ПЕРМСКИЙ
ГОСУДАРСТВЕННЫЙ
НАЦИОНАЛЬНЫЙ
ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

Е. Ю. Еленская,
Ю. Н. Еленский

ОСНОВЫ НЕЛИНЕЙНОГО ФУНКЦИОНАЛЬНОГО АНАЛИЗА

ИЗДАТЕЛЬСКИЙ ЦЕНТР
"Perm University Press"

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное автономное
образовательное учреждение высшего образования
«ПЕРМСКИЙ ГОСУДАРСТВЕННЫЙ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ»

Е. Ю. Еленская, Ю. Н. Еленский

ОСНОВЫ НЕЛИНЕЙНОГО ФУНКЦИОНАЛЬНОГО АНАЛИЗА

*Допущено методическим советом
Пермского государственного национального
исследовательского университета в качестве
учебно-методического пособия для студентов,
обучающихся по направлению подготовки бакалавров
«Математика»*

ИЗДАТЕЛЬСКИЙ ЦЕНТР
"Perm University Press"

Пермь 2022

УДК 517.988(075.8)
ББК 22162я73
Е506

Еленская Е. Ю.

Е506 Основы нелинейного функционального анализа [Электронный ресурс] : учебно-методическое пособие / Е. Ю. Еленская, Ю. Н. Еленский ; Пермский государственный национальный исследовательский университет. – Электронные данные. – Пермь, 2022. – 1 Мб ; 100 с. – Режим доступа: http://www.psu.ru/files/docs/science/books/uchebnie-posobiya/Elenskaya-Elenskij-Osnovy-Nelinejnogo-Funkcionalnogo-Analiza.pdf – Заглавие с экрана.

ISBN 978-5-7944-3843-7

Пособие разработано в соответствии с рабочей программой дисциплины «Нелинейный функциональный анализ». Рассматриваются теоретические и практические аспекты нелинейного функционального анализа. Основное внимание уделяется обобщению понятия производной функции и применению производной к исследованию и решению различных задач. В каждом разделе представлены теоретические сведения, утверждения и теоремы с доказательством, примеры решения задач, а также задачи для самостоятельного решения, что дает преподавателю широкий выбор заданий для организации самостоятельной работы студентов. Также пособие может использоваться при дистанционном обучении и при самостоятельном изучении отдельных разделов курса.

Пособие предназначено для студентов и преподавателей вузов физико-математических специальностей.

УДК 517.988(075.8)
ББК 22162я73

Издается по решению ученого совета механико-математического факультета Пермского государственного национального исследовательского университета

Рецензенты: кафедра «Высшая математика» ПНИПУ (зав. кафедрой – д-р физ.-мат. наук, профессор ***А. Р. Абдуллаев***);

доцент кафедры высшей математики и методики обучения математике ПГГПУ, канд. пед. наук, доцент ***Е. Л. Черемных***

ISBN 978-5-7944-3843-7

Введение

Многие задачи математической физики, механики и других областей знания сводятся к решению операторных уравнений. Важнейшими из них являются линейные уравнения, но на практике большинство задач приводится к нелинейным уравнениям. В данном учебно-методическом пособии излагаются основы нелинейного функционального анализа. Основное внимание уделяется обобщению понятия производной функции и применению производной к исследованию и решению различных задач.

В первом разделе приводятся сведения из общего курса функционального анализа, необходимые для понимания последующего материала.

Во втором разделе излагаются теоретические основы нелинейного функционального анализа с примерами. Предлагаются задачи для самостоятельного решения. В конце даются указания к решению задач и ответы.

1. Основные сведения из функционального анализа

Обозначения:

R – множество действительных чисел,

C – множество комплексных чисел,

P – множество действительных или комплексных чисел.

1.1. Линейные операторы

Определение. Оператор $A: X \to Y$ называется ***линейным***, если

1) X, Y – линейные пространства;

2) $(\forall \lambda_1, \lambda_2 \in \mathrm{P})$ $(\forall x_1, x_2 \in X)$: $A(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 A x_1 + \lambda_2 A x_2$,

где $\mathbf{P} = \mathbf{R}$ или $\mathbf{P} = \mathbf{C}$.

Часто вместо условия 2) требуют выполнения системы из двух условий:

$A(x_1 + x_2) = Ax_1 + Ax_2 \quad (x_1, x_2 \in X)$ − аддитивность,

$A(\lambda x) = \lambda A x \quad (\lambda \in P, x \in X)$ − однородность.

Пример 1

Пусть $X = Y = \mathbf{R}$, $Ax = ax$ (a – фиксированное число). Условия определения, очевидно, выполнены, поэтому A − линейный оператор. Заметим, что в школьном курсе математики функция A, определенная равенством $Ax = ax + b$, называется линейной, но при $b \neq 0$ оператор A не является линейным.

Пример 2

Пусть

$$(Ax)(t) = \int_a^b k(t,s)x(s)ds\,,\ t \in [a;b]\,,$$

где k – фиксированная функция, непрерывная на $[a;b]^2$, $A : C[a;b] \to C[a;b]$, $C[a,b]$ − пространство вещественных функций, непрерывных на отрезке $[a,b]$. Покажем, что оператор A является линейным.

Очевидно, условие 1) выполнено ($C[a;b]$ является линейным пространством).

Проверим условие 2):

$$A(\lambda_1 x_1 + \lambda_2 x_2)(t) = \int_a^b k(t,s)(\lambda_1 x_1(s) + \lambda_2 x_2(s))ds =$$

$$= \lambda_1 \int_a^b k(t,s)x_1(s)ds + \lambda_2 \int_a^b k(t,s)x_2(s)ds = \lambda_1 (Ax_1)(t) + \lambda_2 (Ax_2)(t)\,,$$

т.е. условие 2) выполнено. Тогда согласно определению оператор A является линейным.

Такой оператор называют линейным интегральным оператором Фредгольма. Этот оператор рассматривают и в других функциональных пространствах.

1.2. Метрические пространства

Пусть X – некоторое множество.

Определение. Функция $d : X^2 \to R$ называется ***метрикой*** на множестве X, если

1) $(\forall x, y \in X)$: $d(x,y) = d(y,x)$;

2) $d(x,y)=0 \Leftrightarrow x=y$;

3) $(\forall x,y,z \in X)$: $d(x,y) \le d(x,z)+d(z,y)$ (неравенство треугольника).

Утверждение. $(\forall x,y \in X)$: $d(x,y) \ge 0$.

Определение. Число $d(x,y)$ называется ***расстоянием*** между элементами x и y.

Определение. ***Метрическим пространством*** называется упорядоченная пара $<X,d>$, где X – множество, d – метрика на множестве X.

Примеры метрик и метрических пространств

1. Пусть $X \neq \varnothing$, $d(x_1,x_2)=\begin{cases} 0, x_1 = x_2 \\ 1, x_1 \neq x_2 \end{cases}$.

Тогда $<X,d>$ – метрическое пространство.

2. Пусть $X=R$, $d(x_1,x_2)=|x_1-x_2|$.

Тогда $<X,d>$ – метрическое пространство.

3. Пусть $X=R^n$ (n-мерное евклидово пространство).

Элементами пространства являются

$$x = <x_1, x_2, ..., x_n> \ (x_k \in \mathbf{R}, k=1,2,...,n),$$

$$y = <y_1, y_2, ..., y_n> \ (y_k \in \mathbf{R}, k=1,2,...,n).$$

Зададим метрику формулой

$$d(x,y)=\sqrt{\sum_{k=1}^{n}(x_k-y_k)^2}.$$

Тогда $<X,d>$ – метрическое пространство.

4. Пусть $X=C[a;b]$ (пространство вещественных непрерывных на отрезке [a;b] функций). Зададим метрику формулой

$$d(x,y)=\max_{t\in[a;b]}|x(t)-y(t)|.$$

Тогда $<X,d>$ – метрическое пространство.

5. Пусть $X = C^k[a;b]$ (пространство вещественных функций на отрезке [a;b], имеющих непрерывную производную до порядка k включительно).

Зададим метрику формулой

$$d(x,y) = \max_{t\in[a;b]}\left|x(t)-y(t)\right| + \max_{t\in[a;b]}\left|x'(t)-y'(t)\right| + \ldots + \max_{t\in[a;b]}\left|x^{(k)}(t)-y^{(k)}(t)\right|.$$

Тогда $<X,d>$ – метрическое пространство.

6. Пусть $X = L_p[a;b]$, $p \ge 1$ (пространство функций, для которых существует $\int_a^b |x(t)|^p\,dt$ – интеграл Лебега). Зададим метрику формулой

$$d(x,y) = \left(\int_a^b |x(t)-y(t)|^p\,dt\right)^{1/p}.$$

Тогда $<X,d>$ – метрическое пространство.

1.3. Сходимость в метрических пространствах

Определение. Последовательность $\{x_n\}$ ***сходится к элементу*** x в метрическом пространстве $<X,d>$, если $d(x_n,x)\to 0$ при $n\to\infty$.

В пространстве непрерывных функций $C[a;b]$ определение имеет следующий вид:

$$x_n \to x \Leftrightarrow d(x_n,x) = \max_{t\in[a;b]}\left|x_n(t)-y(t)\right| \to 0 \text{ при } n\to\infty,$$

или

$$x_n \to x \Leftrightarrow (\forall \varepsilon > 0)\ (\exists N \in \mathbf{N})\ (\forall n \in \mathbf{N} \mid n > N) : \max_{t\in[a;b]}\left|x_n(t)-x(t)\right| < \varepsilon$$

.

Так как условие $|x_n(t)-x(t)|<\varepsilon$ выполнено для любого $t\in[a;b]$, то получаем равномерную сходимость последовательности функций $\{x_n(t)\}$ к функции $x(t)$.

Тогда $x_n \to x$ в $C[a;b] \Leftrightarrow x_n(t)\to x(t)$ равномерно на $[a;b]$, т.е. сходимость в $C[a;b]$ – ***равномерная сходимость*** последовательности функций.

В пространстве $L_2[a;b]$ определение имеет следующий вид:

$$x_n \to x \Leftrightarrow d(x_n,x)=\left(\int_a^b |x_n(t)-x(t)|^2\,dt\right)^{1/2} \to 0 \text{ при } n\to\infty \Leftrightarrow$$

$$\Leftrightarrow \int_a^b |x_n(t)-x(t)|^2\,dt \to 0 \text{ при } n\to\infty .$$

Тогда имеем ***сходимость в среднеквадратическом***.

1.4. Полные метрические пространства

Пусть $< X, d >$ – метрическое пространство.

Определение. Последовательность $\{x_n\}\subset X$ называется ***фундаментальной***, если

$$(\forall\varepsilon>0)\ (\exists N\in\mathbf{N})\,(\forall m,n\in\mathbf{N}\,|\,m,\ n>N): d(x_m,x_n)<\varepsilon .$$

Утверждение. Если последовательность $\{x_n\}$ сходится, то последовательность $\{x_n\}$ фундаментальна.

Обратное может не выполняться.

Определение. Метрическое пространство $< X, d >$ называется ***полным***, если в нем каждая фундаментальная последовательность сходится.

Примеры полных метрических пространств

1. $\mathbf{R}$ – пространство действительных чисел.
2. Пространство $\mathbf{R}^{\mathbf{n}}$.
3. Пространства $C[a;b]$, $C^k[a;b]$.

4. Пространство $L_p[a;b]$.

Пример неполного метрического пространства

Метрическое пространство $< X, d >$, где X – множество функций, непрерывных на $[a;b]$, $d(x,y) = \int\limits_a^b |x(t) - y(t)| dt$ – интегральная метрика.

1.5. Критерий полноты метрического пространства

Определение. Метрическое пространство $< X_0, d_0 >$ называется ***метрическим подпространством*** метрического пространства $< X, d >$, если

1) $X_0 \subset X$;

2) $d_0 = d\Big|_{X_0^2}$ – сужение метрики d на X_0^2, т.е. $d_0(x_1, x_2) = d(x_1, x_2)$ при $x_1, x_2 \in X_0$; другими словами, d_0 есть d с более узкой областью определения.

Критерий замкнутости множества. Множество замкнуто тогда и только тогда, когда пределы всех сходящихся последовательностей элементов множества лежат в этом множестве.

$$M \text{ замкнуто } \Leftrightarrow \ (\forall \{x_n\} \subset M \big| \ x_n \to x): \ x \in M .$$

Теорема (критерий полноты подпространства). Для полноты подпространства $< X_0, d_0 >$ полного метрического пространства $< X, d >$ необходимо и достаточно, чтобы X_0 было замкнуто в X.

1.6. Нормированные пространства

Пусть X – линейное пространство.

Определение. Нормой на X называется функция $\|\cdot\|: X \to \mathrm{R}$, удовлетворяющая условиям:

1) $\|x\| = 0 \Leftrightarrow x = \theta$;

2) $\|\lambda x\| = |\lambda| \cdot \|x\| \quad \forall \lambda \in P, \forall x \in X$ (где $P = \mathbf{R}$ или $P = \mathbf{C}$);

3) $\|x + y\| \le \|x\| + \|y\|$ (неравенство треугольника).

Комментарий. $\|x\|$ – аналог длины вектора.

Свойства нормы

Утверждение 1. $\|x\| \ge 0$ ($\forall x \in X$).

Утверждение 2. $\|x - y\| \ge \big| \|x\| - \|y\| \big|$ ($\forall x, y \in X$).

Утверждение 3. Норма – непрерывная функция, т.е. из $x_n \to x$ следует $\|x_n\| \to \|x\|$.

Утверждение 3 показывает, что под знаком нормы можно переходить к пределу.

Утверждение 4. Операция сложения непрерывна, т.е. если $x_n \to x$, $y_n \to y$, то $x_n + y_n \to x + y$ (предел суммы равен сумме пределов).

Утверждение 5. Операция умножения на число непрерывна, т.е. если $\lambda_n \to \lambda$, $\lambda_n, \lambda \in P$, $x_n \to x$, $x_n, x \in X$, то $\lambda_{\mathbf{n}} x_n \to \lambda x$.

Определение. Нормированным пространством называется упорядоченная пара $< X, \|\cdot\| >$, где X – множество, $\|\cdot\|$ – норма на множестве X.

Примеры

1. Пусть $X = \mathbf{R}$, $\|x\| = |x|$.

Тогда $< X, \|\cdot\| >$ – нормированное пространство.

2. Пусть $X = \mathbf{R}^{\mathbf{n}}$ – n-мерное евклидово пространство; элементами этого пространства являются векторы $x = < x_1, \ldots, x_n >$, $x_k \in \mathbf{R}$, $k = \overline{1, n}$.

Зададим норму формулой

$$\|x\| = \sqrt{\sum_{k=1}^{n} |x_k|^2} .$$

Тогда $< X, \|\cdot\| >$ – нормированное пространство.

