(y_2) + \cdots + u_n L(y_n) + f(x).$$

Но y_1, …, y_n — решения уравнения $L(y) = 0$. Поэтому функция y решает уравнение $L(y) = f(x)$, что и требовалось доказать.

Рассмотрим систему уравнений относительно $u_1', u_2', \ldots, u_1'$. Это линейная система, определитель которой есть определитель Вронского $W(x)$, который не равен нулю.

Воспользуемся формулами Крамера. Для каждого $k = 1, 2, \ldots, n$ обозначим через $W_k(x)$ определитель Вронского порядка $n - 1$, построенный по системе функций y_1, …, y_{k-1}, y_{k+1}, …, y_n, Тогда по теореме Крамера

$$u_k'(x) = (-1)^{k+n} \frac{f(x) W_k(x)}{W(x)}, \qquad k = 1, 2, \ldots, n.$$

Интегрируя эти уравнения, получим

$$u_k(x) = (-1)^{k+n} \int \frac{W_k(x)}{W(x)} f(x)\, dx, \qquad k = 1, 2, \ldots, n.$$

Общее решение неоднородного уравнения записывается в виде

$$Y_{\text{н}}(x, C_1, \ldots, C_n) = (C_1 + u_1(x))y_1(x) + (C_2 + u_2(x))y_2(x) + \ldots + (C_n + u_n(x))y_n(x).$$

Пример. Рассмотрим одно уравнение. Пусть $n = 1$ и

$$y' = p(x)y + f(x).$$

Однородное уравнение $y' = p(x)y$ имеет решение

$$y_1(x) = e^{\int p(x)\,dx}.$$

Поскольку $n = 1$, то это и есть фундаментальная система решений.

Применим метод вариации произвольных постоянных. В данном случае $W_1(x) = 1$, $W(x) = y_1(x)$. Следовательно,

$$u_1(x) = \int \frac{f(x)}{y_1(x)}\,dx.$$

Окончательно, общее решение неоднородного уравнения записывается в виде

$$Y(x, C) = e^{\int p(x)\,dx}\left(C + \int f(x)e^{-\int p(x)\,dx}\,dx\right).$$

Линейные уравнения с постоянными коэффициентами

Если коэффициенты линейного дифференциального уравнения есть числа, а не функции, то решение можно непосредственно вычислить.

В этом разделе рассматривается дифференциальное уравнение вида

$$y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1}y' + a_n y = f(x),$$

в котором a_1, a_2, …, a_n — числа, а f — непрерывная функция.

Определение *характеристического многочлена и уравнения.*

Характеристическим многочленом линейного дифференциального уравнения с постоянными коэффициентами называется многочлен

$$\chi(z) = z^n + a_1 z^{n-1} + \cdots + a_{n-1}z + a_n,$$

а *характеристическим уравнением* называется уравнение

$$\chi(z) = 0.$$

□

Линейный дифференциальный оператор L формально получается, если в характеристический многочлен вместо z подставить оператор дифференцирования d/dt:

$$L(y) = \left(\frac{d^n}{dx^n} + a_1\frac{d^{n-1}}{dx^{n-1}} + \ldots + a_{n-1}\frac{d}{dx} + a_n\right)y = \chi\left(\frac{d}{dx}\right)y.$$

Легко вычисляется значения дифференциального оператора на функции $y(x) = e^{zx}$, где z — некоторое фиксированное число, вещественное или комплексное. Действительно, для такой функции $y' = ze^{zx} = zy$, $y'' = z^2e^{zx} = z^2y$, и вообще, при всех $k \geq 1$: $y^{(k)} = z^k y$.

Лемма.

Пусть комплексное (в частном случае, вещественное) число z есть корень характеристического уравнения $\chi(z) = 0$. Тогда функция $y(x) = e^{zx}$ есть решение однородного уравнения $L(y) = 0$.

Доказательство.

Пусть $y(x) = e^{zx}$. Тогда

$$\begin{aligned} L(y) &= \left(\frac{d^n}{dx^n} + a_1\frac{d^{n-1}}{dx^{n-1}} + \ldots + a_{n-1}\frac{d}{dx} + a_n\right)e^{zx} \\ &= (z^n + a_1z^{n-1} + \ldots + a_{n-1}z + a_n)e^{zx} = \chi(z)e^{zx} = 0. \end{aligned}$$

□

Теорема. *Построение решения по корню характеристического уравнения.*

Пусть комплексное число α есть корень характеристического многочлена $\chi(z)$ кратности k и $P(x)$ — произвольный многочлен степени не выше $k-1$. Тогда функция $y(x) = P(x)e^{\alpha x}$ является решением линейного однородного уравнения $L(y) = 0$.

Доказательство.

Пусть $y(x) = P(x)e^{\alpha x}$ и степень $P(x)$ не выше $k-1$. Рассмотрим функцию

$$y_1(x) = \left(\frac{d}{dx} - \alpha\right)y(x).$$

По правилу дифференцирования произведения

$$y_1(x) = y'(x) - \alpha y(x) = P'(x)e^{\alpha x} + P(x)\alpha e^{\alpha x} - \alpha P(x)e^{\alpha x} = P'(x)e^{\alpha x}.$$

Применяя оператор $d/dt - \alpha$ еще раз, получим

$$y_2(x) = \left(\frac{d}{dx} - \alpha\right)^2 y(x) = \left(\frac{d}{dx} - \alpha\right)y_1(x) = P''(x)e^{\alpha x}.$$

По индукции легко доказать, что

$$\left(\frac{d}{dx} - \alpha\right)^m y(x) = P^{(m)}(x)e^{\alpha x}$$

при всех $m \geq 0$. Но $P(x)$ — многочлен, а каждое дифференцирование понижает ровно на 1 его степень. Следовательно, $P^{(k-1)}(x)$ есть константа, а $P^{(k)}(x) = 0$.

По условию, число α является корнем характеристического многочлена $\chi(z)$ кратности k. По определению кратности корня это означает, что $\chi(z) = r(z)(z-\alpha)^k$, где $r(z)$ — некоторый многочлен. Учитывая, что оператор L может быть записан как $\chi(d/dt)$, выполним преобразования:

$$\begin{aligned} L(y) &= \chi\left(\frac{d}{dx}\right)y(x) = \left[r\left(\frac{d}{dx}\right)\left(\frac{d}{dx} - \alpha\right)^k\right]y(x) \\ &= r\left(\frac{d}{dx}\right)\left[\left(\frac{d}{dx} - \alpha\right)^k P(x)e^{\alpha x}\right] = r\left(\frac{d}{dx}\right)P^{(k)}(x)e^{\alpha x} = 0. \end{aligned}$$

□

Следствие. *Общее решение уравнения с постоянными коэффициентами.*

Пусть α_1, α_2, …, α_m — все корни характеристического многочлена $\chi(z)$, и они имеют кратности k_1, k_2, …, k_m, соответственно. Тогда функция

$$Y = P_1(x)e^{\alpha_1 x} + P_2(x)e^{\alpha_2 x} + \cdots + P_m(x)e^{\alpha_m x},$$

в которой при каждом $i = 1, 2 \ldots, m$ степень многочлена $P_i(x)$ не выше $k_i - 1$, есть общее решение однородного уравнения $L(y) = 0$. Произвольными константами в этом решении являются коэффициенты многочленов $P_i(x)$, $1 \le i \le m$.

Доказательство.

Из основной теоремы высшей алгебры следует, что сумма кратностей степеней всех корней $k_1 + k_2 + \ldots + k_m$ равна степени многочлена $\chi(z)$, т.е. n. Каждый многочлен $P_i(x)$ есть сумма степеней x с произвольными коэффициентами:

$$P_i(x) = P_{i0}x^{k_i-1} + P_{i1}x^{k_i-2} + \cdots + P_{i,k_i-2}x + P_{i,k_i-1}.$$

Каждая из функций $x^j e^{\alpha_i x}$ при $0 \le j \le k_i - 1$, $1 \le i \le m$, есть решение уравнения $L(y) = 0$ по предыдущей теореме.

Следовательно, в выражении Y из заключения данного следствия стоит сумма ровно n решений уравнения $L(y) = 0$. По теореме об общем решении линейного однородного уравнения заключение следствия верно, если функции $x^j e^{\alpha_i x}$, где $0 \le j \le k_i - 1$, $1 \le i \le m$, образуют фундаментальную систему, т.е. набор этих функций линейно независим.