3. Пусть $X = C[a;b]$ – пространство вещественных функций, непрерывных на отрезке $[a;b]$. Зададим норму формулой

$$\|x\| = \max_{t\in[a;b]} |x(t)| .$$

Тогда $< X, \|\cdot\| >$ – нормированное пространство.

4. Пусть $X = L_p[a;b]$ – пространство вещественных функций, для которых существует интеграл Лебега $\int_a^b |x(t)|^p dt$; элементы пространства – классы эквивалентных функций. Зададим норму формулой

$$\|x\| = \left(\int_a^b |x(t)|^p dt \right)^{\frac{1}{p}} ; \ L_p[a;b] \text{ – нормированное пространство.}$$

Тогда $< X, \|\cdot\| >$ – нормированное пространство.

Утверждение. Каждое нормированное пространство является метрическим пространством, в котором $d(x, y) = \|x - y\|$.

Определение сходимости в нормированном пространстве вытекает из определения сходимости в метрическом пространстве:

$x_n \to x \quad \Leftrightarrow \quad d(x_n, x) \to 0$ при $n \to \infty \quad \Leftrightarrow \quad \|x_n - x\| \to 0$ при $n \to \infty$.

1.7. Линейные непрерывные операторы

Пусть X, Y – нормированные пространства.

Определение. Оператор $A: X \to Y$ (необязательно линейный) называется ***непрерывным*** в точке $x \in X$, если из $x_n \to x$ следует $Ax_n \to Ax$.

Определение. Оператор A называется ***непрерывным на множестве*** M, если этот оператор непрерывен во всех точках множества M.

Определение. Оператор $A: X \to Y$ (линейный или нелинейный) называется ***ограниченным***, если оператор A преобразует каждое ограниченное множество пространства X в ограниченное множество пространства Y (множество D ограничено $\Rightarrow$ множество $A(D)$ ограничено).

Теорема (о непрерывности линейного оператора).

Пусть X, Y – нормированные пространства, $A: X \to Y$ – линейный оператор. Тогда эквивалентны четыре утверждения:

1) оператор A непрерывен в одной точке;
2) оператор A непрерывен на всем множестве X;
3) оператор A ограничен;
4) $(\exists \lambda \geq 0)$ $(\forall x \in X)$: $\|Ax\| \leq \lambda \|x\|$.

Норма линейного оператора

Определение. Нормой линейного оператора $A: X \to Y$ называется число

$$\|A\| = \min\{\lambda \geq 0 : \|Ax\| \leq \lambda \|x\| \; \forall x \in X\} .$$

Теорема. Норма линейного оператора вычисляется по формуле

$$\|A\| = \sup_{\|\mathbf{x}\| \leq 1} \|Ax\| = \sup_{\|\mathbf{x}\|=1} \|Ax\| = \sup_{x \neq \theta} \frac{\|Ax\|}{\|x\|} .$$

1.8. Линейные функционалы

Определение. Функционалом называется оператор $A: X \to P$, где X – пространство $P = \mathbf{R}$ или $P = \mathbf{C}$.

Другими словами, функционал – это оператор с числовыми значениями.

Определение. Если функционал $F: X \to P$ и $P = \mathbf{R}$, то F называется ***вещественным функционалом***; если $P = \mathbf{C}$, то F – ***комплексный функционал***.

Функционалы могут быть линейными и нелинейными. Так как линейный функционал – это линейный оператор, то теория линейных операторов справедлива для линейных функционалов.

Определение. Пространством, сопряженным к нормированному пространству X, называется множество всех линейных непрерывных функционалов, определенных на пространстве X.

Обозначим X^* – пространство, сопряженное к пространству X.

Примеры.

1. $f(x) = x(0)$, $f: C[0;1] \to \mathbf{R}$ – линейный непрерывный функционал.

2. $f(x) = \int\limits_a^b x(t)dt$, $f: C[a;b] \to \mathbf{R}$ – линейный непрерывный функционал.

1.9. Гильбертовы пространства

Пусть X – линейное пространство (комплексное).

Определение. Скалярным произведением на множестве X называется функция $(\cdot,\cdot): X^2 \to \mathbf{C}$, удовлетворяющая условиям:

1) $(x, y) = \overline{(y, x)}$;

2) $(x_1 + x_2, y) = (x_1, y) + (x_2, y)$;

3) $(\lambda x, y) = \lambda(x, y)$;

4) $(x.x) \in \mathbf{R}$, $(x, x) \geq 0 \Leftrightarrow x = \theta$.

Определение. Предгильбертовым пространством называется упорядоченная пара $< X, (\cdot, \cdot) >$, где X – линейное пространство, $(\cdot, \cdot)$ – скалярное произведение на X .

Утверждение. Функция $\|\cdot\|$, определенная равенством $\|x\| = \sqrt{(x, x)}$, является нормой на X .

Следствие. Каждое предгильбертово пространство является нормированным.

Ортогональность в предгильбертовом пространстве определяется следующим образом: $x \perp y \Leftrightarrow (x, y) = 0$.

Теорема Пифагора. Если $x \perp y$, то $\|x + y\|^2 = \|x\|^2 + \|y\|^2$.

Определение. Гильбертовым пространством называется полное предгильбертово пространство.

Примеры гильбертовых пространств

1. $X = \mathbf{C}^{\mathbf{n}}$, $x =< x_1, \ldots x_n >$, $y =< y_1, \ldots, y_n >$, $(x, y) = \sum_{k=1}^{n} x_k \overline{y_k}$.

2. $X = L_2[a; b]$ (комплексное пространство), $(x, y) = \int_a^b x(t)\overline{y(t)}dt$.

2. Основы нелинейного функционального анализа

Производная вещественной функции – одно из основных понятий математического анализа. С возникновением и развитием функционального анализа были сделаны обобщения понятия «производная» на отображения (операторы) произвольного вида. Рассмотрим такие обобщения и их применение к решению различных задач.

2.1. Дифференциал Фреше. Производная Фреше

Пусть X, Y – нормированные пространства, $x \in X$, G – открытое множество в пространстве X.

Определение. Оператор $A: G \to Y$ ***дифференцируем по Фреше*** в точке $x \in G$, если для любого $h \in X$, для которого $x+h \in G$, приращение оператора можно представить в виде

$$A(x+h) - Ax = Bh + \omega(h), \qquad (1)$$

где оператор $B: X \to Y$ линеен и непрерывен, а ω удовлетворяет условию

$$\frac{\|\omega(h)\|}{\|h\|} \to 0 \text{ при } \|h\| \to 0. \qquad (2)$$

Элемент Bh называется ***дифференциалом Фреше*** оператора A в точке x с приращением h; оператор B называется ***производной Фреше*** оператора A в точке x.

Другими словами, дифференциал Фреше – это линейная часть приращения, главная в смысле условия (2), ω – остаток более высокого порядка малости, чем h. При этом производная Фреше B – это в общем случае не число, а линейный непрерывный оператор. Производная является числом только для вещественных или комплексных функций одной переменной. Но и в этом случае производную Фреше можно считать линейным оператором умножения на это число.

Введем ***обозначения***:

$A'(x)$ – производная Фреше оператора A в точке x;

$dA(x,h)$ – дифференциал Фреше оператора A в точке x с приращением h.

Из условия (1) получаем

$$dA(x,h) = Bh = A'(x)h.$$

Тогда

$$A(x+h) - Ax = A'(x)h + \omega(h).$$

Пример 1. Производная Фреше функции одной переменной.

Пусть

$$f:\mathbf{R}\to\mathbf{R}\,,\ f(x+\Delta x)-f(x)=f'(x)\Delta x+\omega(h)\,.$$

Дифференциал Фреше определяется следующим образом:

$$df(x,\Delta x)=f'(x)\Delta x\,.$$

Получилась формула, известная в математическом анализе. Дифференциал Фреше совпадает с известным дифференциалом вещественной функции. Этот дифференциал можно рассматривать как значение оператора умножения аргумента Δx на число $f'(x)$. Поэтому, с одной стороны, $f'(x)$ − число, с другой – это оператор умножения на число $f'(x)$, что соответствует определению производной Фреше.

Пример 2. Производная Фреше конечномерного оператора.

Пусть $A:\mathbf{R}^{\mathbf{n}}\to\mathbf{R}^{\mathbf{m}}$, элементы пространств – векторы, которые здесь удобно представть в виде матриц-столбцов:

$$x=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}\in\mathbf{R}^{\mathbf{n}}\,,\ y=\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}\in\mathbf{R}^{\mathbf{m}}\,.$$

$$Ax=\begin{pmatrix}a_1(x)\\ \vdots\\ a_m(x)\end{pmatrix},\qquad(3)$$

$x\in\mathbf{R}^{\mathbf{n}}$, $a_k:U\to\mathrm{R}$ ($k=1,...m$), где U – окрестность точки $x\in\mathbf{R}^{\mathbf{n}}$ и $A:U\to\mathbf{R}^{\mathbf{m}}$.

Теорема. Пусть $a_k:U\to\mathrm{R}$, $k=\overline{1,m}$, U – окрестность точки $x\in\mathbf{R}^{\mathbf{n}}$, все a_k дифференцируемы в точке x как функции n переменных (существуют дифференциалы). Тогда оператор A, определенный формулой (3), дифференцируем по Фреше в точке x и справедлива формула

$$A'(x)=\begin{pmatrix} \frac{\partial a_1(x)}{\partial x_1}(x) & \frac{\partial a_1(x)}{\partial x_2}(x) & \ldots & \frac{\partial a_1(x)}{\partial x_n}(x) \\ \frac{\partial a_2(x)}{\partial x_1}(x) & \frac{\partial a_2(x)}{\partial x_2}(x) & \ldots & \frac{\partial a_2(x)}{\partial x_n}(x) \\ \ldots & \ldots & \ldots & \ldots \\ \frac{\partial a_m(x)}{\partial x_1}(x) & \frac{\partial a_m(x)}{\partial x_2}(x) & \cdots & \frac{\partial a_m(x)}{\partial x_n}(x) \end{pmatrix} \text{ (матрица Якоби).} \quad (4)$$

Доказательство. Рассмотрим приращение оператора A в точке x, выделим из него линейную часть, затем покажем, что эта линейная часть является главной. Тогда полученная линейная часть приращения оператора будет дифференциалом Фреше, а из дифференциала найдем производную Фреше.

Возьмем приращение аргумента $h=\begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \in \mathbf{R}^\mathbf{n}$. Тогда

$$A(x+h)-Ax=\begin{pmatrix} a_1(x+h) \\ \vdots \\ a_m(x+h) \end{pmatrix}-\begin{pmatrix} a_1(x) \\ \vdots \\ a_m(x) \end{pmatrix}=\begin{pmatrix} a_1(x+h)-a_1(x) \\ \vdots \\ a_m(x+h)-a_m(x) \end{pmatrix}=$$

$$=\begin{pmatrix} \sum\limits_{j=1}^{n}\frac{\partial a_1}{\partial x_j}(x)h_j+\omega_1(h) \\ \vdots \\ \sum\limits_{j=1}^{n}\frac{\partial a_m}{\partial x_j}(x)h_j+\omega_m(h) \end{pmatrix}=\begin{pmatrix} \sum\limits_{j=1}^{n}\frac{\partial a_1}{\partial x_j}(x)h_j \\ \vdots \\ \sum\limits_{j=1}^{n}\frac{\partial a_m}{\partial x_j}(x)h_j \end{pmatrix}+\begin{pmatrix} \omega_1(h) \\ \vdots \\ \omega_m(h) \end{pmatrix}.$$

Первое слагаемое – линейная по h часть приращения. Второе слагаемое обозначим символом $\omega(h)$.

Тогда $\omega(h)=\begin{pmatrix} \omega_1(h) \\ \vdots \\ \omega_m(h) \end{pmatrix}$,

$$A(x+h)-Ax=\begin{pmatrix} \frac{\partial a_1(x)}{\partial x_1}(x) & \frac{\partial a_1(x)}{\partial x_2}(x) & \dots & \frac{\partial a_1(x)}{\partial x_n}(x) \\ \frac{\partial a_2(x)}{\partial x_1}(x) & \frac{\partial a_2(x)}{\partial x_2}(x) & \dots & \frac{\partial a_2(x)}{\partial x_n}(x) \\ \dots & \dots & \dots & \dots \\ \frac{\partial a_m(x)}{\partial x_1}(x) & \frac{\partial a_m(x)}{\partial x_2}(x) & \cdots & \frac{\partial a_m(x)}{\partial x_n}(x) \end{pmatrix}\begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}+\omega(h)\,. \tag{5}$$

Из определения дифференцируемости функций многих переменных следует, что

$$\frac{\omega_k(h)}{\|h\|}\to 0 \text{ при } \|h\|\to 0 \text{ для всех } k = 1, 2, \dots \text{m}.$$

Тогда

$$\frac{\|\omega(h)\|}{\|h\|}=\frac{1}{\|h\|}\sqrt{\sum_{k=1}^{m}\omega_k^2(h)}=\sqrt{\sum_{k=1}^{m}\left(\frac{\omega_k(h)}{\|h\|}\right)^2}\to 0 \text{ при } \|h\|\to 0\,.$$

Отсюда следует, что линейная часть приращения оператора является главной частью. Кроме того, известно, что матрица – линейный непрерывный оператор. Поэтому первое слагаемое в правой части равенства (5) является дифференциалом Фреше

$$dA(x,h)=A'(x)h\,,$$

а отсюда следует, что производная Фреше есть матрица Якоби (4). **Теорема доказана.**

Пример 3. Производная Фреше функции n переменных.

Пусть f − вещественная функция n переменных, т.е. $f:R^n\to R$. Здесь $m=1$, поэтому матрица Якоби состоит из одной строки, т.е.

$$f'(x)=\left(\frac{\partial f}{\partial x_1}(x) \quad \frac{\partial f}{\partial x_2}(x) \quad \dots \quad \frac{\partial f}{\partial x_n}(x)\right).$$

Итак, производная Фреше функции многих переменных представляет собой матрицу-строку из частных производных.

Выпишем дифференциал Фреше. Для этого представим приращение аргумента в виде матрицы-столбца:

$$\Delta x = \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \\ \\ \Delta x_n \end{pmatrix}.$$

Тогда согласно формуле для дифференциала Фреше:

$$df(x, \Delta x) = f'(x)\Delta x =$$

$$= \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) & \frac{\partial f}{\partial x_2}(x) & \ldots & \frac{\partial f}{\partial x_n}(x) \end{pmatrix} \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \\ \\ \Delta x_n \end{pmatrix} =$$

$$= \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(x)\Delta x_k \ .$$

Дифференциал Фреше совпадает с известным дифференциалом вещественной функции многих переменных.

Пример 4. Рассмотрим оператор $Ax = \begin{pmatrix} x_1^3 - 2x_2 \\ x_2^3 + x_1 \\ 5x_1 - x_2 \end{pmatrix}$, $A : \mathbf{R^2} \to \mathbf{R^3}$.

По формуле (4) имеем

$$A'(x) = \begin{pmatrix} 3x_1^2 & -2 \\ 1 & 3x_2^2 \\ 5 & -1 \end{pmatrix}.$$

Возьмем элемент $x^* = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Тогда $A'(x^*) = \begin{pmatrix} 3 & -2 \\ 1 & 0 \\ 5 & -1 \end{pmatrix}$.

Пример 5. Дифференциал Фреше нелинейного интегрального оператора.

$X = Y = C[a;b]$ – пространство непрерывных на отрезке $[a;b]$ функций.

Рассмотрим нелинейный интегральный оператор Урысона:

$$(Ax)(t) = \int_a^b k(t,s,x(s))ds\,,\ t \in [a;b]\,,$$

где $k:[a;b]\times[a;b]\times\mathbf{R} \to \mathbf{R}$ – заданная функция.

Предположим, что функция k непрерывна, тогда $A: C[a;b] \to C[a;b]$.

Теорема. Пусть функция $k:[a;b]\times[a;b]\times\mathbf{R} \to \mathbf{R}$ непрерывна в области определения, дифференцируема по третьему аргументу, производная по третьему аргументу непрерывна в области определения функции k. Тогда оператор A дифференцируем по Фреше на С$[a;b]$, при этом

$$dA(x,h)(t) = A'(x,h)(t) = \int_a^b \frac{\partial k}{\partial x}(t,s,x(s))h(s)ds\,,\ t \in [a;b]\,.$$

Доказательство изложено, например в книге Л.А.Люстерника и В.И.Соболева [2].

Свойства производной Фреше

Пусть X, Y − нормированные пространства, $G \subset X$, G − открытое множество.

Утверждение 1. Если $Ax = y_0\ \ \forall x \in G$ (т.е. A − оператор-константа), то оператор A дифференцируем по Фреше на G и $A'(x)$ – нулевой оператор ($A'(x)h = \theta\ \ \forall x \in G,\ h \in X$).

Доказательство. Возьмем произвольные $x \in G,\ h \in X$ с условием $x + h \in G$. Тогда $A(x+h) - Ax = y_0 - y_0 = \theta$. Здесь и линейная часть приращения, и остаточный член равны нулю. Условие вида (2), очевидно, выполнено. Поэтому оператор A дифференцируем по Фреше на G, и $A'(x)$ – нулевой оператор ($A'(x)h = \theta\ \ \forall x \in G,\ h \in X$). ***Утверждение доказано.***

Это свойство является обобщением известного утверждения для вещественных функций одной переменной: производная константы равна нулю.

Утверждение 2. Если операторы

$$A_1 : G \to Y \ , \ A_2 : G \to Y$$

дифференцируемы по Фреше в точке $x \in G$, то оператор $A_1 + A_2$ дифференцируем по Фреше в точке x, при этом

$$(A_1 + A_2)'(x) = A_1'(x) + A_2'(x) \, .$$

Идея доказательства. Это утверждение почти очевидно. Для доказательства приращение оператора $A_1 + A_2$ надо разбить на сумму приращений каждого из двух операторов, из этой суммы выделить остаточный член для суммы операторов и показать, что выполнено условие вида (2).

Утверждение 3 (о производной композиции операторов). Пусть X, Y, Z – нормированные пространства, U – окрестность точки $x \in X$, V – окрестность точки $y = Bx$, оператор $B : U \to Y$ дифференцируем по Фреше в точке x, оператор $A : V \to Z$ дифференцируем по Фреше в точке $y = Bx$. Тогда оператор $AB : U \to Z$ дифференцируем по Фреше в точке x и справедлива формула

$$(AB)'(x) = A'(Bx) \cdot B'(x) \, . \tag{6}$$

Замечание. Правая часть равенства (6) представляет собой композицию двух линейных операторов.

Доказательство. Возьмем произвольный элемент $h \in X$, и рассмотрим приращение оператора AB, т.е.

$$\Delta AB = (AB)(x+h) - (AB)x \ = \ A(B(x+h)) - A(Bx) \, .$$

Введем обозначение: $\Delta y = B(x+h) - Bx$, тогда

$$B(x+h) = Bx + \Delta y = y + \Delta y \ , \ \text{и} \ \ \Delta AB = A(y + \Delta y) - Ay \, .$$

Воспользуемся дифференцируемостью оператора A в точке y. Получим равенство

$$\Delta AB = A(y + \Delta y) - Ay = \ A'(y)\Delta y + \omega_1(\Delta y)$$

с условием

$$\frac{\|\omega_1(\Delta y)\|}{\|\Delta y\|} \to 0 \quad \text{при} \quad \|\Delta y\| \to 0\,.$$

Теперь используем дифференцируемость оператора B в точке x. Получим равенство

$$\Delta y = B(x+h) - Bx = B'(x)h + \omega_2(h)$$

с условием

$$\frac{\|\omega_2(h)\|}{\|h\|} \to 0 \quad \text{при} \quad \|h\| \to 0\,.$$

Подставим это выражение для Δy в предыдущее равенство. Получим такое равенство:

$$\Delta AB = A'(y)\Delta y + \omega(\Delta y) = A'(y)(B'(x)h + \omega_2(h)) + \omega_1(\Delta y) =$$

$$= A'(y)B'(x)h + A'(y)\omega_2(h) + \omega_1(\Delta y)\,.$$

Введем обозначение: $\omega(h) = A'(y)\omega_2(h) + \omega_1(\Delta y)$.

Тогда $A'(y)B'(x)h + \omega(h)$.

Оператор $A'(y)B'(x)$ линеен и непрерывен как композиция линейных непрерывных операторов. Остаточный член $\omega(h)$ удовлетворяет условию

$$\frac{\|\omega(h)\|}{\|h\|} \to 0 \quad \text{при} \quad \|h\| \to 0\,,$$

что легко доказывается. Поэтому $A'(y)B'(x)h$ − главная линейная часть приращения ΔAB, т.е. это дифференциал Фреше оператора AB. Поэтому

$$d(AB)(x,h) = A'(y)B'(x)h = A'(Bx)B'(x)h\,,$$

откуда следует равенство (6). ***Утверждение доказано.***

Равенство (6) представляет собой обобщение известного правила дифференцирования сложных функций.

Утверждение 4. Если $L : X \to Y$ – линейный непрерывный оператор, то оператор L дифференцируем по Фреше на X и $L'(x) = L$ для любого $x \in X$.

Доказательство. Зафиксируем произвольный элемент $x \in X$. Тогда для любого $h \in X$ выполняется равенство

$$L(x+h) - Lx = Lx + Lh - Lx = Lh .$$

Здесь Lh − линейная часть приращения, а остаточный член $\omega = \theta$. Условие вида (2) для него, очевидно, выполняется, поэтому $dL(x,h) = Lh$, а отсюда следует равенство $L'(x) = L$. ***Утверждение доказано.***

Утверждение 5. Пусть X, Y, Z – нормированные пространства, G − открытое множество в пространстве X , $A : G \to Y$, $L : Y \to Z$ – линейный непрерывный оператор, оператор A дифференцируем по Фреше в точке $x \in G$.

Тогда оператор $LA : U \to Z$ дифференцируем по Фреше в точке x , и справедлива формула

$$(LA)'(x) = LA'(x) .$$

(символ линейного оператора выносится за знак производной, как постоянный множитель).

Доказательство. Это равенство вытекает из утверждений 3 и 4.

$$(LA)'(x) = L'(Ax)A'(x) = LA'(x) .$$

Утверждение доказано.

Рассмотрим пример с использованием свойств производной Фреше.

Задача. Даны операторы $A, B : R^2 \to R^2$, дифференцируемые по Фреше в точке $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Найти производную Фреше композиции AB в точке x двумя способами:

1. Сначала найти ABx (представить в виде матрицы-столбца вида (3)), затем найти $(AB)'(x)$ по известному правилу (4).

2. Сначала найти $A'(x)$, затем $A'(Bx)$, $B'(x)$ и, наконец, $(AB)'(x)$ по правилу дифференцирования композиции (утверждение 3, равенство (6)).

Решим задачу для операторов A, B, определенных равенствами

$$Ax = \begin{pmatrix} x_1^2 \\ x_1 + 2x_2 \end{pmatrix}, \quad Bx = \begin{pmatrix} 2x_1 \\ x_2^2 \end{pmatrix}.$$

.

Решение

1) Для нахождения $ABx = A(Bx)$ в выражении для Ax на место x_1 ставим первую координату вектора (столбца) B, т.е. $2x_1$, а на место x_2 ставим вторую координату вектора B, т.е. x_2^2.

Получаем

$$ABx = A(Bx) = \begin{pmatrix} 4x_1^2 \\ 2x_1 + 2x_2^2 \end{pmatrix},$$

затем по равенству (4)

$$(AB)'(x) = \begin{pmatrix} 8x_1 & 0 \\ 2 & 4x_2 \end{pmatrix}.$$

2) $A'(x) = \begin{pmatrix} 2x_1 & 0 \\ 1 & 2 \end{pmatrix}$, $A'(Bx) = \begin{pmatrix} 4x_1 & 0 \\ 1 & 2 \end{pmatrix}$, $B'(x) = \begin{pmatrix} 2 & 0 \\ 0 & 2x_2 \end{pmatrix}$,

$$(AB)'(x) = A'(Bx)B'(x) = \begin{pmatrix} 4x_1 & 0 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & 2x_2 \end{pmatrix} = \begin{pmatrix} 8x_1 & 0 \\ 2 & 4x_2 \end{pmatrix}.$$

Результаты решения двумя способами совпадают.

Задачи для самостоятельного решения

Решить предыдущую задачу для операторов, определенных такими равенствами:

1. $Ax = \begin{pmatrix} x_1^2 \\ x_1 - x_2 \end{pmatrix}$, $Bx = \begin{pmatrix} x_2^2 \\ x_1 \end{pmatrix}$.

2. $$Ax = \begin{pmatrix} x_2^2 \\ x_1 - x_2 \end{pmatrix}, \quad Bx = \begin{pmatrix} x_2^2 \\ x_1 \end{pmatrix}.$$

3. $$Ax = \begin{pmatrix} x_1 + x_2 \\ x_1 - x_2 \end{pmatrix}, \quad Bx = \begin{pmatrix} x_2^2 \\ x_1 \end{pmatrix}.$$

4. $$Ax = \begin{pmatrix} x_1 + x_2 \\ x_1^2 \end{pmatrix}, \quad Bx = \begin{pmatrix} x_2^2 \\ x_1 \end{pmatrix}.$$

2.2. Дифференциал Гато. Производная Гато

Пусть X, Y – нормированные пространства, $x \in X$, U – окрестность точки x.

Определение. Оператор $A : U \to Y$ будем называть ***дифференцируемым по Гато*** в точке $x \in U$, если для любого $h \in X$ существует предел $\lim\limits_{t\to 0} \dfrac{A(x+th) - Ax}{t}$ ($t \in \mathbf{R}$); этот предел называется ***дифференциалом Гато*** оператора A в точке x.

Дифференциал Гато оператора A в точке x с приращением h обозначают символом $DA(x,h)$ (в отличие от дифференциала Фреше, который обозначается так: $dA(x,h)$).

Отметим отличие свойств дифференциала Гато от свойств дифференциала Фреше. Оператор $dA(x,\cdot) = A'(x)$ линейный и непрерывный по определению. Напротив, оператор $DA(x,\cdot)$ может не быть линейным.

Установим связь дифференциала Гато с производной по направлению. Пусть f − вещественная функция n переменных, и при некоторых $x = (x_1, x_2, \ldots, x_n)$, $h = (h_1, h_2, \ldots, h_n)$, $\|h\| = 1$ существует предел

$$\lim_{t\to 0} \frac{f(x+th) - f(x)}{t} \quad (t \in \mathbf{R}).$$

Введем обозначение: $\varphi(t)=f(x+th)$. Зафиксируем x,h. Тогда φ − вещественная функция одной переменной. При этом

$$\lim_{t\to 0}\frac{f(x+th)-f(x)}{t}=\lim_{t\to 0}\frac{\varphi(t)-\varphi(0)}{t}=\varphi'(0)=$$

$$=\frac{\partial}{\partial t}f(x+th)\Big|_{t=0}=\sum_{k=1}^{n}\frac{\partial f}{\partial x_k}(x)h_k,$$

а это производная функции f по направлению h при $\|h\|=1$.

Если $\|h\|\neq 1$, то производная по направлению определяется равенством

$$\frac{\partial f}{\partial h}(x)=\frac{1}{\|h\|}\sum_{k=1}^{n}\frac{\partial f}{\partial x_k}(x)h_k.$$

Тогда получается равенство

$$\frac{\partial f}{\partial h}(x)=\frac{1}{\|h\|}\lim_{t\to 0}\frac{f(x+th)-f(x)}{t},$$

поэтому дифференцируемость по Гато функции f в точке x равносильна существованию последнего предела при всех $h\neq\theta$.

Пусть теперь A − оператор, действующий из нормированного пространства X или из некоторой его части в нормированное пространство Y. По аналогии с предыдущим предел $\dfrac{1}{\|h\|}\lim\limits_{t\to 0}\dfrac{A(x+th)-Ax}{t}$ при фиксированных x,h естественно назвать производной оператора A по направлению h в точке x.

Определение. Пусть X,Y − нормированные пространства, G − открытое выпуклое множество в X. Производной оператора $A:G\to Y$ в точке x по направлению $h\in X$ назовем элемент

$$\frac{\partial A}{\partial h}(x) = \frac{1}{\|h\|}\lim_{t\to 0}\frac{A(x+th)-A(x)}{t}.$$

С учетом определения дифференциала Гато $\lim\limits_{t\to 0}\dfrac{A(x+th)-Ax}{t} = DA(x,h)$ при всех $h\in X$ из этого соотношения получаем следующее равенство:

$$\frac{\partial A}{\partial h}(x) = \frac{1}{\|h\|}DA(x,h).$$

Поэтому можно сказать, что дифференцируемость по Гато оператора A в точке x равносильна дифференцируемости оператора A в точке x по всем направлениям $h\in X$.

Пример

$A:\mathbf{R^2}\to\mathbf{R}$, $x=(x_1,x_2)\in \mathrm{R}^2$, $Ax=\sqrt[3]{x_1}\cdot\sqrt[3]{x_2^2}$.

Найдем дифференциал Гато $DA(\theta,h)$.