Докажем линейную независимость этого набора функций. Если $m = 1$, то это очевидно, так как многочлен с ненулевыми коэффициентами не может быть тождественно равен нулю. Далее проведем индукцию по m. Пусть для $m - 1$ утверждение доказано.

Докажем его для m методом от противного. Предположим, что функция $x^{k_l-1}e^{\alpha_l x}$ есть линейная комбинация всех остальных:

$$x^{k_l-1}e^{\alpha_l x} = \sum_{i=1}^{m}\sum_{j=0}^{k_i-1} C_{ij}x^j e^{\alpha_i x},$$

где C_{ij} — некоторые коэффициенты. Для каждого $i \ne l$ хотя бы один из коэффициентов C_{ij} не равен нулю, так как иначе получается линейная комбинация из $m - 1$ функции, которая не существует по индукционному предположению.

Разделим последнее равенство на $e^{\alpha_{i_0} x}$ и продифференцируем k_l раз:

$$[x^{k_l-1}]^{(k_l)} = \left[\sum_{i=1}^{m}\sum_{j=0}^{k_i-1} C_{ij}x^j e^{(\alpha_i - \alpha_{i_0})x}\right]^{(k_l)}.$$

Но слева получается ноль, а справа — сумма произведений функций $e^{(\alpha_i - \alpha_l)x}$ на многочлены степени не выше $k_i - 1$. При $i = l$ после вычисления производной порядка k_l получится ноль. Поэтому справа стоит сумма не более, чем $m - 1$ экспоненты с полиномиальными множителями. Это противоречит индукционному предположению, о том, что такая система функций линейно независима. □

Частные решения при специальных правых частях

Рассмотрим линейное неоднородное уравнение с постоянными коэффициентами

$$y^{(n)} + a_1 y^{(n-1)} + \cdots + a_{n-1} y' + a_n y = P(x)e^{\gamma x},$$

где $P(x)$ — многочлен степени m, а γ — некоторое число, вещественное или комплексное.

По теореме об общем решении линейного неоднородного уравнения, это решение складывается из общего решения линейного однородного уравнения и частного решения данного уравнения. Первая часть описана в предыдущем разделе. Остается определить какое–либо частное решение данного неоднородного уравнения.

Теорема *о частном решении неоднородного уравнения.*

Пусть число γ является корнем характеристического многочлена $\chi(z)$ кратности k. Тогда существует частное решение линейного неоднородного уравнения вида

$$y(x) = x^k M(x) e^{\gamma x},$$

где $M(x)$ — многочлен степени не выше m.

В частности, если γ не является корнем $\chi(z)$, то $k = 0$ и частное решение имеет вид $y(x) = M(x)e^{\gamma x}$.

Доказательство.

Пусть $f(x) = P(x)e^{\gamma x}$ — правая часть уравнения. По правилу вычисления производных высших порядков от произведения функций:

$$L(x^i e^{\gamma x}) = e^{\gamma x} \sum_{j=0}^{i} C_i^j x^{i-j} \chi^{(j)}(\gamma).$$

По условию число γ есть корень многочлена $\chi(z)$ кратности k, поэтому $\chi^{(j)}(\gamma) = 0$ при $0 \le j \le k-1$ и суммирование можно ограничить пределами $k \le j \le i$.

Подставим в предыдущее уравнение $i = k + l$ при $0 \le l \le m$:

$$L(x^{k+l} e^{\gamma x}) = e^{\gamma x} \sum_{j=k}^{k+l} C_{k+l}^j x^{k+l-j} \chi^{(j)}(\gamma) = e^{\gamma x} \sum_{j=0}^{l} C_{k+l}^{k+j} x^{l-j} \chi^{(k+j)}(\gamma) = e^{\gamma x} p_l(x),$$

где $p_l(x)$ — многочлен степени l.

Найдем такие многочлены $p_l(x)$ для каждого $l = 0, 1, \ldots, m$. Многочлен $P(x)$ степени m однозначно раскладывается по ним:

$$P(x) = \sum_{l=0}^{m} c_l p_l(x),$$

где c_0, c_1, …, c_m — некоторые числа. Умножим предыдущее уравнение на c_l и просуммируем по l:

$$\begin{aligned} L(x^k \sum_{l=0}^{m} c_l x^l e^{\gamma x}) &= \sum_{l=0}^{m} c_l L(x^k x^l e^{\gamma x}) \\ = \quad & e^{\gamma x} \sum_{l=0}^{m} c_l p_l(x) = e^{\gamma x} P(x) = f(x). \end{aligned}$$

Следовательно, функция $y(x) = x^k M(x) e^{\gamma x}$ является частным решением линейного неоднородного уравнения при

$$M(x) = \sum_{l=0}^{m} c_l x^l.$$

□

"Угадав"общий вид частного решения, можно найти недостающие параметры — коэффициенты многочлена $M(x)$ — подстановкой этого решения в уравнение и приравниванием коэффициентов при одинаковых функциях.

Особое значение имеют конкретные значения числа γ, лежащие на мнимой оси. Если $i = \sqrt{-1}$ — мнимая единица, то по определению, $e^{i\alpha x} = \cos\alpha x + i \sin\alpha x$. По этой формуле удается найти частные решения, если в правой части присутствуют тригонометрические функции.

Следствие 1. *Полиномиальная правая часть.*

Уравнение

$$y^{(n)} + a_1 y^{(n-1)} + \dots + a_{n-1} y' + a_n y = P(x),$$

где $P(x)$ — многочлен степени k и $a_n \neq 0$, имеет частным решением многочлен $y(x) = M(x)$ степени не выше k. □

Следствие 1'. *Полиномиальная правая часть при нулевом корне.*

Уравнение

$$y^{(n)} + a_1 y^{(n-1)} + \dots + a_{n-1} y' + a_n y = P(x),$$

где $P(x)$ — многочлен степени k, $a_n = a_{n-1} = \dots = a_{n-j+1} = 0$ и $a_{n-j} \neq 0$, имеет частным решением многочлен $y(x) = x^j M(x)$, где степень $M(x)$ не выше k. □

Доказательство.

Следствия 1 и 1' получаются из теоремы в частном случае $\gamma = 0$.

Следствие 2. *Тригонометрические функции в правой части.*

Уравнение

$$y^{(n)} + a_1 y^{(n-1)} + \dots + a_{n-1} y' + a_n y = Q(x) \cos\alpha x + R(x) \sin\alpha x,$$

где $Q(x)$ и $R(x)$ — многочлены степени не выше k, а чисто мнимое число $i\alpha$ не является корнем характеристического уравнения, имеет частным решением функцию

$$y(x) = M(x) \cos\alpha x + N(x) \sin\alpha x,$$

где $M(x)$ и $N(x)$ — многочлены степени не выше k. □

Следствие 2'. *Тригонометрические функции в правой части при чисто мнимом корне.*

Уравнение

$$y^{(n)} + a_1 y^{(n-1)} + \dots + a_{n-1} y' + a_n y = Q(x) \cos\alpha x + R(x) \sin\alpha x,$$

где $Q(x)$ и $R(x)$ — многочлены степени не выше k, а чисто мнимое число $i\alpha$ является корнем характеристического уравнения кратности j, имеет частным решением функцию

$$y(x) = x^j (M(x) \cos\alpha x + N(x) \sin\alpha x),$$

где $M(x)$ и $N(x)$ — многочлены степени не выше k. $\square$

Доказательство следствий 2 и 2'.

Подставим в условия теоремы значение $\gamma = i\alpha$ и $P(x) = Q(x) - iR(x)$. Тогда правая часть будет равна

$$P(x)e^{i\alpha x} = (Q(x)\cos\alpha x + R(x)\sin\alpha x) + i(Q(x)\sin\alpha x - R(x)\cos\alpha x) = f_1(x) + if_2(x),$$

где f_1 и f_2 — вещественные функции, а $f_1(x)$ совпадает с правой частью исходного дифференциального уравнения.

Согласно теореме, частным решением будет

$$y(x) = x^j H(x)(\cos\alpha x + i\sin\alpha x),$$

но коэффициенты многочлена $H(x)$ комплексные. Другими словами, $H(x) = M(x) - iN(x)$, где многочлены M и N имеют уже вещественные коэффициенты и по–прежнему их степени не выше k.