Решение

По определению

$$DA(\theta,h) = \lim_{t\to 0}\frac{A(\theta+th)-A\theta}{t} = \lim_{t\to 0}\frac{A(th)}{t} =$$

$$=\lim_{t\to 0}\frac{\sqrt[3]{th_1}\cdot\sqrt[3]{t^2h_2^2}}{t} = \sqrt[3]{h_1}\cdot\sqrt[3]{h_2^2}\ \text{ – нелинейное отображение.}$$

Отсюда следует, что, $DA(\theta,\cdot):R^2\to R$ – нелинейный оператор.

Определение. Если $DA(x,\cdot):X\to Y$ – линейный непрерывный оператор ($X,Y-$ нормированные пространства), то он называется ***производной Гато*** оператора A в точке x.

Приведенный выше пример показывает, что оператор A может быть дифференцируемым по Гато, но не иметь производной Гато.

Обозначим символом $A_c'(x)$ производную Гато оператора A в точке x (обозначение из книги А.Н.Колмогорова и С.В.Фомина [1]).

Если производная Гато $A_c'(x)$ существует, то дифференциал Гато связан с производной Гато равенством:

$$DA(x,h) = A_c'(x)h .$$

Установим теперь связь дифференцируемости по Гато с дмфференцируемостью по Фреше.

Пусть X, Y – нормированные пространства, $x \in X$, U – окрестность точки x, $A : U \to Y$.

Теорема. Пусть оператор $A : U \to Y$ дифференцируем по Фреше в точке $x \in U$. Тогда существует производная Гато оператора A в точке x, которая совпадает с производной Фреше ($A_c'(x) = A'(x)$).

Доказательство. Зафиксируем произвольный элемент $h \in X$. Возьмем произвольное $t \neq 0$. Так как оператор A дифференцируем по Фреше в точке x, то

$$\frac{A(x+th) - Ax}{t} = = \frac{A'(x)(th) + \omega(th)}{t} = A'(x)h + \frac{\omega(th)}{t} .$$

При этом

$$\left\| \frac{\omega(th)}{t} \right\| = \frac{\|\omega(th)\|}{|t|} = \frac{\|\omega(th)\|}{|t|\|h\|}\|h\| = \frac{\|\omega(th)\|}{\|th\|}\|h\| \to 0$$

при $\|th\| \to 0$, или при $t \to 0$.

Поэтому существует

$$DA(x,h) = \lim_{t \to 0} \frac{A(x+th) - Ax}{t} = A'(x)h .$$

Так как $A'(x)$ − линейный непрерывный оператор, то оператор $DA(x,\cdot)$ тоже линеен и непрерывен. Поэтому $DA(x,\cdot) = A_c'(x)$ − производная Гато. При этом $A_c'(x) = A'(x)$. ***Теорема доказана***.

Итак, из дифференцируемости по Фреше следует дифференцируемость по Гато; при этом существует и производная Гато. Примеры показывают, что обратное неверно.

Пример. $A:\mathbf{R}^2 \to \mathbf{R}$, $x = < x_1, x_2 > \in \mathbf{R}^2$, $Ax = \sqrt[3]{x_1} \cdot \sqrt[3]{x_2^2}$.

Выше показано, что $DA(\theta, h) = \sqrt[3]{h_1} \cdot \sqrt[3]{h_2^2}$ – это оператор дифференцирования по Гато в точке θ. Покажем, что оператор A не дифференцируем по Фреше в точке θ.

Предположим, что оператор A дифференцируем по Фреше в точке θ. Тогда $A'(\theta)$ – линейный непрерывный оператор. По теореме оператор A имеет в точке θ производную Гато $A'_c(\theta) = A'(\theta)$. Тогда $DA(\theta,\cdot) = A'_c(\theta)$ – линейный непрерывный оператор. Но $DA(\theta, h) = \sqrt[3]{h_1} \cdot \sqrt[3]{h_2^2}$ не является линейным по h, поэтому $DA(\theta,\cdot)$ не является линейным. Получили противоречие, следовательно, A не дифференцируем по Фреше в точке θ.

2.3. Теорема Хана-Банаха и следствие из нее

Определение. **Функционалом** называется отображение $F: X \to \mathbf{P}$, где $\mathbf{P} = \mathbf{R}$ или $\mathbf{P} = \mathbf{C}$ (оператор, значениями которого являются числа).

Теорема Хана-Банаха (один из вариантов).

Пусть X – нормированное пространство, X_0 – подпространство пространства X, f_0 – линейный непрерывный функционал, определенный на X_0.

Тогда функционал f_0 можно продолжить на все пространство X с сохранением нормы, т.е. существует линейный непрерывный функционал f, определенный на X, такой, что $f(x) = f_0(x)$ при $x \in X_0$ и $\|f\| = \|f_0\|$.

Следствие. Пусть X – вещественное нормированное пространство, $X \neq \{\theta\}$. Тогда для любого $x_0 \in X$, $x_0 \neq \theta$ существует линейный непрерывный функционал $f: X \to \mathbf{R}$, для которого $\|f\| = 1$ и $f(x_0) = \|x_0\|$.

Доказательства теоремы и следствия изложены, например, в книге А.Н.Колмогорова и С.В.Фомина [1] (гл. 4, §1, п. 3).

Следствие из теоремы Хана – Банаха нам понадобится в следующем разделе.

2.4. Формулы конечных приращений

Для вещественных функций известна формула Лагранжа (формула конечных приращений):

$$f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1), \ x_1 < \xi < x_2 .$$

Обобщим эту формулу на случай операторов в нормированном пространстве.

Напомним определение отрезка прямой в линейном пространстве.

Пусть X − линейное пространство.

Определение. Отрезком прямой, соединяющим точки x_0, x_1 в линейном пространстве X, называется множество $[x_0, x_1] =$ $= \{x \in X: \ x = (1-\alpha)x_0 + \alpha x_1, \ 0 \le \alpha \le 1\}$.

Определение. Множество $M \subset X$ называется ***выпуклым***, если из условия $x_0, x_1 \in M$ следует включение $[x_0, x_1] \subset M$.

Пусть X, Y – вещественные нормированные пространства, M – выпуклое множество в пространстве X, оператор $A: M \to Y$ имеет производную Гато на множестве M, f – линейный непрерывный функционал, определенный на Y.

Зафиксируем точку $x \in M$ и $h \in X$ так, что $x + h \in M$.

Рассмотрим функцию $\varphi: [0;1] \to \mathbf{R}$,

$$\varphi(t) = f(A(x + th), \quad t \in [0;1]. \qquad (1)$$

Утверждение 1. Если оператор A имеет производную Гато $A'_c(x)$ для любого $x \in M$, то функция φ, определенная формулой (1), дифференцируема на отрезке $[0;1]$ и

$$\varphi'(t) = f(A'_c(x+th)h), \; t \in [0;1]. \qquad (2)$$

Доказательство. Зафиксируем произвольное $t \in [0;1]$. Возьмем любое Δt.

Тогда согласно определению производной

$$\varphi'(t) = \lim_{\Delta t \to 0} \frac{\varphi(t+\Delta t) - \varphi)t)}{\Delta t} = \lim_{\Delta t \to 0} \frac{f(A(x+(t+\Delta t)h) - A(x+th))}{\Delta t} =$$

$$= f(\lim_{\Delta t \to 0} \frac{A(x+(t+\Delta t)h) - A(x+th)}{\Delta t}) = f(A'_c(x+th)h).$$

Утверждение доказано.

Теорема (о формуле конечных приращений).

Пусть X, Y – вещественные нормированные пространства, M – выпуклое множество в пространстве X, оператор $A: M \to Y$ имеет производную Гато на множестве M.

Тогда

1) для любых $x_0, x_1 \in M$ и для любого линейного непрерывного функционала $f: Y \to \mathbf{R}$ существует такая точка $x^* \in [x_0; x_1]$, что

$$f(Ax_1) - f(Ax_0) = f(A'_c(x^*)(x_1 - x_0)) \qquad (3)$$

(первая формула конечных приращений);

2) для любых $x_0, x_1 \in M$

$$\|Ax_1 - Ax_0\| \le \sup_{x \in [x_0; x_1]} \|A'_c(x)\| \cdot \|x_1 - x_0\| \qquad (4)$$

(вторая формула конечных приращений).

Доказательство

1. Выведем первую формулу конечных приращений. Возьмем произвольные $x_0, x_1 \in M$. Пусть $f: Y \to R$ − линейный непрерывный

функционал. Рассмотрим функцию $\varphi:[0;1]\to\mathbf{R}$, определенную равенством

$$\varphi(t)=f(A(x_0+th),\quad t\in[0;1],$$

где $h=x_1-x_0$. Тогда $f(Ax_1)-f(Ax_0)=\varphi(1)-\varphi(0)$.

По теореме Лагранжа найдется такая точка $\xi\in(0;1)$, что

$$\varphi(1)-\varphi(0)=\varphi'(\xi)(1-0)=\varphi'(\xi).$$

По предыдущему утверждению $\varphi'(\xi)=f(A_c'(x_0+\xi h)h)$. В итоге получается равенство

$$f(Ax_1)-f(Ax_0)=\ \varphi'(\xi)=f(A_c'(x_0+\xi h)h).$$

Очевидно, $x_0+\xi h=x_0+\xi(x_1-x_0)\in[x_0,x_1]$.

Введем обозначение: $x^*=x_0+\xi h$.

Тогда $f(Ax_1)-f(Ax_0)=\ f(A_c'(x^*)(x_1-x_0))$, т.е. выполнено равенство (3). Первая формула конечных приращений доказана.

2. Выведем вторую формулу конечных приращений. Из первой формулы конечных приращений следует, что для любого линейного непрерывного функционала $f:Y\to\mathbf{R}$ существует такая точка $x^*\in[x_0;x_1]$, что

$$f(Ax_1-Ax_0)=\ f(Ax_1)-f(Ax_0)=f(A_c'(x^*)(x_1-x_0)).$$

По следствию из теоремы Хана-Банаха найдется такой линейный непрерывный функционал $f:Y\to\mathbf{R}$, что

$$\|f\|=1\quad\text{и}\quad f(Ax_1-Ax_0)=\ \|Ax_1-Ax_0\|.$$

Тогда для такого функционала будет выполнено равенство

$$\|Ax_1-Ax_0\|\ =\ f(A_c'(x^*)(x-x_0)).$$

Из него следует такая оценка:

$$\|Ax_1 - Ax_0\| \le \left|f(A_c'(x^*)(x_1 - x_0)\right| \le \|f\| \cdot \|A_c'(x^*)(x_1 - x_0)\| \le$$

$$\le \|A_c'(x^*)\| \cdot \|x_1 - x_0\| \le \sup_{x \in [x_0, x_1]} \|A_c'(x)\| \cdot \|x_1 - x_0\| .$$

Получилась вторая формула конечных приращений. ***Теорема доказана.***

2.5. Полилинейные операторы

Пусть $X_1, X_2, \ldots, X_n, Y$ – линейные пространства.

Определение. Оператор $A: X_1 \times X_2 \times \ldots \times X_n \to Y$ называется ***n-линейным***, если все операторы вида

$$A(x_1, x_2, \ldots, x_{k-1}, \cdot, x_{k+1}, \ldots, x_n) : X_k \to Y, \ k = \overline{1, n}$$

являются линейными (все x_j, $j \ne k$ – фиксированные элементы).

Говорят, что оператор A линеен по каждому аргументу.

Если не требуется указывать количество аргументов, то оператор A называют ***полилинейным***.

Если оператор n-линейный и $n = 2$, то такой оператор называется ***билинейным***.

Пример

1. Скалярное произведение в $\mathbf{R}^3$.

$x = (x_1, x_2, x_3)$, $y = (y_1, y_2, y_3)$.

$(x, y) = x_1 y_1 + x_2 y_2 + x_3 y_3$.

Это билинейный оператор, действующий из $\mathbf{R}^3 \times \mathbf{R}^3$ в $\mathbf{R}$.

2. Пусть k – непрерывная функция.

Тогда $A(x_1, x_2)(t) = \int_a^b k(t, s) x_1(s) x_2(s) ds$ – билинейный оператор, действующий из $C[a; b] \times C[a; b]$ в $C[a; b]$.

2.6. Степенные операторы

Пусть X, Y – нормированные пространства.

Определение. Оператор $B: X \to Y$ называется оператором n-й степени, если существует такой n-линейный оператор $A: X^n \to Y$, что оператор B можно представить в виде $Bx = A(\underbrace{x, x, \ldots, x}_{n})$.

Замечание. Если B − оператор n-й степени, то часто вместо Bx пишут Bx^n. Здесь x^n − символ, показывающий, что B − оператор n-й степени.

Пример. Рассмотрим оператор

$$A(x_1, x_2, x_3)(t) = \int_a^b k(t,s)x_1(s)x_2(s)x_3(s)ds,$$

$A: (C[a;b])^3 \to C[a;b]$, оператор A – 3-линейный оператор.

Тогда оператор $(Bx)(t) = \int_a^b k(t,s)x^3(s)ds$ – оператор третьей степени.

Если $B: X \to Y$ – оператор n-й степени, то его норма определяется формулой

$$\|B\| = \{\min M \geq 0 : \|Bx^n\| \leq M\|x\|^n \ \forall x \in X\}.$$

2.7. Дифференциалы Фреше второго и более высоких порядков

Пусть X, Y – нормированные пространства, U – окрестность точки $x \in X$, оператор $A: U \to Y$ дифференцируем по Фреше в U.

Тогда $\forall x \in U, h \in X$ существует $dA(x,h)$. При этом $A'(x) = dA(x, \cdot)$ – линейный непрерывный оператор.

Рассмотрим оператор $dA(\cdot, h): U \to Y$, где h – фиксированный элемент.

Пусть этот оператор дифференцируем по Фреше в U. Тогда существует дифференциал Фреше этого оператора $d(dA(x,h),g)$, где g – приращение аргумента.

Оператор $dA(x,h)$ линеен и непрерывен по h. Тогда оператор $d(dA(x,h),g)$ линеен и непрерывен по h.

По определению дифференциала Фреше оператор $d(dA(x,h),g)$ линеен и непрерывен по g.

Более точно, оператор $d(dA(x,\cdot),g): X \to Y$ линеен и непрерывен; оператор $d(dA(x,h),\cdot): X \to Y$ линеен и непрерывен; оператор $d(dA(x,\cdot),\cdot): X^2 \to Y$ билинеен и непрерывен.

Определение. Вторым дифференциалом Фреше оператора A в точке x с приращением $h \in X$ называется элемент $d(dA(x,h),h)$.

Заметим, что если x зафиксировать, а h считать аргументом, то получится оператор $d(dA(x,\cdot),\cdot): X^2 \to Y$, который при одинаковых аргументах является оператором второй степени (квадратичным оператором).