После умножения и разделения вещественной и мнимой частей, получим

$$y(x) = x^j(M(x)\cos\alpha x + N(x)\sin\alpha x) + ix^j(M(x)\sin\alpha x - N(x)\cos\alpha x) = y_1(x) + iy_2(x),$$

где функции $y_1(x)$ и $y_2(x)$ вещественные.

Приравняем вещественные части в уравнении $L(y) = f$. Получится искомое равенство: $L(y_1) = f_1(x)$. $\square$

Пример. Найти общее решение линейного неоднородного уравнения

$$y'' - 2y' + y = x^2 + e^x + \cos 3x.$$

Сначала найдем общее решение соответствующего линейного однородного уравнения. Характеристический многочлен получается заменой производной d/dx на независимую переменную z:

$$\chi(z) = z^2 - 2z + 1 = (z-1)^2.$$

Поэтому общим решением однородного уравнения будет

$$Y_{\text{однородн}}(x, C_1, C_2) = C_1 x e^x + C_2 e^x.$$

Найдем частное решение неоднородного уравнения. В соответствии с теоремой, оно может быть взято в виде

$$y_{\text{частн}}(x) = Ax^2 + Bx + C + Dx^2e^x + E\cos 3x + F\sin 3x.$$

Действительно, число $z = 0$ — не корень $\chi(z)$, поэтому степень многочлена в $y_{\text{частн}}(x)$ не больше 2 по следствию 1. Далее, функция e^x в правой части удовлетворяет теореме при $\gamma = 1$. Но это корень $\chi(z)$ кратности 2, поэтому решение De^x умножается еще на x^2. Наконец, последние два слагаемых в $y_{\text{частн}}(x)$ получаются из следствия 2.

Подставим функцию $y_{\text{частн}}(x)$ в дифференциальное уравнение и приравняем коэффициенты при одинаковых функциях. Получится система линейных алгебраических уравнений, из которой все эти коэффициенты можно найти.

Окончательно, общее решение есть

$$Y(x, C_1, C_2) = C_1xe^x + C_2e^x + x^2 + 4x + 6 + \frac{1}{2}x^2e^x + \frac{4}{35}\cos 3x - \frac{3}{35}\sin 3x.$$

Ряды

Определение.

Пусть $(a_i)_{i=1}^{\infty}$ — последовательность вещественных чисел. Тогда выражение

$$a_1 + a_2 + \ldots + a_n + \ldots = \sum_{i=1}^{\infty} a_i$$

называется *числовым рядом.*

Для каждого натурального числа n конечная сумма

$$S_n = \sum_{i=1}^{n} a_i$$

называется *частичной суммой* этого ряда. □

Каждый элемент a_i обычно называют членом ряда. Частичные суммы полностью характеризуют ряд, так как $a_i = S_i - S_{i-1}$. Последовательность частичных сумм $(S_n)_{n=1}^{\infty}$ может иметь или не иметь предел при $n \to \infty$.

Определение.

Ряд называется *сходящимся*, если последовательность его частичных сумм имеет предел. В этом случае предел

$$S = \lim_{n\to\infty} S_n = \sum_{i=1}^{\infty} a_i$$

называется *суммой ряда* и обозначается так же, как и ряд.

Если последовательность частичных сумм не имеет предела, то ряд называется расходящимся. □

Пример. *Геометрическая прогрессия.*

$a_i = a_0 q^i$. Если $|q| < 1$, то ряд сходится, и его сумма равна $S = a_0 q/(1-q)$. Если $|q| \geq 1$, то ряд расходится. При этом если $|q| > 1$, то последовательность частичных сумм

$$S_n = a_0 q \frac{q^n - 1}{q - 1}$$

неограничена. Если $q = 1$, то $S_n = a_0 n$, и это тоже неограниченная последовательность. Если $q = -1$, то $S_n = -a_0$ при нечетном n, и $S_n = 0$ при четном n. В последнем случае последовательность (S_n) ограничена, но все–таки не имеет предела.

Теорема *о свойствах числовых рядов.*

1. Если ряд $\sum a_n$ сходится и имеет сумму S, то для любого числа c ряд $\sum c a_n$ также сходится, и он имеет сумму cS.

2. Если ряды $\sum a_n$ и $\sum b_n$ сходятся и имеют суммы S и R, соответственно, то ряд, получающийся почленным сложением, $\sum(a_n + b_n)$, также сходится и имеет сумму $Q = S + R$.

3. Если в каком–либо ряде изменить конечное число членов, то его сходимость или расходимость не изменится.

Доказательство.

Утверждения 1 и 2 непосредственно следуют из свойств линейности предела, примененных для последовательности частичных сумм.

Докажем утверждение 3. Пусть конечное число членов последовательности (a_n) изменено. Среди измененных членов выберем тот, у которого максимальный индекс. Пусть это индекс N. Обозначим частичные суммы до изменения через S_n, а после изменения через S'_n. Тогда разности

$$R_n = S_n - S_N = S'_n - S'_N = \sum_{i=N+1}^{n} a_i, \qquad n > N,$$

не зависят от проведенных изменений. Если последовательность (R_n) имеет предел, то последовательности (S_n) и (S'_n) также имеют пределы, так как отличаются от нее на константы S_N и S'_N, соответственно. Если же (R_n) не имеет предела, то оба ряда расходятся. □

При анализе сходимости ряда часто бывает удобен критерий Коши сходимости последовательности. А именно, последовательность (S_n) сходится тогда и только тогда, когда она сходится в себе. Последнее означает, что для любого $\varepsilon > 0$ существует такое число $N > 0$, что если $n_1 > N$ и $n_2 > 0$, то $|S_{n_1} - S_{n_2}| < \varepsilon$. В применении к последовательностям получается следующий результат.

Теорема. *Критерий Коши для числовых рядов.*

Для того, чтобы ряд $\sum_{n=1}^{\infty} a_n$ сходился, необходимо и достаточно, чтобы что для любого $\varepsilon > 0$ существовало такое число $N > 0$, что если $n > N$ и $p > 0$, то $|a_n + a_{n+1} + \ldots + a_{n+p}| < \varepsilon$. □

Иногда удается доказать расходимость числового ряда, если оказываются нарушеннными некоторые необходимые условия его сходимости.

Теорема. *Необходимые условия сходимости числовых рядов.*

1. Если ряд сходится, то его члены стремятся к нулю.

2. Если ряд сходится, то его частичные суммы ограничены.

Доказательство.

Утверждение 1 следует из предельного перехода

$$a_n = S_n - S_{n-1} \to S - S = 0.$$

Утверждение 2 следует из ограниченности любой сходящейся последовательности. □

Примеры.

1. Ряд $\sum(-1)^n$ расходится, так как его общий член $(-1)^n$ не стремится к нулю при $n \to \infty$.

2. Если $a_n \to 0$, то это еще не значит, что ряд $\sum a_n$ сходится. Наример, ряд

$$1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{3} + \frac{1}{3} + \frac{1}{3} + \cdots + \underbrace{\frac{1}{n} + \frac{1}{n} + \cdots + \frac{1}{n}}_{n \text{ слагаемых}} + \cdots$$

расходится, так как к частичным суммам добавляется по 1 через n шагов. При этом общий член ряда стремится к нулю.

Ряды с неотрицательными членами

У таких рядов следующая частичная сумма $S_{n+1} = S_n + a_{n+1}$ больше либо равна предыдущей S_n, так как $a_n \geq 0$. Получается монотонная последовательность частичных сумм, для которой сходимость установить значительно легче, чем для общей последовательности.

Теорема.
Сходимость ряда с неотрицательными членами равносильна ограниченности последовательности частичных сумм.
Доказательство.
Монотонная ограниченная последовательность всегда имеет предел. Для последовательности частичных сумм это будет сумма ряда. Если же последовательность частичных сумм неограничена, то у нее нет предела. □

Теорема. *Принцип сравнения рядов с неотрицательными членами.*
Пусть $\sum a_n$ и $\sum b_n$ — ряды с неотрицательными членами. Пусть $a_n \leq b_n$ для любого $n \geq 1$. Тогда:

1. Если $\sum b_n$ сходится, то и $\sum a_n$ сходится.
2. Если $\sum a_n$ расходится, то и $\sum b_n$ расходится.