Введем обозначение: $d^2 A(x,h)$ – второй дифференциал Фреше оператора A в точке x с приращением h. По определению $d^2 A(x,h) = d(dA(x,h),h)$.

$d^2 A(x,\cdot)$ – квадратичный оператор.

Определение. Второй производной Фреше оператора A в точке x называется оператор $d^2 A(x,\cdot): X \to Y$.

Введем ***обозначения***:

$A''(x)$ – вторая производная Фреше оператора A в точке x;

$A''(x)$ – непрерывный квадратичный оператор;

значение $A''(x)h^2$ – второй дифференциал Фреше, т.е. $A''(x)h^2 = d^2 A(x,h)$ (здесь h^2 – символическое обозначение, которое показывает, что $A''(x)$ – квадратичный оператор).

Аналогично определяются дифференциалы более высоких порядков. Определим третий дифференциал Фреше.

Рассмотрим оператор $d^2 A(\cdot, h)$, где $h \in X$ – фиксированный элемент.

Предположим, что оператор $d^2 A(\cdot, h) : U \to Y$ дифференцируем по Фреше в окрестности U точки x.

Рассмотрим оператор $d(d^2 A(x, h), p)$; он является линейным по p и квадратичным по h.

Определение. Третьим дифференциалом Фреше оператора A в точке x с приращением h называется элемент

$$d^3 A(x, h) = d(d^2 A(x, h), h) \in Y .$$

Аналогично предыдущему $d^3 A(x, \cdot)$ – оператор третьей степени (кубичный оператор).

Тогда для произвольной степени

$$d^{n+1} A(x, h) = d(d^n A(x, h), h) ;$$

$$d^n A(x, h) = A^{(n)}(x) h^n ,$$

где $A^{(n)}$ – n-я производная, $A^{(n)}(x) = d^n A(x, \cdot)$ – непрерывный оператор n-й степени.

Пример. Рассмотрим оператор A, заданный формулой

$$Ax = \begin{pmatrix} x_2^3 - 3x_1 x_3 + 2x_2 \\ x_1^2 - 3x^2 x_3 + 1 \end{pmatrix}.$$

Обозначим $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbf{R}^3$, $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbf{R}^2$; $A : \mathbf{R}^3 \to \mathbf{R}^2$.

Найдем все дифференциалы в оператора A в произвольной точке x с приращением $h = \begin{pmatrix} h_1 \\ h_2 \\ h_3 \end{pmatrix}$.

Решение

Воспользуемся формулами для первого дифференциала и дифференциалов более высоких порядков.

$$dA(x,h)=\begin{pmatrix}-3x_3 & 3x_2^2+2 & -3x_1\\ 2x_1 & -3x_3 & -3x_2\end{pmatrix}\begin{pmatrix}h_1\\ h_2\\ h_3\end{pmatrix}=\begin{pmatrix}-3x_3h_1+\left(3x_2^2+2\right)h_2-3x_1h_3\\ 2x_1h_1-3x_3h_2-3x_2h_3\end{pmatrix},$$

$$d^2A(x,h)=d(dA(x,h),h)=\begin{pmatrix}-3h_3 & 6x_2h_2 & -3h_1\\ 2h_1 & -3h_3 & -3h_2\end{pmatrix}\begin{pmatrix}h_1\\ h_2\\ h_3\end{pmatrix}=$$

$$=\begin{pmatrix}-3h_1h_3+6x_2h_2^2-3h_1h_3\\ 2h_1^2-3h_2h_3-3h_2h_3\end{pmatrix}=\begin{pmatrix}-6h_1h_3+6x_2h_2^2\\ 2h_1^2-6h_2h_{33}\end{pmatrix},$$

$$d^3A(x,h)=d(d^2A(x,h),h)=\begin{pmatrix}0 & 6h_2^2 & 0\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}h_1\\ h_2\\ h_3\end{pmatrix}=\begin{pmatrix}6h_2^3\\ 0\end{pmatrix},$$

$d^4A(x,h)=\theta$ и все $d^nA(x,h)=\theta$ ($n\ge 4$).

2.8. Формула Тейлора

Пусть X,Y – нормированные пространства.

Теорема. Пусть U – окрестность точки $x_0\in X$, оператор $A:U\to Y$ дифференцируем по Фреше n раз, оператор $A^{(n)}(\cdot)$ равномерно непрерывен в окрестности точки x_0. Тогда для любого $h\in X$, такого что $x_0+h\in U$, справедлива формула

$$A(x_0+h)=Ax_0+A'(x_0)h+\frac{1}{2!}A''(x_0)h^2+\ldots+\frac{1}{n!}A^{(n)}(x_0)h^n+\omega_n(x_0,h),\quad(1)$$

где $\omega_n(x_0,h)$ удовлетворяет условию

$$\frac{\left\|\omega_n(x_0,h)\right\|}{\|h\|^n}\to 0 \text{ при } \|h\|\to 0.\quad(2)$$

Доказательство этой теоремы изложено в книге Н.Колмогорова и С.В.Фомина [1] (гл.Х, §1, п. 10).

Определение. Формула (1) при условии (2) называется ***формулой Тейлора***.

2.9. Ряд Тейлора

Пусть X, Y – нормированные пространства, U – окрестность точки $x_0 \in X$, оператор $A: U \to Y$.

Пусть оператор A имеет в U производную Фреше всех порядков; тогда для любого n справедлива формула Тейлора.

Элементу $A(x_0+h)$ можно поставить в соответствие ряд

$$A(x_0+h) = Ax_0 + A'(x_0)h + \frac{1}{2!}A''(x_0)h^2 + \ldots + \frac{1}{n!}A^{(n)}(x_0)h^n + \ldots \,.$$

Или

$$A(x_0+h) = Ax_0 + \sum_{n=1}^{\infty}\frac{1}{n!}A^{(n)}(x_0)h^n \,. \qquad (1)$$

Определение. Правая часть формулы (1) называется ***рядом Тейлора*** оператора A в окрестности точки x_0.

Ряд Тейлора может как сходиться, так и расходиться. Вопрос о сходимости ряда Тейлора требует дополнительных исследований.

Пример

1. Разложим оператор $A: \mathbf{R}^2 \to \mathbf{R}^3$ в ряд Тейлора в окрестности точки x^*.

$$Ax = \begin{pmatrix} x_1^4 - 2x_1x_2 + 3 \\ x_2^2 - 3x_1 + 1 \\ x_1^2 + 1 \end{pmatrix}, \; x^* = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Воспользуемся формулой (1).

$$dA(x,h)=\begin{pmatrix} 4x_1^3 - 2x_2 & -2x_1 \\ -3 & 2x_2 \\ 2x_1 & 0 \end{pmatrix}\begin{pmatrix} h_1 \\ h_2 \end{pmatrix}=\begin{pmatrix} (4x_1^3-2x_2)h_1-2x_1h_2 \\ -3h_1+2x_2h_2 \\ 2x_1h_1 \end{pmatrix}.$$

$$d^2A(x,h)=d(dA(x,h),h)=\begin{pmatrix} 12x_1^2h_1-2h_2 & -2h_1 \\ 0 & 2h_2 \\ 2h_1 & 0 \end{pmatrix}\begin{pmatrix} h_1 \\ h_2 \end{pmatrix}=$$

$$=\begin{pmatrix} (12x_1^2h_1-2h_2)h_1-2h_1h_2 \\ 2h_2^2 \\ 2h_1^2 \end{pmatrix}=$$

$$=\begin{pmatrix} 12x_1^2h_1^2-4h_1h_2 \\ 2h_2^2 \\ 2h_1^2 \end{pmatrix}.$$

$$d^3A(x,h)=\begin{pmatrix} 24x_1h_1^2 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} h_1 \\ h_2 \end{pmatrix}=\begin{pmatrix} 24x_1h_1^3 \\ 0 \\ 0 \end{pmatrix}.$$

$$d^4A(x,h)=\begin{pmatrix} 24h_1^3 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} h_1 \\ h_2 \end{pmatrix}=\begin{pmatrix} 24h_1^4 \\ 0 \\ 0 \end{pmatrix}.$$

$d^5A(x,h)=\theta$ и все $d^nA(x,h)=\theta$ ($n\ge 5$).

Тогда ряд Тейлора оператора A в окрестности точки x^* имеет вид

$$A(x^*+h)=\begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}+\begin{pmatrix} -2h_1 \\ -3h_1+2h_2 \\ 0 \end{pmatrix}+\frac{1}{2!}\begin{pmatrix} -4h_1h_2 \\ 2h_2^2 \\ 2h_1^2 \end{pmatrix}+\frac{1}{3!}\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}+\frac{1}{4!}\begin{pmatrix} 24h_1^4 \\ 0 \\ 0 \end{pmatrix}.$$

2. Найдем значение $A\theta$ непосредственно и с помощью разложения в ряд Тейлора.

Непосредственно подставим точку θ в формулу для оператора A:

$$A\theta = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}.$$

Теперь, если $x^* = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ и $x^* + h = \theta$, то $h = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$. Подставим h в ряд Тейлора оператора A:

$$A\theta = \begin{pmatrix} 3+0+0+0+0 \\ 2-2+1+0+0 \\ 1+0+0+0+0 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}.$$

Результат совпал с предыдущим.

Задачи для самостоятельного решения

Разложить оператор $A: R^2 \to R^2$ в ряд Тейлора в окрестности заданной точки $x*$. Вычислить значение $A\theta$ непосредственно и через разложение в ряд Тейлора.

Варианты задач

1. $Ax = \begin{pmatrix} x_1^3 + 2x_1 + 3 \\ x_1x_2 - x_1 + 1 \end{pmatrix}, \quad x^* = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$

2. $Ax = \begin{pmatrix} x_1^2x_2 - 2x_1x_2 + 1 \\ x_1^2 + 3x_1 + 2 \end{pmatrix}, \quad x^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$

3. $Ax = \begin{pmatrix} x_1^3 + x_1x_2 + 3 \\ x_1x_2 - 3x_1 + 2 \end{pmatrix}, \quad x^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$

4. $$Ax = \begin{pmatrix} x_1 x_2^2 + 4x_2 + 2 \\ x_1^2 - x_1 x_2 + 3 \end{pmatrix}, \qquad x^* = \begin{pmatrix} -1 \\ 0 \end{pmatrix}.$$

2.10. Необходимые условия экстремума функционала

Одна из важнейших задач нелинейного анализа – нахождение точек экстремума вещественных функционалов. В математическом анализе рассматриваются такие задачи для функционалов частного вида – для функций одной или нескольких переменных. Изложим основы теории экстремальных задач для вещественных функций (функционалов), определенных на множествах в нормированных пространствах произвольных размерностей, среди которых могут быть и бесконечномерные пространства.

Пусть X – нормированное пространство, $x_0 \in X$, G – открытое множество в X, $x_0 \in G$.

Рассмотрим вещественный функционал $f : G \to \mathbf{R}$.

Определение. Точка x_0 называется точкой ***локального минимума*** функционала $f : G \to \mathbf{R}$, если существует такая окрестность U точки x_0, что $f(x) \ge f(x_0)$ для любого $x \in U$.

Точка $x_0 \in G$ называется точкой ***локального максимума*** функционала $f : G \to \mathbf{R}$, если существует такая окрестность U точки x_0, что $f(x) \le f(x_0)$ для любого $x \in U$.

Точка x_0 называется точкой ***локального экстремума*** функционала $f : G \to \mathbf{R}$, если x_0 – точка локального минимума или локального максимума.

Теорема 1. Пусть X – нормированное пространство, $G \subset X$ – открытое множество, точка $x_0 \in G$, функционал $f : G \to \mathbf{R}$ дифференцируем по Фреше в точке x_0, x_0 – точка локального экстремума функционала f. Тогда $f'(x_0)$ – нулевой функционал ($f'(x_0)h = 0 \ \forall h \in X$).

Доказательство. Будем рассматривать отдельно 2 случая: x_0 − точка локального минимума функционала f и x_0 − точка локального максимума функционала f.

Случай 1: x_0 − точка локального минимума функционала f.

Пусть U − окрестность точки x_0, в которой выполнено неравенство

$$f(x) \geq f(x_0), \qquad (1),$$

δ − радиус этой окрестности.

Возьмем произвольный элемент $h \in X$, $h \neq \theta$. Рассмотрим точки $x = x_0 + th$, $t \in R$.

Если $|t| < \dfrac{\delta}{\|h\|}$, то $\|x - x_0\| = \|th\| = |t|\|h\| < \delta$ и, следовательно, $x \in U$. Тогда выполнено неравенство (1), или $f(x) - f(x_0) \geq 0$ при $|t| < \dfrac{\delta}{\|h\|}$.

Преобразуем неравенство с применением производной Фреше:

$$0 \leq f(x) - f(x_0) = f(x_0 + th) - f(x_0) = f'(x_0)(th) + \omega(th),$$

причем $\dfrac{\omega(th)}{\|th\|} \to 0$ при $\|th\| \to 0$, а это равносильно условию

$$\frac{\omega(th)}{t} \to 0. \qquad (2)$$

Итак, получилось неравенство

$$f'(x_0)(th) + \omega(th) \geq 0. \qquad (3)$$

Пусть $0 < t < \dfrac{\delta}{\|h\|}$. Разделим неравенство (3) на *t*. Получим такое неравенство:

$$f'(x_0)h + \frac{\omega(th)}{t} \geq 0.$$

Перейдем к пределу при $t \to 0$. Учитывая свойство (2), получим неравенство

$$f'(x_0)h \geq 0 . \qquad (4)$$

Пусть теперь $-\frac{\delta}{\|h\|} < t < 0$. Разделим неравенство (3) на t. Так как $t < 0$, то $f'(x_0)h + \frac{\omega(th)}{t} \leq 0$. После предельного перехода при $t \to 0$ получается неравенство

$$f'(x_0)h \leq 0 . \qquad (5)$$

Из двух последних неравенств (4) и (5) получается равенство

$$f'(x_0)h = 0 . \qquad (6)$$

Это равенство выполняется при любых $h \in X$, $h \neq \theta$, так как такое предположение было в начале доказательства. Равенство выполнено, очевидно, и при $h = \theta$. Итак, равенство (6) выполнено при всех $h \in X$, откуда следует, что $f'(x_0)$ − нулевой функционал. Утверждение теоремы для случая 1 доказано.

Случай 2: x_0 − точка локального максимума функционала f.

Этот случай легко сводится к случаю 1.

Рассмотрим вспомогательную функцию $f_1 = -f$. Тогда x_0 − точка локального минимума функции f_1. Из доказанного в случае 1 следует, что $f_1'(x_0)$ − нулевой функционал. Так как $f'(x_0) =$ =- $f_1'(x_0)$, то $f'(x_0)$ − тоже нулевой функционал. Итак, утверждение теоремы верно и в случае 2. ***Теорема доказана.***

Определение. Точки x, в которых функционал $f'(x)$ является нулевым, называются ***стационарными***.