Доказательство.
Если больший ряд $\sum b_n$ сходится, то последовательность его частичных сумм ограничены. Следовательно, ограничена и последовательность частичных сумм меньшего ряда $\sum a_n$, и по предыдущей теореме он сходится. Аналогично доказывается и второе утверждение. □

Следствие.
Пусть существуют такие положительные числа α и β, что

$$\alpha \leq \frac{a_n}{b_n} \leq \beta$$

при всех $n \geq 1$, и $\sum a_n$ и $\sum b_n$ — ряды с неотрицательными членами. Тогда эти ряды сходятся или расходятся одновременно.
Доказательство.
Достаточно заметить, что $a_n \leq \beta b_n$ и $b_n \leq \alpha^{-1} a_n$, и применить предыдущую теорему. □

Признаки сходимости и расходимости рядов с положительными членами

Сюда относятся признаки Даламбера, Коши и интегральный признак.

Теорема. *Признак сходимости Даламбера.*

Пусть $\sum a_n$ — ряд с положительными членами и

$$\frac{a_{n+1}}{a_n} \to p$$

при $n \to \infty$. Тогда ряд сходится, если $p < 1$; расходится, если $p > 1$; (а при $p = 1$ вопрос остается открытым).

Доказательство.

Пусть $p < 1$. Выберем какое-нибудь число q между p и 1, т.е. $p < q < 1$. Тогда по свойствам предела последовательности неравенство $a_{n+1}/a_n < q$ будет выполнено, начиная с некоторого номера N. Таким образом,

$$a_{N+1} < qa_N, \quad a_{N+2} < qa_{N+1} < q^2 a_N,$$

и по индукции легко доказать, что $a_{N+k} < a_N q^k$

Но ряд $\sum_{k=1}^{\infty} a_N q^k$ сходится, так как это геометрическая прогрессия и $0 < q < 1$. По теореме о сравнении рядов отсюда следует, что ряд $\sum_{k=0}^{\infty} a_{N+k}$ сходится. Но сходимость ряда не зависит от конечного набора его членов $(a_i)_{i=1}^{N}$. Следовательно, ряд $(a_i)_{i=1}^{\infty}$ также сходится.

Утверждение о расходимости ряда при $p > 1$ доказывается аналогично сведением к геометрической прогрессии со знаменателем $q > 1$. $\square$

Пример. Пусть x - некоторое положительное число. Рассмотрим ряд

$$\sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

Заметим, что если суммировать до N, а не до бесконечности, то получается многочлен Маклорена для функции e^x степени N. Докажем, что этот ряд сходится для любых $x > 0$, а его сумма, как будет доказано позже, совпадает с функцией e^x. Получается разложение экспоненты в степенной ряд.

Для доказательства сходимости ряда применим признак Даламбера:

$$\frac{a_{n+1}}{a_n} = \frac{x^{n+1}}{(n+1)!} \frac{n!}{x^n} = \frac{x}{n+1} \to 0$$

при $n \to \infty$. Ряд сходится, так как предел меньше 1.

Теорема. *Признак сходимости Коши.*

Пусть $\sum a_n$ — ряд с положительными членами и

$$\sqrt[n]{a_n} \to p$$

при $n \to \infty$. Тогда ряд сходится, если $p < 1$; расходится, если $p > 1$; (а при $p = 1$ вопрос остается открытым).

Доказательство.

Пусть $p < 1$. Выберем как в доказательстве предыдущей теоремы $q \in (p, 1)$. Тогда $\sqrt[n]{a_n} < q$, начиная с некоторого номера N, т.е. при $n > N$. Но это значит, что $a_n < q^n$, и сравнение с геометрической прогрессией снова доказывает сходимость ряда $\sum a_n$.

Аналогично доказывается расходимость ряда при $p > 1$. $\square$

Теорема. *Интегральный признак сходимости ряда.*

Пусть $\kappa(x)$ — положительная, убывающая функция, заданная на промежутке $[1, \infty)$. Тогда числовой ряд и несобственный интеграл

$$\sum_{i=1}^{\infty} \kappa(i), \qquad \int_1^{\infty} \kappa(x)\, dx$$

сходятся или расходятся одновременно.

Доказательство.

Пусть $i \geq 1$ и рассмотрим отрезок $[i, i+1]$. На этом отрезке функция $\kappa(x)$ убывает, следовательно,

$$\kappa(i) \leq \kappa(x) \leq \kappa(i+1), \qquad i \leq x \leq i+1.$$

Интегрируя эти неравенства по x на отрезке $[i, i+1]$, получим

$$\kappa(i) \leq \int_i^{i+1} \kappa(x)\, dx \leq \kappa(i+1).$$

Сложим неравенства по i от 1 до n:

$$\sum_{i=1}^{n} \kappa(i) \leq \int_1^{n+1} \kappa(x)\, dx \leq \sum_{i=2}^{n+1} \kappa(i).$$

Это основные два неравенства, которые имеют простой геометрический смысл и из которых легко доказываются утверждения теоремы.

Пусть ряд $\sum_{i=1}^{\infty} \kappa(i)$ сходится. Тогда его частичные суммы ограничены, и из правого неравенства в последней формуле видно, что интегралы по конечному промежутку тоже равномерно ограничены. Но интегрируется положительная функция, следовательно, несобственный интеграл сходится.

Если же сходится несобственный интеграл, то из левого неравенства следует, что частичные суммы ряда $\sum \kappa(i)$ ограничены. По теореме о сходимости рядов с неотрицательными членами этот ряд сходится. $\square$

Пример. Пусть $\alpha > 0$. Рассмотрим ряд

$$\sum_{n=1}^{\infty} \frac{1}{n^{\alpha}}.$$

Элементы этого ряда есть значения в целых точках функции $\kappa(x) = 1/x^{\alpha}$. Эта функция монотонно убывает. Следовательно, можно применить интегральный критерий сходимости ряда.

Соответствующий несобственный интеграл

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

анализируется легко, так как неопределенный интеграл вычисляется явно. Как известно, этот интеграл сходится при $\alpha > 1$ и расходится при $0 < \alpha \leq 1$. Отсюда получается следующий результат.

Ряд
$$\sum_{n=1}^{\infty} \frac{1}{n^\alpha}.$$
сходится при $\alpha > 1$ и расходится при $0 < \alpha \leq 1$.

Отметим, что расходящийся ряд при $\alpha = 1$ называется гармоническим. Он также записывается в виде
$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \dots + \frac{1}{n} + \cdots$$

Знакочередующиеся числовые ряды

Ряд называется знакочередующимся, если знаки его членов чередуются, т.е. четные члены имеют один знак, а нечетные — противоположный. Если нечетные члены имеют знак $+$, то ряд можно записать в виде
$$a_1 - a_2 + a_3 - a_4 + \dots,$$
где $a_i > 0$ при всех i.

Теорема. *Критерий Лейбница сходимости знакочередующихся рядов.*

Пусть у знакочередующегося ряда последовательность абсолютных величин его членов $(|a_i|)_{i=1}^{\infty}$ — убывающая и стремится к нулю, т.е. $|a_1| > |a_2| > |a_3| > \dots$, $a_n \to 0$. Тогда ряд сходится.

Доказательство.

Пусть нечетные члены положительны, а четные — отрицательные:
$$a_1 - a_2 + a_3 - a_4 + \dots,$$
где $a_i > 0$ при всех i. Обозначим частичную сумму ряда через S_n. Если n — четное число, $n = 2k$, то
$$S_{2k} = (a_1 - a_2) + (a_3 - a_4) + \ldots + (a_{2k-1} - a_{2k}).$$
Нетрудно видеть, что последовательность $(S_{2k})_{k=1}^{\infty}$ — возрастающая по k. С другой стороны,
$$S_{2k} = a_1 - (a_2 - a_3) - (a_4 - a_5) - \ldots - (a_{2k-2} - a_{2k-1}) - a_{2k} \leq a_1.$$
Таким образом, последовательность $(S_{2k})_{k=1}^{\infty}$ — возрастающая и ограниченная числом a_1. Поэтому она имеет предел, который обозначим через S. Поскольку $a_i \to 0$ по условию, то и последовательность нечетных частичных сумм имеет предел:
$$S_{2k+1} = S_{2k} + a_{2k+1} \to S$$
при $k \to \infty$. Таким образом, последовательность $(S_i)_{i=1}^{\infty}$ имеет предел S, который и называется суммой ряда по определению. □

Замечание. Пусть выполнены условия теоремы Лейбница. Тогда легко оценить скорость сходимости ряда. Последовательность частичных сумм с нечетными индексами убывает:
$$S_{2k+1} = a_1 - (a_2 - a_3) - (a_4 - a_5) - \ldots - (a_{2k} - a_{2k+1}),$$

и сходится к S. Следовательно, $S_{2k} < S < S_{2k+1}$ при всех k. Но длина промежутка $[S_{2k}, S_{2k+1}]$, содержащего S, равна a_{2k+1}. Отсюда следует следующий результат:

для знакочередующегося ряда с убывающими абсолютными величинами его членов точность приближения суммы ряда его частичной суммой меньше абсолютной величины последнего члена:

$$|S_n - S| < a_n.$$

Пример. Ряд

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{(-1)^{n+1}}{n} + \cdots$$

сходится по теореме Лейбница, тогда как ряд из абсолютных величин — гармонический, и он расходится. В дальнейшем будет доказано, что сумма этого ряда равна $\ln 2$. О таком ряде говорят, что он условно сходится.

Определение.

Ряд $\sum a_i$ называется *абсолютно сходящимся*, если ряд из абсолютных величин $\sum |a_i|$ сходится.

Ряд $\sum a_i$ называется *условно сходящимся*, если он сходится, но не абсолютно, т.е. $\sum a_i$ сходится, а $\sum |a_i|$ расходится. □

Теорема.

Абсолютно сходящийся ряд сходится.

Доказательство.

Воспользуемся критерием Коши сходимости произвольных рядов. Поскольку ряд $\sum |a_i|$ сходится, то для любого $\varepsilon > 0$ существует такое число N, что если $n > N$ и $p > 0$, то

$$|a_n| + |a_{n+1}| + \ldots + |a_{n+p}| < \varepsilon.$$

Однако в этом случае и

$$|a_n + a_{n+1} + \ldots + a_{n+p}| \leq |a_n| + |a_{n+1}| + \ldots + |a_{n+p}| < \varepsilon,$$

следовательно, по критерию Коши ряд $\sum a_i$ сходится. □

Ряды с комплексными членами

Определение.

Пусть $c_k = a_k + ib_k$ — комплексные числа при каждом $k \geq 1$. Тогда рядом $\sum c_k$ по определению называется сумма рядов

$$\sum_{k=1}^{\infty} c_k = \sum_{k=1}^{\infty} a_k + i \sum_{k=1}^{\infty} b_k.$$

Говорят, что ряд $\sum c_k$ сходится, если сходится его вещественная часть $\sum a_k$ и мнимая часть $\sum b_k$. Если из суммы, соответственно, равны S_a и S_b, то сумма комплексного ряда равна по определению числу $S_a + iS_b$.

Комплексный числовой ряд $\sum c_k$ называется *абсолютно сходящимся*, если сходится вещественный числовой ряд $\sum |c_k|$ или, что то же самое, абсолютно сходится каждый из рядов $\sum a_k$ и $\sum b_k$. □

Для комлексных числовых рядов остаются верными многие утверждения, сформулированные для вещественных числовых рядов. Это: сходимость абсолютно сходящегося ряда, признак Даламбера, признак Коши.

Сходимость абсолютно сходящегося ряда.
Пусть комплексный числовой ряд $\sum c_k$ абсолютно сходится. Тогда он сходится. □

Признак Даламбера.
Пусть $(c_k)_{k=1}^{\infty}$ — последовательность комплексных чисел и существует предел

$$\lim_{n\to\infty}\left|\frac{c_{n+1}}{c_n}\right| = p.$$

Тогда при $p < 1$ ряд $\sum c_k$ сходится, а при $p > 1$ — расходится. □

Признак Коши (радикальный).
Пусть $(c_k)_{k=1}^{\infty}$ — последовательность комплексных чисел и существует предел

$$\lim_{n\to\infty}\sqrt[n]{|c_n|} = p.$$

Тогда при $p < 1$ ряд $\sum c_k$ сходится, а при $p > 1$ — расходится. □

Доказательства последних трех утверждений повторяют доказательства для вещественных рядов. Они предлагаются слушателям в качестве упражнения.

Функциональные ряды. Равномерная сходимость

Пусть $(\phi_k(x))_{k=1}^{\infty}$ — последовательность функций, заданных на одном и том же множестве D. Тогда выражение

$$\phi_1(x) + \phi_2(x) + \ldots + \phi_k(x) + \ldots \qquad \text{или} \qquad \sum_{k=1}^{\infty}\phi_k(x)$$

называется функциональным рядом. Суммы вида

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

называются частичными суммами этого ряда.

Если последовательность частичных сумм $(S_n(x))_{n=1}^{\infty}$ имеет предел при каждом $x \in D$, то этот предел $S(x)$ называется суммой функционального ряда, а сам ряд называется сходящимся на D.

Определение *поточечной и равномерной сходимости функций.*
Пусть последовательность функций $(f_n(x))_{n=1}^{\infty}$ задана на некотором множестве $D \subseteq \mathsf{R}$.

1. Если для любого $x = x_0 \in D$ числовая последовательность $(f_n(x_0))_{n=1}^{\infty}$ имеет предел $F(x_0)$, то говорят, что последовательность функций $(f_n(x))_{n=1}^{\infty}$ *сходится к* $F(x)$ *поточечно на* D.

По определению предела это значит, что для каждого числа $x \in D$

$$\forall \varepsilon > 0\, \exists N \geq 1: \text{ если } n > N, \text{ то } |f_n(x) - f(x)| < \varepsilon.$$

При этом при каждом ε число N может зависеть от x.

2. Последовательность $(f_n(x))_{n=1}^{\infty}$ называется *равномерно сходящейся* к функции $f(x)$ на D, если это число N можно выбрать одним и тем же для всех $x \in D$ сразу (но зависящем от ε по–прежнему), т.е.

$$\forall \varepsilon > 0\, \exists N \geq 1: \text{ если } n > N, \text{ то } \forall x \in D\ \ |f_n(x) - f(x)| < \varepsilon.$$

Функциональный ряд $\sum_{k=1}^{\infty} \phi_k(x)$ называется *равномерно сходящимся на* D, если равномерно на D сходится последовательность его частичных сумм. □

Равномерная сходимость существенна при анализе сходимости интегралов от функций, производных от функций, а также при определении свойств предельной функции.

Теорема. *Непрерывность предельной функции.*

Если последовательность непрерывных на D функция сходится равномерно на D, то предельная функция также непрерывна.

Контрпример, показывающий существенность равномерной сходимости для непрерывности предельной функции.

Пусть $f_n(x) = x^n$, $x \in [0, 1]$. При $x < 1$ последовательность x^n сходится к нулю при $n \to \infty$, а при $x = 1$ последовательность x^n состоит из одних единиц. Поэтому предельная функция имеет вид: $f(x) = 0$ при $0 \leq x < 1$ и $f(1) = 1$. Эта предельная функция разрывна в точке $x = 1$. Заключение предыдущей теоремы не выполнено, так как функции $f_n(x)$ сходятся к $f(x)$ поточечно на $[0, 1]$, но не равномерно.

Теорема. *Предельный переход под знаком интеграла.*

Пусть последовательность непрерывных на $[a, b]$ функций $(f_n(x))_{n=1}^{\infty}$ сходится равномерно на $[a, b]$ к функции $f(x)$. Тогда

$$\int_a^b f_n(x)\, dx \to \int_a^b f(x)\, dx$$

при $n \to \infty$, т.е. предел интегралов равен интегралу от предела.

Контрпример, показывающий существенность равномерной сходимости для перестановки предела и интеграла.

Пусть $f_n(x) = 0$ при $x \in [\frac{2}{n}, 1]$, а на промежутке $[0, \frac{2}{n}]$ график функции $f_n(x)$ имеет вид равнобедренного треугольника с основанием на оси абсцисс и высотой n, достигаемой в точке $x = \frac{1}{n}$. Нетрудно видеть, что площадь равнобедренного треугольника равна 1 при любом n. В то же время значения $f_n(x)$ при любом фиксированном $x > 0$ становятся равными нулю при всех достаточно больших n. Следовательно, функции $f_n(x)$ сходятся поточечно (но не равномерно!) к функции $f(x) = 0$.