Замечание. Нулевой функционал часто обозначают нулем θ. Тогда утверждение теоремы можно записать в виде $f'(x_0) = \theta$.

Комментарий. Стационарная точка может быть точкой локального экстремума, а может и не быть. Для того чтобы узнать, есть ли экстремум в данной стационарной точке, нужны дополнительные исследования. Такие исследования могут быть проведены с помощью достаточных условий экстремума.

Пример. Пусть $f:\mathbf{R}^n \to \mathbf{R}$, т.е. f − вещественная функция n переменных.

Пусть x^* − точка локального экстремума функции f, и f дифференцируема по Фреше в точке x^*. Тогда по предыдущей теореме производная Фреше $f'(x^*)$ − нулевой функционал. В этом случае $f'(x^*) = \begin{pmatrix} \dfrac{\partial f}{\partial x_1}(x^*) & \dfrac{\partial f}{\partial x_2}(x^*) & \dots & \dfrac{\partial f}{\partial x_n}(x^*) \end{pmatrix}$ − матрица Якоби, состоящая из одной строки и n столбцов (теорема о производной Фреше конечномерного оператора, п. 2.1).

Так как $f'(x^*)$ − нулевой функционал, то все элементы матрицы Якоби равны 0, т.е. все частные производные функции f в точке x^* равны нулю. Получились необходимые условия экстремума функции n переменных, известные из математического анализа.

Если функция дифференцируема по Фреше дважды, то из того, что точка x_0 − точка локального экстремума функции f, вытекают свойства не только первой, но и второй производной Фреше этой функции. Сформулируем и докажем соответствующую теорему.

Теорема 2. Пусть X – нормированное пространство, $G \subset X$ – открытое множество, $x_0 \in G$, x_0 – точка локального экстремума функции $f:G \to \mathbf{R}$, эта функция дифференцируема по Фреше дважды в некоторой окрестности точки x_0, при этом функция $f''(\cdot)$ равномерно непрерывна в указанной окрестности. Тогда $f'(x_0)$ – нулевой функционал ($f'(x_0)h = 0\ \forall h \in X$).

Если x_0 − точка локального минимума функции f, то $f''(x_0)h \geq 0$ $(\forall h \in X)$.

Если x_0 − точка локального максимума функции f, то $f''(x_0)h \leq 0$ $(\forall h \in X)$.

Доказательство. Будем рассматривать отдельно 2 случая: x_0 − точка локального минимума функционала f и x_0 − точка локального максимума функционала f.

Случай 1: x_0 − точка локального минимума функционала f.

Пусть U − окрестность точки x_0, в которой выполнено неравенство

$$f(x) \geq f(x_0), \tag{7}$$

δ − радиус этой окрестности (согласно определению точки минимума).

Возьмем произвольный элемент $h \in X$, $h \neq \theta$. Рассмотрим точки $x = x_0 + th$, $t \in R$.

Если $|t| < \dfrac{\delta}{\|h\|}$, то $\|x - x_0\| = \|th\| = |t|\|h\| < \delta$ и, следовательно, $x \in U$. Тогда выполнено неравенство (7), или $f(x) - f(x_0) \geq 0$ при $|t| < \dfrac{\delta}{\|h\|}$. Преобразуем неравенство по формуле Тейлора.

$$0 \leq f(x) - f(x_0) = f(x_0 + th) - f(x_0) =$$

$$= f'(x_0)(th) + \frac{1}{2} f''(x_0)(th)^2 + \omega_2(th),$$

причем $\dfrac{\|\omega_2(th)\|}{\|th\|^2} \to 0$ при $\|th\| \to 0$, а это эквивалентно условию

$$\frac{\omega_2(th)}{t^2} \to 0 \qquad \text{при} \qquad t \to 0. \tag{8}$$

Так как $f'(x_0)$ − нулевой функционал, то

$$\frac{1}{2} f''(x_0)(th)^2 + \omega_2(th) \geq 0 \tag{9}$$

при $|t| < \dfrac{\delta}{\|h\|}$.

Разделим неравенство (9) на t^2.

Тогда $\frac{1}{2} f''(x_0)h^2 + \frac{\omega_2(th)}{t^2} \geq 0$. Перейдем к пределу при $t \to 0$ и учтем условие (8). Получим неравенство

$$f''(x_0)h^2 \geq 0 .$$

Так как $h \in X$, $h \neq \theta$, то это неравенство выполняется при всех таких h. При $h = \theta$ оно, очевидно, тоже выполняется. Итак, $f''(x_0)h^2 \geq 0$ при всех $h \in X$. Утверждение теоремы для случая локального минимума доказано.

Случай 2: пусть теперь x_0 − точка локального максимума функционала f. Рассмотрим вспомогательную функцию $f_1 = -f$. Очевидно, x_0 − точка локального минимума функционала $f_1 = -f$. Тогда в соответствии с доказанным в случае 1 выполнено неравенство $-f''(x_0)h^2 \geq 0$ при всех $h \in X$, а это эквивалентно тому, что $f''(x_0)h^2 \leq 0$. ***Теорема доказана.***

В дальнейшем будет видно, что необходимые условия экстремума, изложенные в теореме 2, совсем немного отличаются от доствточных условий экстремума.

Замечание. В теореме 2 рассматриваются функции, определенные на некотором множестве из нормированного пространства произвольного вида. В этом случае требуется равномерная непрерывность $f''(\cdot)$ в некоторой окрестности точки экстремума. Это объясняется тем, что в теореме о формуле Тейлора есть такое условие (А.Н.Колмогоров, С.В.Фомин).

Если же рассматриваются вещественные функции n переменных, то это условие можно заменить другим, менее жестким. Достаточно потребовать, чтобы функция была дифференцируема в некоторой окрестности точки экстремума $n-1$ раз, а n-я производная существовала только в самой точке экстремума. Таковы условия в теореме о формуле Тейлора для функций многих переменных, что изложено в учебниках по математическому анализу.

2.11. Достаточные условия экстремума функционала

Пусть X – нормированное пространство.

Определение. Квадратичный функционал $\varphi : X \to \mathbf{R}$ будем называть ***неотрицательным***, если для любого $x \in X$ выполняется неравенство

$$\varphi(x) \geq 0. \tag{1}$$

Определение. Квадратичный функционал $\varphi : X \to \mathbf{R}$ будем называть ***положительным***, если для любого $x \in X$, $x \neq \theta$ выполняется неравенство

$$\varphi(x) > 0. \tag{2}$$

Определение. Квадратичный функционал $\varphi : X \to \mathbf{R}$ будем называть ***сильно положительным***, если существует такое число $c > 0$, что для любого $x \in X$ выполняется неравенство

$$\varphi(x) \geq c\|x\|^2. \tag{3}$$

Теорема 1. Пусть выполнены условия:

1) X – нормированное пространство, $x_0 \in X$, U – окрестность точки x_0;

2) функционал $f : U \to \mathbf{R}$ дифференцируем в окрестности точки x_0 и $f'(x_0)$ – нулевой функционал ($f'(x_0) = \theta$);

3) в окрестности U точки x_0 существует $f''(x)$, причем оператор $f''(\cdot)$ равномерно непрерывен на окрестности U.

Тогда если $f''(x_0)$ – сильно положительный квадратичный функционал, то x_0 – точка локального минимума функционала f.

Если же $-f''(x_0)$ – сильно положительный квадратичный функционал, то x_0 – точка локального максимума функционала f.

Доказательство. Будем рассматривать отдельно 2 случая: $f''(x_0)$ – сильно положительный квадратичный функционал и $-f''(x_0)$ – сильно положительный квадратичный функционал.

Случай 1. $f''(x_0)$ – сильно положительный квадратичный функционал.

Возьмем произвольный элемент $x \in U$, $x \neq x_0$. Введем обозначение: $h = x - x_0$. Преобразуем разность $f(x) - f(x_0)$ по формуле Тейлора:

$$f(x) - f(x_0) = f(x_0 + h) - f(x_0) = f'(x_0)h + \frac{1}{2}f''(x_0)h^2 + \omega_2(h),$$

причем

$$\frac{\omega_2(h)}{\|h\|^2} \to 0 \quad \text{при} \quad \|h\| \to 0 \tag{4}$$

Так как $f'(x_0)$ − нулевой функционал, то

$$f(x) - f(x_0) = \frac{1}{2}f''(x_0)h^2 + \omega_2(h) \geq \frac{1}{2}f''(x_0)h^2 - |\omega_2(h)|. \tag{5}$$

Так как $f''(x_0)$ – сильно положительный квадратичный функционал, то существует такая константа $c > 0$, что $f''(x_0)h^2 \geq c\|h\|^2$ для любых $h \in X$. Тогда

$$f(x) - f(x_0) \geq \frac{c}{2}\|h\|^2 - |\omega_2(h)| = \frac{c}{2}\|h\|^2 - \frac{|\omega_2(h)|}{\|h\|^2}\|h\|^2.$$

Так как выполняется условие (4), то найдется такое $\delta > 0$, что $\frac{|\omega_2(h)|}{\|h\|^2} < \frac{c}{2}$ при $\|h\| < \delta \iff \|x - x_0\| < \delta$. Отсюда и из соотношений (5) следует неравенство $f(x) - f(x_0) > 0$ при всех x из δ-окрестности точки x_0. Поэтому x_0 − точка локального минимума функционала f.

Утверждение теоремы для случая локального минимума доказано.

Случай 2. Пусть теперь $-f''(x_0)$ – сильно положительный квадратичный функционал.

Рассмотрим вспомогательную функцию $f_1 = -f$. Тогда $f_1''(x)$ $-$ сильно положительный квадратичный функционал. По доказанному утверждению в случае 1 x_0 $-$ точка локального минимума функционала $f_1 = -f$. Отсюда следует, что x_0 - точка локального максимума функционала $f = -f_1$. ***Теорема доказана.***

Комментарий. Сравнение необходимых и достаточных условий экстремума показывает, что необходимые условия экстремума, изложенные в соответствующем разделе, совсем немного отличаются от достаточных условий экстремума.

В достаточных условиях требуется сильная положительность второй производной Фреше, а в необходимые условия входит неотрицательность второй производной, т.е. менее жесткое требование.

Замечание 1. В доказанной теореме рассматриваются функции, определенные на некотором множестве из нормированного пространства произвольного вида. В этом случае требуется равномерная непрерывность $f''(\cdot)$ в некоторой окрестности точки экстремума. Это объясняется тем, что в теореме о формуле Тейлора есть такое условие (А.Н.Колмогоров, С.В.Фомин).

Если же рассматриваются вещественные функции n переменных, то это условие можно заменить другим, менее жестким. Достаточно потребовать, чтобы функция была дифференцируема в некоторой окрестности точки экстремума $n-1$ раз, а n-я производная существовала только в самой точке экстремума.

Таковы условия в теореме о формуле Тейлора для функций многих переменных, что изложено в учебниках по математическому анализу.

Замечание 2. В достаточных условиях экстремума для функций n переменных вместо сильной положительности второй производной достаточно потребовать ее положительности в следующем смысле: $f''(x_0)h > 0$ при $h \neq \theta$. Это объясняется тем, что положительность и сильная положительность квадратичного функционала, определенного на конечномерном пространстве, эквивалентны.

Кроме достаточных условий экстремума, изложенных в теореме 1 этого раздела, известны и другие условия, которые получены с помощью выпуклого анализа.

Сначала изложим без доказательства теорему о достаточных условиях экстремума для случая, когда вещественная функция определена на некотором множестве из нормированного пространства произвольной размерности (конечной или бесконечной).

Затем рассмотрим частный случай, когда функция определена на некотором множестве конечномерного пространства, т.е. когда это функция конечного числа переменных.

Теорема 2. Пусть выполнены условия:

1) X – нормированное пространство, $x_0 \in X$, U – окрестность точки x_0;

2) функционал $f : U \to \mathbf{R}$ дифференцируем в окрестности точки x_0 и $f'(x_0)$ – нулевой функционал ($f'(x_0) = \theta$);

3) в окрестности U точки x_0 существует $f''(x)$, причем оператор $f''(\cdot)$ равномерно непрерывен на окрестности U.

Тогда если для всех x из некоторой окрестности точки x_0

$$d^2 f(x,h) \geq 0 \quad (\forall h \in X),$$

то x_0 − точка локального миимума функционала f.

Если

$$d^2 f(x,h) \leq 0 \quad (\forall h \in X),$$

то x_0 − точка локального максимума функционала f.

Обратим внимание на различия в условиях теорем 1 и 2. Так, для локального минимума в теореме 1 требуется сильная положительность второй производной в стационарной точке, а в теореме 2 требуется лишь неотрицательность второй производной, но не в одной точке, а в некоторой окретности стационарной точки. Похожие различия в этих теоремах и для локального максимума.

Теперь рассмотрим случай, когда f − функция n переменных. В этом случае теоремы 1 и 2 тоже верны, только условие 3) теорем можно заменить менее жестким: достаточно потребовать существование f'' только в одной точке x_0.

Так как в конечномерном пространстве обычно нижние индексы используют для обозначения координат точки, то далее точку экстремума будем обозначать символом x^*.

В условиях теоремы 1 в конечномерном пространстве R^n для локального минимума функционала f в точке x^* с учетом замечания 2 к этой теореме можно потребовать лишь положительность, а не сильную положительность функционала $f''(x^*)$, что эквивалентно условию $d^2 f(x^*, h) > 0$ при $h \in R^n$, $h \neq \theta$. Известно, что

$$d^2 f(x^*, h) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x^*) h_i h_j ,$$

т.е. это квадратичная форма по h. А положительность такой формы устанавливается с помощью критерия Сильвестра. Напомним его.

Рассмотрим квадратичную форму $\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} h_i h_j$ с симметричной матрицей $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$. Нам понадобятся так называемые главные миноры матрицы A.

Введем ***обозначение***: $A_{i_1 i_2 \dots i_m}$ − минор, состоящий из всех элементов строк и столбцов с номерами $i_1, i_2, \dots, i_m$.