Таким образом, в этом примере

$$\lim_{n \to \infty} \int_0^1 f_n(x)\, dx = 1, \qquad \int_0^1 f(x)\, dx = 0.$$

Заключение предыдущей теоремы не выполнено, так как сходимость неравномерная.

Теорема. *Предельный переход под знаком производной.*

Пусть последовательность непрерывно дифференцируемых функций $(f_n(x))_{n=1}^{\infty}$ на интервале D (т.е. $f_n \in C^1(D)$, $n{\geq}1$) сходится поточечно к функции $f(x)$ на D, а последовательность производных $(f'_n(x))_{n=1}^{\infty}$ сходится равномерно на D к некоторой функции.

Тогда эта функция есть производная от f, т.е.

$$f'_n(x) \to f'(x), \text{ если } f_n(x) \to f(x)$$

при $n \to \infty$.

Контрпример, показывающий существенность равномерной сходимости для перестановки предела и производной.

Пусть $f_n(x) = \sin(nx)/\sqrt{n}$. Эта последовательность функций равномерно сходится при $n \to \infty$ к функции $f(x) = 0$, у которой $f'(x) = 0$ при всех x. Однако последовательность производных $f'_n(x) = \sqrt{n}\cos x$ вообще не ограничена по n и поэтому не может иметь никакого предела.

В этом примере производная от предела не равна пределу производных, потому что последовательность производных не сходится ни к какому пределу даже поточечно, а тем более, равномерно.

Теорема. *Признак равномерной сходимости Вейерштрасса.*

Пусть последовательность функций $(\phi_k(x))_{k=1}^{\infty}$ задана на множестве D и мажорируется последовательностью чисел $(a_k)_{k=1}^{\infty}$, т.е. $|\phi_k(x)| \leq a_k$ при всех $k \geq 1$ и при всех $x \in D$. Пусть числовой ряд $\sum a_k$ сходится.

Тогда функциональный ряд $\sum \phi_k(x)$ сходится равномерно на D.

Степенные ряды

Определение *степенного ряда.*

Пусть $(a_k)_{k=0}^{\infty}$ — последовательность вещественных (или комплексных) чисел. Ряд

$$a_0 + a_1 z + \cdots + a_n z^n + \cdots = \sum_{k=0}^{\infty} a_k z^k$$

называется *степенным рядом.* Числа a_k называются *коэффициентами* степенного ряда. Комплексное число z называется *аргументом* степенного ряда.

Множество комплексных чисел z, для которых степенной ряд сходится, называется *областью сходимости степенного ряда.* В этой области сумма степенного ряда является функцией от z. □

Теорема *Абеля.*

Если степенной ряд сходится при некотором значении аргумента, то он абсолютно сходится при всех значениях аргумента с меньшим модулем.

Доказательство.

Пусть ряд сходится при $z = z_0$. Тогда общий член ряда $a_k z_0^k$ стремится к нулю при $k \to \infty$. Тем более, последовательность $(a_k z_0^k)_{k=1}^{\infty}$ ограничена. Существует такое число $M > 0$, что

$$|a_k z_0^k| \le M$$

при всех k.

Пусть z_1 — некоторое комплексное число и $|z_1| < |z_0|$. Введем обозначение

$$q = \frac{|z_1|}{|z_0|} < 1.$$

Оценим общий член ряда при $z = z_1$:

$$|a_k z_1^k| = |a_k z_0^k| \left(\frac{|z_1|}{|z_0|}\right)^k \le Mq^k.$$

Ряд $\sum_{k=0}^{\infty} Mq^k$ есть сумма убывающей геометрической прогрессии, он сходится. По теореме сравнения рядов отсюда следует, что ряд

$$\sum_{k=0}^{\infty} a_k z_1^k$$

сходится абсолютно, что и требовалось доказать. □

Следствие.

Если степенной ряд расходится при некотором значении аргумента, то он расходится при всех значениях аргумента с большим модулем. □

Действительно, если бы это было не так, т.е. ряд расходился бы при некотором z_1, сходился при некотором z_2 и $|z_1| < |z_2|$, то это прямо противоречило бы теореме Абеля.

Круг сходимости и радиус сходимости степенного ряда

Пусть степенной ряд сходится при некотором значении $z_1 \ne 0$. На комплексной плоскости C нарисуем окружность с центром в начале координат и радиуса $r_1 = |z_1|$. По теореме Абеля степенной ряд сходится в любой внутренней точке круга, ограниченного этой окружностью.

Предположим, что в некоторой точке $z_2 \in \mathsf{C}$ степенной ряд расходится. Нарисуем окружность с центром в начале координат и радиусом $r_2 = |z_2|$. По следствию из теоремы Абеля вне пределов круга, ограниченного этой окружностью, ряд расходится. Кроме того, из теоремы Абеля следует, что $r_1 \le r_2$.

Пусть $r_1 < r_2$. В кольце между концентрическими окружностями радиусов r_1 и r_2 ряд может как сходиться, так и расходиться. Выберем точку z_3 внутри этого кольца. Если при $z = z_3$ степенной ряд сходится, то внутренний круг сходимости можно расширить до радиуса $r_3 = |z_3|$. Если же степенной ряд расходится, то внешний круг можно сжать до радиуса $r_3 = |z_3|$. После этого кольцо, внутри которого сходимость ряда не установлена, станет у́же.

Поступая и далее таким же образом, будем сужать кольцо между кругом сходимости и внешностью круга расходимости, пока оно не сожмется до нулевой толщины.

Теорема. *Круг сходимости степенного ряда.*

Пусть степенной ряд сходится при некотором $z \ne 0$ и расходится при некотором комплексном z. Тогда существует такой круг с центром в начале координат на комплексной плоскости, что внутри этого круга степенной ряд сходится, а вне круга — расходится. □

Определение.

Этот круг называется *кругом сходимости степенного ряда.* Его радиус называется *радиусом сходимости степенного ряда.* □

Отметим, что на границе круга сходимости, на окружности степенной ряд может в некоторых точках сходиться, а в некоторых расходиться.

Пусть не выполнены условия теоремы о круге сходимости. Если ряд сходится только при $z = 0$, то говорят, что радиус сходимости этого ряда равен нулю.

Если степенной ряд сходится при всех комплексных z, то кругом сходимости называют всю комплексную плоскость, а радиус сходимости считают равным бесконечности.

Для определения радиуса сходимости степенной ряда часто используют признаки Даламбера и Коши.

Следствие. *Признак Даламбера для степенных рядов.*

Пусть последовательность коэффициентов степенного ряда $(a_k)_{k=0}^{\infty}$ обладает следующим свойством: существует предел

$$\lim_{k\to\infty}\left|\frac{a_{k+1}}{a_k}\right| = L.$$

Тогда радиус сходимости степенного ряда равен $1/L$.

Доказательство.

Пусть $|z| < 1/L$. Тогда существует предел отношения соседних членов степенного ряда:

$$\lim_{k\to\infty}\left|\frac{a_{k+1}z^{k+1}}{a_k z^k}\right| = L|z| < 1.$$

По признаку Даламбера ряд абсолютно сходится.

Если $|z| > 1/L$, то из того же условия следует, что ряд расходится. Следовательно, радиус сходимости степенного ряда равен $1/L$. □

Следствие. *Признак Коши для степенных рядов.*

Пусть последовательность коэффициентов степенного ряда $(a_k)_{k=0}^{\infty}$ обладает следующим свойством: существует предел

$$\lim_{k\to\infty}\sqrt[k]{|a_k|} = L.$$

Тогда радиус сходимости степенного ряда равен $1/L$. □

Доказательство аналогично.

Почленное интегрирование и дифференцирование степенных рядов

Степенной ряд является частным случаем функционального ряда. Поэтому можно применить теоремы о почленном интегрировании и дифференцировании функциональных рядов. Важную роль при этом играют вопросы равномерной сходимости.

Следствие *о равномерной сходимости степенного ряда.*

Степенной ряд равномерно сходится в любом круге с центром в начале координат и радиусом меньшим радиуса сходимости.

Доказательство по существу полностью повторяет доказательство теоремы Абеля. Пусть ρ — радиус сходимости и $r < \rho$. Тогда при изучении сходимости в круге радиуса r можно выбрать $q = r/\rho$. Это число не зависит от точки z внутри круга радиуса r и в то же время

$$|a_k z^k| \leq M q^k.$$

По теореме Вейерштрасса функциональный ряд сходится равномерно. □

Следствием общих теорем о интегрировании и дифференцировании функциональных рядов являются два следующих утверждения.