Например, если

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix},$$

то

$$A_1 = 1,\ A_2 = 5,\ A_3 = 9,\ A_{12} = \begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix}, A_{13} = \begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix},$$

$$A_{23}=\begin{vmatrix}5 & 6\\ 8 & 9\end{vmatrix},\quad A_{123}=\begin{vmatrix}1 & 2 & 3\\ 4 & 5 & 6\\ 7 & 8 & 9\end{vmatrix}.$$

Выпишем сначала $d^2 f$ в произвольной точке x:

$$d^2 f(x,h)=\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 f}{\partial x_i \partial x_j}(x)h_i h_j .$$

Матрица этой квадратичной формы называется матрицей Гессе функции f в точке x. Эта матрица имеет вид

$$A(x)=\begin{pmatrix}a_{11} & \dots & a_{1n}\\ \dots & \dots & \dots\\ a_{n1} & \dots & a_{nn}\end{pmatrix},\ \text{где}\ \ a_{ij}=\frac{\partial^2 f}{\partial x_i \partial x_j}(x) .$$

Для установления положительности квадратичной формы нужны не все главные миноры, а только так называемые последовательные миноры $\Delta_1 = A_1$, $\Delta_2 = A_{12}$, $\Delta_3 = A_{123}$ и т.д.

Критерий Сильвестра утверждает, что для положительности квадратичной формы необходимо и достаточно, чтобы все последовательные главные миноры ее матрицы были положительными.

Отсюда следует, что для отрицательности квадратичной формы необходимо и достаточно, чтобы все последовательные главные миноры нечетного порядка были отрицательными, а миноры четного порядка были положительными.

Тогда из теоремы 1 следует, что если все последовательные главные миноры матрицы Гессе в точке x^* положительны при выполнении условий 1) – 3) теоремы 1, то x^*- точка минимума функции f.

Выведем теперь достаточные условия экстремума функции n переменных, которые получаются на основе выпуклого анализа. Эти условия получим из теоремы 2 и замечания 1 к ней.

Из этой теоремы следует, что для локального минимума при условии выполнения условий 1) – 3) достаточно, чтобы 2-я производная функции f была неотрицательной, а для максимума – неположительной. Но здесь требуется выполнение этих условий не в одной точке, а в некоторой окрестности точки x^*.

Неотрицательность соответствующей квадратичной формы можно установить с помощью другого варианта критерия Сильвестра, который утверждает , что для неотрицательности квадратичной формы необходимо и достаточно, чтобы все главные миноры матрицы, а не только последовательные, были неотрицательными (Галлеев Э.М., Тихомиров В.М. Оптимизация). Отсюда следует, что для неположительности квадратичной формы необходимо и достаточно, чтобы все главные миноры матрицы нечетного порядка были неположительными, а все главные миноры четного порядка были неотрицательными.

Отсюда, а также из теоремы 2 и замечания 1 к теореме 1 получаются новые достаточные условия экстремума.

***Теорема 3*.** Пусть выполнены условия:

1) $x^* \in \mathbf{R}^n$, U – окрестность точки x^*;

2) $f : U \to \mathbf{R}$, функция f дифференцируема по Фреше в окрестности U точки x^*, и $f'(x_0)$ – нулевой функционал;

3) в окрестности U точки x^* существует $f''(x)$, причем оператор $f''(\cdot)$ равномерно непрерывен на окрестности U.

Тогда, если для всех x из некоторой окрестности точки x^* все главные миноры матрицы Гессе функции f неотрицательны, то x^* − точка локального минимума функции f.

Если же для всех x из некоторой окрестности точки x^* все главные миноры нечетного порядка матрицы Гессе функции f неположительны, а все главные миноры четного порядка неотрицательны, то x^* − точка локального максимума функции f.

Пусть f − вещественная функция многих переменных. Для нахождения точек экстремума − сначала те же действия, что в алгоритме, который мы уже изучали. Находим все частные производные, приравниваем их к 0, Получается система, из которой находим все стационарные точки функции f. Выписываем матрицу Гессе в произвольной точке.

Берем одну из стационарных точек и находим матрицу Гессе в этой точке. Вычисляем последовательные главные миноры этой матрицы. Если все они $\neq 0$, то работает теория, которую мы уже изучали, и нет ничего нового в решении. А дополнительные теоретические сведения нужны, если среди последовательных главных миноров есть миноры, равные нулю. Опишем такой вариант.

Рассмотрим 2 случая.

1. Все последовательные главные миноры ≥ 0. Для определенных выводов этого недостаточно, поэтому надо рассматривать не только последовательные главные миноры, но и все главные миноры. Покажем, что это такое. Будем использовать такие обозначения:

$A_{ij...m}$ − миноры из элементов строк и столбцов матрицы Гессе с номерами $i, j, ...m$.

Например, если

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix},$$

то

$$A_1 = 1,\ A_2 = 5,\ A_3 = 9,\ \ A_{12} = \begin{vmatrix} 1 & 2 \\ 4 & 5 \end{vmatrix}, A_{13} = \begin{vmatrix} 1 & 3 \\ 7 & 9 \end{vmatrix},\ \ A_{123} = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}.$$

Известно продолжение критерия Сильвестра: для неотрицательной определенности квадратичного функционала $\varphi : R^n \to R$, $\varphi(h) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}h_i h_j$ (где a_{ij} − элементы симметричной матрицы A), необходимо и достаточно, чтобы все главные миноры матрицы A были ≥ 0 (Галеев Э.М., Тихомиров В.М. Оптимизация [3]).

Для матрицы Гессе функции это означает неотрицательную определенность 2-го дифференциала в стационарной точке, которую обозначим символом x^*, т.е.

$$d^2 f(x^*, h) \geq 0 \qquad (1)$$

для любых $h \in R^n$.

Если бы мы установили, что функция f выпукла в окрестности точки x^*, то сразу сделали бы вывод о том, что стационарная точка – точка минимума (Васильев Ф.П. Методы решения экстремальных задач [4]).

Но для выпуклости функции f нужно, чтобы неравенство типа (1) выполнялось не только в точке x^*, но и при всех x из некоторой окрестности точки x^* (Васильев Ф.П. Методы решения экстремальных задач [4]).

Итак, надо проверить, не выполняется ли неравенство при всех x из некоторой окрестности точки x^* и при всех $h \in \mathbf{R}^n$. А это можно увидеть из всех главных миноров матрицы Гессе при всех x из некоторой окрестности точки x^*.

Если все такие миноры неотрицательны, то дифференциал $d^2 f$ неотрицательно определен в окрестности стационарной точки, отсюда следует выпуклость функции в этой окрестности. Тогда эта стационарная точка – точка минимума функции f (Васильев Ф.П. Методы решения экстремальных задач [4]).

2. Среди главных последовательных миноров матрицы Гессе в стационарной точке есть ≥ 0 и ≤ 0. Если все главные миноры нечетного порядка ≤ 0, а миноры четного порядка ≥ 0 в некоторой окрестности стационарной точки, то эта стационарная точка – точка локального максимума функции f. Это легко получается, если использовать вспомогательную функцию $f_1 = -f$ и применить теорию п.1.

Описанное исследование надо провести в каждой стационарной точке функции f.

Пример. Дана функция

$$f(x, y, z) = x^2 + y^2 + 3z^2 + xy + 3yz + 3zx + x + y + 2z.$$

Найдем точки экстремума этой функции.

Решение

Найдем все частные производные и приравняем их к 0. Получим систему для нахождения стационарных точек:

$$\begin{cases} 2x+y+3z=-1 \\ x+2y+3z=-1 \\ 3x+3y+6z=-2 \end{cases}.$$

Третье уравнение есть сумма первого и второго, поэтому третье уравнение отбрасываем. Остается система из двух уравнений с тремя неизвестными. Она имеет бесконечное множество решений, которое можно записать в таком виде:

$$\begin{cases} x=t \\ y=t \\ z=-t-\dfrac{1}{3} \end{cases}, \quad t\in\mathbf{R} \quad \text{или} \quad (x,y,z)=(t,\ t,-t-\frac{1}{3}),\ \ t\in\mathbf{R}.$$

Это множество стационарных точек. Проверим, есть ли среди них точки локального экстремума.

Матрица Гессе имеет вид

$$A=\begin{pmatrix} 2 & 1 & 3 \\ 1 & 2 & 3 \\ 3 & 3 & 6 \end{pmatrix}.$$

Она не зависит от координат точек, поэтому матрица будет такой во всех стационарных точках.

Так как минор $A_{123}=\begin{vmatrix} 2 & 1 & 3 \\ 1 & 2 & 3 \\ 3 & 3 & 6 \end{vmatrix}=0$ (третья строка есть сумма первой и второй), то надо рассматривать все главные миноры. Расчеты показывают, что все они ≥ 0. Тогда f − выпуклая на всем $\mathbf{R}^3$ функция. Согласно теории все стационарные точки – точки минимума.

Ответ. Множество $(x,y,z)=(t,\ t,-t-\frac{1}{3})$, $t\in\mathbf{R}$ есть множество точек минимума функции f.

Рассмотрим порядок действий при решении задач на экстремум функционала. В таких задачах требуется найти все точки локального экстремума на заданном множестве и указать их тип – точки минимума или точки максимума. Если нужно, то еще и вычислить значения функционала в точках экстремума.

Опишем кратко метод решения задачи на экстремум в случае, когда функционал дифференцируем по Фреше до второго порядка включительно. Пусть f − вещественный функционал, определенный на открытом множестве G нормированного пространства X.

Метод решения задачи на экстремум

1. Находим стационарные точки функционала f, решая уравнение $f'(x)=\theta$.

2. В каждой стационарной точке проверяем выполнение достаточных условий экстремума. Так мы находим все токи минимума и максимума. Если нужно, вычисляем значения функционала в найденных точках.

При решении задач на экстремум наиболее сложным этапом, как правило, является проверка выполнения достаточных условий экстремума. Если функционал – вещественная функция n переменных, то метод проверки выполнеиия достаточных условий экстремума хорошо известен и описан в учебниках по математическому анализу.

Напомним, что положительная определенность второго дифференциала устанавливается с помощью критерия Сильвестра. Если же функционал определен на некотором множестве из бесконечномерного нормированного пространства, то проверка выполнения достаточных условий экстремума, в основном, значительно сложнее. Такую ситуацию еще обсудим в разделе «Основная задача вариационного исчисления».

2.12. Достаточные условия отсутствия экстремума в заданной точке

Пусть X – нормированное пространство, $x_0 \in X$ − точка в пространстве X, G – открытое множество в X, $x_0 \in G$.

Рассмотрим вещественный функционал

$$f : G \to \mathbf{R},$$
дифференцируемый по Фреше в точке x_0.

Утверждение. Если точка x_0 не является стационарной точкой функционала f (т.е. если $f'(x_0) \neq \theta$), то в этой точке нет экстремума.

Это следует из теоремы о необходимых условиях экстремума (раздел 2.10, теорема 1).

А теперь изложим достаточные условия отсутствия экстремума в стационарной точке.

Теорема. Пусть X – нормированное пространство, $G \subset X$ – открытое множество, $x_0 \in G$, x_0 – стационарная точка функции $f : G \to \mathbf{R}$, эта функция дифференцируема по Фреше дважды в некоторой окрестности точки x_0, при этом функция $f''(\cdot)$ равномерно непрерывна в указанной окрестности.

Если $f''(x_0)$ − знакопеременный квадратичный функционал, то в точке x_0 нет экстремума.

Доказательство. Предположим противное, т.е. что x_0 − точка локального эстремума. Тогда по теореме 2 раздела 2.10 функционал $f''(x_0)$ является либо неотрицательным, либо неположительным, т.е. знакопеременным быть не может, а это противоречит условию теоремы. Поэтому в точке x_0 нет экстремума. ***Теорема доказана.***

Замечание. Если $X = \mathbf{R}^n$, то условие существования $f''(\cdot)$ в некоторой окрестности точки x_0 можно заменить существованием $f''(\cdot)$ только в одной точке x_0.

Если f является функцией нескольких переменных, то для знакопеременности $f''(\cdot)$ необходимо и достаточно, чтобы среди последовательных главных миноров матрицы Гессе нечетного порядка в точке x_0 были миноры разных знаков, или среди таких миноров четного порядка были миноры, меньшие нуля.

Но если это не так и среди последовательных главных миноров матрицы Гессе есть хотя бы один, равный нулю, то надо рассматривать главные миноры в окрестностях стационарной точки, причем все главные миноры, а не только последовательные.

Исследование будет закончено, если выполнены достаточные условия локального минимума или локального максимума. Если это не так, то легко сделать ошибку, если решить, что здесь нет экстремума.

Рассмотрим подробнее один из вариантов такой ситуации. Пусть все последовательные главные миноры матрицы Гессе в стационарной точке неотрицательны. Этого, как уже известно, недостаточно для локального минимума в стационарной точке x_0 (локального максимума здесь точно нет).

Тогда рассматриваем все главные миноры матрицы Гессе в окрестностях стационарной точки x_0.

Если найдется такая окрестность, в которой все главные миноры неотрицательны, то x_0 − точка локального минимума. Если же в любой окрестности точки x_0 найдется точка, отличная от x_0, в которой хотя бы один из главных миноров < 0, то есть опасность сделать вывод о том, что здесь нет локального минимума. На самом деле это не так, что легко подтверждается простым примером.

Пример. Пусть $f : R^2 \to R$, $f(x) = x_1^2 x_2^2$.

Очевидно, что $\theta = (0;0)$ − точка локального минимума. Посмотрим, какие выводы можно сделать о поведении функции в окрестности точки θ на основе теории, изложенной выше.

Матрица Гессе в произвольной точке x имеет вид

$$A(x) = \begin{pmatrix} 2x_2^2 & 4x_1x_2 \\ 4x_1x_2 & 2x_1^2 \end{pmatrix}.$$

Тогда $A(\theta) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$. Все главные миноры этой матрицы $= 0$. Этого недостаточно для каких-либо определенных выводов.

Рассмотрим теперь главные миноры в окрестностях точки θ. $A_1(x) = 2x_2^2$ $A_2(x) = 2x_1^2$, $A_{12}(x) = -12x_1^2 x_2^2$.

В любой окрестности точки θ есть точки, в которых $A_{12}(x) < 0$. Может показаться, что в точке θ минимума нет. Но мы знаем, что θ – точка минимума.

Аналогичный пример для точки максимума получится, если вместо функции f из только что рассмотренного примера взять функцию $(-f)$.

Пример показывает, что наличие минимума или максимума в стационарной точке могут сочетаться со значениями некоторых главных миноров матрицы Гессе любых знаков в точках, отличных от стационарной точки.

Заметим, что проведенное исследование традиционным методом не позволяет сделать вывод о том, что θ – точка минимума. Рассмотренный пример показывает, что изложенная теория исследования функций на экстремум дает результат не всегда.

Задачи для самостоятельного решения

Найти точки экстремума заданной функции $f: \mathrm{R}^2 \to \mathrm{R}$ и указать их тип (точка минимума, точка максимума).

1. $f(x, y) = x^3 - 2xy - \frac{1}{4}y^2 + \frac{7}{3}x$.
2. $f(x, y) = x^3 + xy + y^2 - \frac{1}{16}x$.
3. $f(x, y) = x^2 + 4xy + y^2 - 8x - 10y$.
4. $f(x, y) = x^4 + x^2 + y^2 + 2xy - 2x - 2y + 1$.