Теорема *о почленном интегрировании степенного ряда.*

Степенной ряд можно почленно интегрировать вдоль любого промежутка внутри круга сходимости. □

Теорема *о почленном дифференцировании степенного ряда.*

Степенной ряд можно почленно дифференцировать в любой точке внутри круга сходимости. □

В качестве z можно взять и вещественное число, т.е. комплексное число, лежащее на оси абсцисс. Поэтому для вещественных чисел областью сходимости степенного ряда является промежуток с симметричными относительно нуля границами.

Пример 1. *Разложение логарифма в ряд.*

Рассмотрим геометрическую прогрессию

$$\sum_{k=0}^{\infty} z^k.$$

Это степенной ряд. Пусть $z = x$ — вещественное число. Как известно, ряд сходится для убывающей прогрессии ($|x| < 1$) и расходится для возрастающей прогрессии ($|x| > 1$). Таким образом, радиус сходимости этого ряда равен 1.

Пусть $|x| < 1$. Тогда

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}.$$

Проинтегрируем это ряд почленно по промежутку $[0, x]$:

$$\sum_{k=0}^{\infty} \frac{x^{k+1}}{k+1} = -\ln(1-x).$$

Полученное равенство после замены x на $-x$ можно записать следующим образом:

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots + (-1)^{n+1}\frac{x^n}{n} + \cdots = \sum_{n=1}^{\infty}(-1)^{n+1}\frac{x^n}{n}.$$

Эта формула называется разложением логарифма в ряд Маклорена. Равенство верно при любых x из промежутка $(-1, 1)$.

Рассмотрим этот ряд на границе промежутка сходимости. При $x = -1$ получается гармонический ряд, который расходится. И действительно, в левой части получается $\ln 0$, что не существует.

При $x = 1$ ряд сходится по признаку Лейбница. Получается следующее равенство:

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + \frac{(-1)^{n+1}}{n} + \cdots = \ln 2.$$

Таким образом, множество сходимости данного ряда на вещественной прямой есть промежуток $(-1, 1]$.

Ряд Тейлора

Определение *ряда Тейлора.*

Пусть функция f имеет в точке a производные всех порядков. Тогда степенной ряд

$$f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \cdots$$

называется *рядом Тейлора функции f в точке a.* □

Частичной суммой порядка n ряда Тейлора функции f является многочлен Тейлора. Поэтому сходимость ряда Тейлора — это сходимость многочленов Тейлора при неограниченном увеличении порядка.

Особое значение имеет вопрос о сходимости ряда Тейлора к самой функции f. Пусть $\mathcal{P}_{f,a,n}(x)$ — многочлен Тейлора функции f в точке a степени n. Запишем остаточный член многочлена Тейлора в форме Лагранжа:

$$R_n(x) = f(x) - \mathcal{P}_{f,a,n}(x) = \frac{f^{(n+1)}(\theta)}{(n+1)!}(x-a)^{n+1},$$

где число θ лежит между a и x.

Отсюда следует, что если остаточный член многочлена Тейлора стремится к нулю, то ряд Тейлора сходится и сумма ряда Тейлора равна $f(x)$. Другими словами, получается разложение функции f в ряд Тейлора.

При $a = 0$ ряд Тейлора называется рядом Маклорена. Если остаточный член многочлена Маклорена стремится к нулю, то функция f раскладывается в ряд Маклорена:

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \cdots = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!}x^n.$$

Пример 2. *Разложение экспоненты в ряд Маклорена.*

Запишем остаточный член многочлена Маклорена для функции $f(x) = e^x$:

$$R_n(x) = e^{\theta}\frac{x^{n+1}}{(n+1)!}, \qquad |\theta| < |x|.$$

В примере к признаку сходимости Даламбера было доказано, что ряд с общим членом $x^n/n!$ сходится. Следовательно, $x^n/n! \to 0$ при $n \to \infty$.

Остаточный член многочлена Маклорена стремится к нулю, поэтому функция e^x разлагается в ряд Маклорена:

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \cdots = \sum_{n=1}^{\infty} \frac{x^n}{n!},$$

и ряд сходится при всех x.

Пример 3. *Разложение синуса и косинуса в ряд Маклорена.*

Легко убедиться в том, что производные от синуса и косинуса удовлетворяют уравнениям

$$\begin{aligned} \sin^{(n)}(x) &= \sin(x + \pi n/2), \qquad n = 1, 2, \ldots \\ \cos^{(n)}(x) &= \cos(x + \pi n/2), \qquad n = 1, 2, \ldots \end{aligned}$$

Остаточные члены многочленов Маклорена стремятся к нулю по той же причине, что и для функции e^x. Поэтому эти функции разлагаются в ряды Маклорена:

$$\begin{aligned} \sin x &= x - \frac{x^3}{6} + \frac{x^5}{120} - \cdots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}, \\ \cos x &= 1 - \frac{x^2}{2} + \frac{x^4}{24} - \cdots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}, \end{aligned}$$

и ряды сходятся при всех вещественных (и комплексных) x.

Ряды Фурье

В начале XIX века французский математик J. Fourier удивил научный мир, заявив, что всякая непрерывная функция на промежутке есть сумма одних синусоид, да еще с кратными частотами (первая публикация — 1807 г.). Последовавшее доказательство этого утверждения убедило всех сомневавшихся и открыло новую область исследования функция — анализ Фурье. В рамках этого анализа остаются только простейшие функции — синусоиды или простые гармоники, а все прочие функции представляются последовательностью коэффициентов их ряда Фурье.

Функция $f(x) = A\cos(\omega x + \phi)$ называется простой гармоникой. Она имеет период, равный $2\pi/\omega$. Число A называется ее амплитудой, ω — частотой, а ϕ — фазой. Функция $f_1(x) = A\sin(\omega x + \phi_1)$ также является простой гармоникой, так как получается из $f(x)$ заменой значения фазы: $\phi = \phi_1 - \pi/2$.

Пусть частота $\omega = n$ — натуральное число. Тогда период $T = 2\pi/n$. На промежутке $[-\pi, \pi]$ укладывается ровно n периодов, а за пределами этого промежутка простая гармоника продолжается периодически. Поэтому мы ограничимся изучением функции только на промежутке $[-\pi, \pi]$.

Пусть $(a_n)_{n=0}^{\infty}$ и $(b_n)_{n=1}^{\infty}$ — последовательности вещественных чисел. Пусть функциональный ряд

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

сходится на промежутке $[-\pi,\pi]$, и его сумма обозначается $f(x)$. Ввиду периодичности каждого слагаемого ряд сходится и за пределами промежутка $[-\pi,\pi]$ и функция f периодическая. Числа a_n и b_n называются коэффициентами Фурье.

Вычисление коэффициентов Фурье

Вычислим интеграл от простых гармоник и их произведений по промежутку $[\pi,\pi]$. При $n,m\geq 1$:

$$\begin{aligned}
\int_{-\pi}^{\pi}\cos nx\,dx &= \frac{1}{n}\sin nx\Big|_{-\pi}^{\pi}=0,\\
\int_{-\pi}^{\pi}\sin nx\,dx &= -\frac{1}{n}\cos nx\Big|_{-\pi}^{\pi}=0,\\
\int_{-\pi}^{\pi}\sin nx\cos mx\,dx &= \frac{1}{2}\int_{-\pi}^{\pi}(\sin(n+m)x-\sin(n-m)x)\,dx=0,\\
\int_{-\pi}^{\pi}\cos nx\cos mx\,dx &= \frac{1}{2}\int_{-\pi}^{\pi}(\cos(n-m)x+\cos(n+m)x)\,dx=0,\quad (n\neq m),\\
\int_{-\pi}^{\pi}\sin nx\sin mx\,dx &= \frac{1}{2}\int_{-\pi}^{\pi}(\cos(n-m)x-\cos(n+m)x)\,dx=0,\quad (n\neq m),\\
\int_{-\pi}^{\pi}\cos^2 nx\,dx &= \frac{1}{2}\int_{-\pi}^{\pi}(1+\cos 2nx)\,dx=\pi,\\
\int_{-\pi}^{\pi}\sin^2 nx\,dx &= \frac{1}{2}\int_{-\pi}^{\pi}(1-\cos 2nx)\,dx=\pi.
\end{aligned}$$

Предположим, что ряд сходится равномерно. Тогда по теореме о почленном интегрировании ряда интеграл от суммы ряда равен сумме ряда из интегралов. В данном случае

$$\begin{aligned}
\int_{-\pi}^{\pi}f(x)\,dx &= a_0\pi,\\
\int_{-\pi}^{\pi}f(x)\cos nx\,dx &= a_n\pi,\qquad n\geq 1\\
\int_{-\pi}^{\pi}f(x)\sin nx\,dx &= b_n\pi,\qquad n\geq 1.
\end{aligned}$$

Действительно, все слагаемые в ряде после интегрирования оказываются равными нулю, кроме одного, "своего" слагаемого, содержащего ту функцию, на которую $f(x)$ умножается перед интегрированием.

Теперь предположим, что функция f задана, а ряд строится по коэффициентам, вычисляемым через интегралы.

Определение *ряда Фурье.*

Пусть функция f задана и интегрируема на промежутке $[-\pi,\pi]$. *Рядом Фурье* для функции f называется функциональный ряд

$$F(x)=\frac{a_0}{2}+\sum_{n=1}^{\infty}(a_n\cos nx+b_n\sin nx),$$

в котором коэффициенты определяются равенствами

$$\begin{aligned} a_0 &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\,dx, \\ a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx\,dx, \qquad n \geq 1, \\ b_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx\,dx, \qquad n \geq 1. \end{aligned}$$

Эти коэффициенты называются *коэффициентами Фурье* для функции f. □

Разложимость функции в ряд Фурье

Ряд Фурье на промежутке $[-\pi, \pi]$ можно написать для любой интегрируемой по этому промежутку функции. Он может сходиться или нет, а сумма F может быть равна функции f (если ряд сходится) или не равна. Если $f = F$, то говорят, что функция f раскладывается в ряд Фурье. Оказывается, что при слабых ограничениях любая функция раскладывается в ряд Фурье.

Поскольку сумма ряда Фурье $F(x)$ — всегда периодическая функция, $F(x+2\pi) = F(x)$, то естественно предположить, что f обладает тем же свойством $f(x+2\pi) = f(x)$. В дальнейшем это свойство будет предполагаться.

Определение.

1. Если функция f имеет в точке x_0 разрыв первого рода (скачок), т.е. существуют односторонние пределы

$$\lim_{x\to x_0+0} f(x) = f(x_0+0), \qquad \lim_{x\to x_0-0} f(x) = f(x_0-0),$$

и если в точке x_0 значение f есть полусумма этих пределов,

$$f(x_0) = \frac{f(x_0+0)+f(x_0-0)}{2},$$

то разрыв функции f в точке x_0 называется *правильным*.

2. Функция называется *кусочно–гладкой*, если промежуток ее задания можно разбить на конечное число промежутков, в каждом из которых эта функция имеет непрерывную производную.

3. Функция называется *кусочно–монотонной*, если промежуток ее задания можно разбить на конечное число промежутков, в каждом из которых эта функция монотонна. □

Теорема *об условиях разложения функции в ряд Фурье.*

Пусть функция $f(x)$ либо кусочно–гладкая, либо кусочно–монотонная на промежутке $[-\pi, \pi]$, а все ее разрывы правильные, то она разлагается в ряд Фурье. □

(Без доказательства.)

Разложение четных и нечетных функций

Напомним, что функция f называется нечетной, если $f(-x) = -f(x)$ и четной, если $f(-x) = f(x)$ при всех x. Функции $\sin nx$ нечетные, а $\cos nx$ четные при всех

$n \geq 1$. Кроме того, произведение двух четных функций или двух нечетных функций четно, а произведение четной функции на нечетную нечетно.

Интеграл от нечетной функции по симметричному промежутку равен нулю, так как интегралы по положительной половине промежутка и по отрицательной половине промежутка равны по абсолютной величине и отличаются знаком:

$$\int_{-L}^{L} f(x)\,dx = \int_{-L}^{0} f(x)\,dx + \int_{0}^{L} f(x)\,dx = 0.$$

Если же функция f четная, то интегралы по половинам промежутка равны, и поэтому

$$\int_{-L}^{L} f(x)\,dx = \int_{-L}^{0} f(x)\,dx + \int_{0}^{L} f(x)\,dx = 2\int_{0}^{L} f(x)\,dx.$$

Применим эти свойства к интегралам, определяющим коэффициенты Фурье.

Если функция f четная, то все коэффициенты b_n равны нулю. Поэтому

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nx,$$

где

$$a_n = \frac{2}{\pi}\int_{0}^{\pi} f(x)\cos nx\,dx, \qquad n \geq 0.$$

Этот ряд Фурье называется разложением функции f по косинусам.

Если функция f нечетная, то все коэффициенты a_n равны нулю. Поэтому

$$f(x) = \sum_{n=1}^{\infty} b_n \sin nx,$$

где

$$b_n = \frac{2}{\pi}\int_{0}^{\pi} f(x)\sin nx\,dx, \qquad n \geq 1.$$

Этот ряд Фурье называется разложением функции f по синусам.

Предположим, что функция f задана на промежутке $(0, \pi]$. Ее можно продолжить на промежуток $[-\pi, 0)$ как четную, т.е. по правилу $f(-x) = f(x)$, и тогда f может быть разложена только по косинусам.

Но функцию f можно продолжить на промежуток $[-\pi, 0)$ и как нечетную, т.е. по правилу $f(-x) = -f(x)$, и тогда f может быть разложена только по синусам.

Таким образом, каждую функцию на промежутке $(0, \pi]$ можно разложить в ряд Фурье:

- по синусам и косинусам, $\cos nx$ и $\sin nx$, если она произвольным образом доопределена на промежутке $[-\pi, 0)$;
- только по синусам $\sin nx$;
- только по косинусам $\cos nx$.

В последних двух случаях количество гармоник не согласуется с длиной промежутка, в два раза больше в следующем смысле: минимальный период гармоник равен 2π, а функция f задана на промежутке длиной π.

Пример. Разложить на промежутке $(-\pi, \pi)$ в ряд Фурье функцию $f(x) = \operatorname{sign} x$. По определению, $f(x) = 1$, если $x > 0$; $f(x) = -1$, если $x < 0$; и $f(0) = 0$.

Эта функция кусочно–гладкая, а разрыв в точке $x = 0$ правильный. Поэтому функция разлагается в ряд Фурье. Функция f нечетная, следовательно, она разлагается по синусам, т.е. $a_n = 0$ при всех $n \geq 0$.

$$b_n = \frac{2}{\pi} \int_0^{\pi} f(x) \sin nx \, dx = \frac{2}{\pi} \int_0^{\pi} \sin nx \, dx = \frac{4}{n\pi}.$$

Окончательно,

$$\operatorname{sign} x = \frac{4}{\pi} \left(\sin x + \frac{1}{3} \sin 3x + \frac{1}{5} \sin 5x + \cdots \right), \qquad -\pi < x < \pi.$$

Заметим, что при $x = \pi/2$ получается уже известная формула, которую можно использовать для приближенного вычисления числа π:

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots$$

Разложение в ряд Фурье по произвольному промежутку

Пусть функция f задана на промежутке $[-L, L]$. Для того, чтобы ее разложить в ряд Фурье, сделаем замену переменной. Определим функцию $g(y)$ на промежутке $[-\pi, \pi]$ равенством $g(y) = f(x)$ при $y = \pi x / L$.

Функция $g(y)$ может быть разложена в ряд Фурье на промежутке $[-\pi, \pi]$. После этого вместо y нужно подставить $\pi x / L$. В итоге получится следующий ряд:

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos \frac{\pi n x}{L} + b_n \sin \frac{\pi n x}{L} \right).$$

Коэффициенты Фурье рассчитываются по формулам

$$\begin{aligned}
a_0 &= \frac{1}{L} \int_{-L}^{L} f(x) \, dx, \\
a_n &= \frac{1}{L} \int_{-L}^{L} f(x) \cos \frac{\pi n x}{L} \, dx, \qquad n \geq 1, \\
b_n &= \frac{1}{L} \int_{-L}^{L} f(x) \sin \frac{\pi n x}{L} \, dx, \qquad n \geq 1.
\end{aligned}$$