2.13. Основная задача вариационного исчисления

Вариационное исчисление – раздел математики, в котором изучаются задачи на экстремум функционалов.

Рассмотрим функционал

$$f(x) = \int_a^b \varphi(t, x(t), x'(t))dt, \tag{1}$$

где $\varphi:[a;b]\times\mathbf{R}\times\mathbf{R}\to\mathbf{R}$ – заданная функция.

Тогда $f:X\to\mathbf{R}$, где X – функциональное пространство.

Или $x\in X\xrightarrow{f} f(x)\in\mathbf{R}$.

Будем считать, что $X=C^1[a;b]$.

В пространстве $C^1[a;b]$ норма элемента x определяется формулой

$$\|x\|=\max_{t\in[a;b]}|x(t)|+\max_{t\in[a;b]}|x'(t)|.$$

Если дополнительно потребовать, чтобы функция φ была непрерывной, то получим $f:C^1[a;b]\to\mathbf{R}$.

Потребуем, чтобы функция φ имела непрерывные частные производные до второго порядка включительно.

Рассмотрим вопрос о необходимых условиях экстремума функционала (1).

Известно необходимое условие экстремума: если $x\in C^1[a;b]$ – точка локального экстремума функционала f, то $f'(x)$ – нулевой функционал.

Используем определение функционала $f'(x)$.

Возьмем произвольный элемент $h\in C^1[a;b]$.

$$f(x+h)-f(x)=\int_a^b\varphi(t,x(t)+h(t),x'(t)+h'(t))dt-\int_a^b\varphi(t,x(t),x'(t))dt$$

$$f(x+h)-f(x)=\int_a^b\Big(\frac{\partial\varphi}{\partial x}(t,x(t),x'(t))h(t)+\frac{\partial\varphi}{\partial x'}(t,x(t),x'(t))h'(t)+$$
$$+\omega_1(t,x(t),x'(t),h(t),h'(t))\Big)dt=$$

$$=\int_a^b\left(\frac{\partial\varphi}{\partial x}h(t)+\frac{\partial\varphi}{\partial x'}h'(t)\right)dt+\int_a^b\omega_1(t,x(t),x'(t),h(t),h'(t))dt\quad.$$

Можно доказать, что

$$\frac{|\omega(x,h)|}{\|h\|} \to 0 \text{ при } \|h\| \to 0,$$

где

$$\omega(x,h) = \int_a^b \omega_1(t, x(t), x'(t), h(t), h'(t))dt. \qquad (2)$$

Из формулы (2) следует, что линейная часть – это дифференциал Фреше:

$$df(x,h) = f'(x)h = \int_a^b \left(\frac{\partial\varphi}{\partial x} h(t) + \frac{\partial\varphi}{\partial x'} h'(t) \right) dt.$$

По необходимому условию экстремума

$$f'(x) \text{ – нулевой функционал } \Leftrightarrow$$

$$\Leftrightarrow \quad f'(x)h = 0 \quad \forall h \in C^1[a;b] \quad \Leftrightarrow$$

$$\Leftrightarrow \quad \int_a^b \left(\frac{\partial\varphi}{\partial x} h(t) + \frac{\partial\varphi}{\partial x'} h'(t) \right) dt = 0 \quad \forall h \in C^1[a;b]. \qquad (3)$$

Используем формулу интегрирования по частям для второго слагаемого в формуле (3).

$$\int_a^b \underbrace{\frac{\partial\varphi}{\partial x'}}_{u} \underbrace{h'(t)dt}_{dv} = \frac{\partial\varphi}{\partial x'} h(t)\Big|_a^b - \int_a^b h(t) \frac{d}{dt}\left(\frac{\partial\varphi}{\partial x'} \right) dt.$$

Подставим это выражение в формулу (3).

$$\frac{\partial\varphi}{\partial x'} h(t)\Big|_a^b + \int_a^b \left(\frac{\partial\varphi}{\partial x} - \frac{d}{dt}\left(\frac{\partial\varphi}{\partial x'} \right) \right) h(t)dt = 0 \quad \forall h \in C^1[a;b]. \qquad (4)$$

Возьмем такое $\forall h \in C^1[a;b]$, для которого $h(a) = h(b) = 0$. Тогда

$$\int_a^b \left(\frac{\partial \varphi}{\partial x} - \frac{d}{dt}\left(\frac{\partial \varphi}{\partial x'} \right) \right) h(t)dt = 0 . \tag{5}$$

Лемма 1 (основная лемма вариационного исчисления).

Пусть функционал f непрерывен на отрезке $[a;b]$ и для любой непрерывной функции h выполнено условие $\int_a^b f(x)h(x)dx = 0$. Тогда $f(x) \equiv 0$ на отрезке $[a;b]$.

Лемма 2 (усиленный вариант основной леммы вариационного исчисления).

Пусть функционал f непрерывен на отрезке $[a;b]$ и для любой функции $h \in C^1[a;b]$, для которой $h(a) = h(b) = 0$, выполнено условие $\int_a^b f(x)h(x)dx = 0$. Тогда $f(x) \equiv 0$ на отрезке $[a;b]$.

Из формулы (5) по лемме 2 следует, что

$$\frac{\partial \varphi}{\partial x} - \frac{d\varphi}{dt}\frac{\partial \varphi}{\partial x'} = 0 . \tag{6}$$

Подставим формулу (6) в формулу (4):

$$\frac{\partial \varphi}{\partial x'} h(t)\Big|_a^b = 0 . \tag{7}$$

Возьмем $h \in C^1[a;b]$, $h(a) \neq 0$, $h(b) = 0$. По формуле (7) получим

$$-\frac{\partial \varphi}{\partial x'} \underbrace{h(a)}_{\neq 0} = 0 .$$

Тогда

$$\frac{\partial \varphi}{\partial x'}\Big|_{t=a} = 0 . \tag{8}$$

Возьмем $h \in C^1[a;b]$, $h(a) = 0$, $h(b) \neq 0$. Аналогично по формуле (7) получим

$$\frac{\partial \varphi}{\partial x'}\Big|_{t=b} = 0 . \tag{9}$$

Если $x \in C^1[a;b]$ – точка локального экстремума функционала f, определенного формулой (1), то x удовлетворяет условиям (6), (8), (9), которые являются необходимыми условиями экстремума функционала (1).

Если x удовлетворяет условию (6), то x является решением уравнения (6). Тогда можно рассматривать условие (6) как уравнение относительно x.

Определение. Уравнение (6) называется ***уравнением Эйлера***; при этом элемент x называется ***решением уравнения Эйлера***, удовлетворяющим граничным условия (8) и (9).

Теорема 1. Пусть функция $\varphi : [a;b] \times \mathbf{R} \times \mathbf{R} \to \mathbf{R}$ имеет непрерывные частные производные до второго порядка включительно, $x \in C^1[a;b]$ – точка локального экстремума функционала f, определенного формулой (1). Тогда x является решением уравнения Эйлера (6), удовлетворяющим граничным условиям (8) и (9).

Задача с закрепленными границами

В предыдущей задаче экстремум рассматривался на всем пространстве $C^1[a;b]$ без дополнительных условий.

Рассмотрим другую задачу с дополнительными условиями: найти точку экстремума $x(1)$ при условии, что

$$x(a) = A, \ x(b) = B \tag{10}$$

(концы закреплены).

Теорема 2. Пусть функция $\varphi:[a;b]\times\mathbf{R}\times\mathbf{R}\to\mathbf{R}$ имеет непрерывные частные производные до второго порядка включительно, $x\in C^1[a;b]$ – точка локального экстремума функционала f, определенного формулой (1), при дополнительных условиях (10). Тогда x является решением уравнения Эйлера (6), удовлетворяющим граничным условиям (10).

Частные случаи уравнения Эйлера

Уравнение Эйлера является обыкновенным дифференциальным уравнением второго порядка относительно x.

Существуют случаи, когда порядок уравнения Эйлера можно понизить.

1. $\varphi=\varphi(t,x')$, т.е. φ явно не зависит от x.

Тогда $\dfrac{\partial\varphi}{\partial x}=0$; $-\dfrac{d}{dt}\dfrac{\partial\varphi}{\partial x'}=0$; $\dfrac{\partial\varphi}{\partial x'}=c$ – уравнение первого порядка ($c=const$).

2. $\varphi=\varphi(x,x')$, т.е. φ явно не зависит от t.

Утверждение. Пусть x – решение уравнения Эйлера в случае 2. Тогда x является решением уравнения первого порядка

$$\varphi-x'\frac{\partial\varphi}{\partial x'}=c. \tag{12}$$

Пусть x – решение уравнения (12).

Подставим x в (12) и продифференцируем полученное тождество по t.

$$\frac{\partial\varphi}{\partial x}x'+\frac{\partial\varphi}{\partial x'}x''-x''\frac{\partial\varphi}{\partial x'}-x'\frac{d}{dt}\frac{\partial\varphi}{\partial x'}=0.$$

Заметим, что второе и третье слагаемые взаимно уничтожаются. Вынесем за скобку x'; получим

$$x'\left(\frac{\partial\varphi}{\partial x}-\frac{d}{dt}\frac{\partial\varphi}{\partial x'}\right)=0\,. \qquad (13)$$

Возможно и обратное преобразование к уравнению (12). Тогда уравнения (12) и (13) эквивалентны.

Очевидно, что если x – решение уравнения (6), то x – решение уравнения (13), а следовательно, и уравнения (12).

На практике вместо уравнения (6) решают уравнение (12).

2.14. Пример решения вариационной задачи

Задача. Требуется изготовить трубу заданной длины; ее форма – некоторая поверхность вращения, поперечные сечения – окружности, радиусы сечений на концах равны заданным числам и равны между собой. Расход материала пропорционален площади поверхности трубы. Найти такую форму трубы, чтобы расход материала был наименьшим.

Математическая постановка задачи

Дана поверхность вращения, определяющаяся некоторой кривой, вращением которой она образована; кривая может быть задана в виде функции $x = x(t)$.

Пусть A – радиус сечений на концах трубы; l – длина трубы. Введем число $a=\dfrac{l}{2}$. Расположим график функции x так, чтобы $t\in[-a;a]$.

Обозначим S – площадь поверхности.

Согласно известной формуле

$$S=2\pi\int_{-a}^{a}x(t)\sqrt{1+(x'(t))^2}\,dt\,.$$

Рассмотрим функционал

$$f(x)=2\pi\int_{-a}^{a}x(t)\sqrt{1+(x'(t))^2}\,dt\,. \qquad (14)$$

Точка минимума функционала f является точкой минимума площади поверхности S.

Тогда ***задача*** принимает следующий вид: найти в $C^1[a;b]$ точку минимума функционала (14) при дополнительных условиях

$$x(-a) = x(a) = A \qquad (15)$$

(вариационная задача с закрепленными концами).

Воспользуемся теоремой 2.

Если x – точка локального экстремума, то x – решение уравнения Эйлера при условиях (15).

Здесь $\varphi = x\sqrt{1+x'^2}$ (для функционала (14)); φ явно не зависит от t (случай 2).

Требуется решить уравнение (12).

$$\varphi - x'\frac{\partial\varphi}{\partial x'} = c\,.$$

Найдем частную производную $\frac{\partial\varphi}{\partial x'}$ от функции $\varphi = x\sqrt{1+x'^2}$:

$$\frac{\partial\varphi}{\partial x'} = x\frac{2x'}{2\sqrt{1+x'^2}}\,.$$

Подставим $\frac{\partial\varphi}{\partial x'}$ в уравнение (12):

$$x\sqrt{1+x'^2} - \frac{xx'^2}{\sqrt{1+x'^2}} = c\,.$$

Умножим обе части уравнения на $\sqrt{1+x'^2}$:

$$x(1+x'^2) - xx'^2 = c\sqrt{1+x'^2}\,.$$

Раскрыв скобки, получим $x = c\sqrt{1+x'^2}$. Возведем обе части в квадрат и разрешим уравнение относительно x'.

$$x^2 = c^2\sqrt{1+x'^2} \ ,$$

$$x'^2 = \frac{x^2}{c^2} - 1 \ ,$$

$$x'^2 = \frac{x^2 - c^2}{c^2} \ \text{ при } \ c \neq 0 \ ,$$

$$x' = \pm\frac{\sqrt{x^2 - c^2}}{|c|} \ .$$

Раскрывая модуль, при любых ненулевых значениях *c* получим

$$x' = \frac{\sqrt{x^2 - c^2}}{c} \ .$$

Иначе $\dfrac{dx}{dt} = \dfrac{\sqrt{x^2 - c^2}}{c}$.

Решим полученное дифференциальное уравнение в разделяющихся переменных:

$$\frac{dx}{\sqrt{x^2 - c^2}} = \frac{1}{c}dt \ ,$$

$$\int\frac{dx}{\sqrt{x^2 - c^2}} = \frac{1}{c}t + c_1 \ . \qquad (16)$$

Сделаем замену переменной $x = c\operatorname{ch} z$. Тогда

$$x^2 - c^2 = c^2(\operatorname{ch}^2 z - 1) = c^2 \operatorname{sh}^2 z \ ,$$

$$dx = c\operatorname{sh} z dz \ ,$$

$$\int \frac{dx}{\sqrt{x^2 - c^2}} = \int \frac{c\,\mathrm{sh}\, z dz}{|c| \cdot |\mathrm{sh}\, z|} = \pm \int dz\,.$$

Подставим этот интеграл в уравнение (16).

$\pm \int dz = \frac{1}{c} t + c_1$, где c и c_1 – произвольные постоянные.

$$\int dz = \frac{1}{c} t + c_1\ ,$$

$$z = \frac{1}{c} t + c_1\,.$$

Тогда

$$x = c\mathrm{ch}\, z = c\,\mathrm{ch}(\frac{1}{c} t + c_1) \qquad (17)$$

(общее решение уравнения Эйлера).

Используем условия (15):

$$x(-a) = x(a) = A\,,$$

$$c\,\mathrm{ch}(-\frac{a}{c} + c_1) = c\,\mathrm{ch}(\frac{a}{c} + c_1)\,.$$

Функция ch является четной, поэтому $c_1 = 0$.
Окончательно получаем

$$x(t) = c\,\mathrm{ch}(\frac{t}{c})\,. \qquad (18)$$

Используем условие $x(a) = A$ (из (15)):

$$c\,\mathrm{ch}(\frac{a}{c}) = A\,. \qquad (19)$$

Теперь требуется найти c .

Если уравнение (19) имеет решения относительно c, то из равенства (18) получаем стационарные точки (экстремали) функционала. Только в этих точках может быть экстремум.

Теперь для определения того, являются ли стационарные точки точками экстремума, нужны дополнительные исследования. Достаточные условия экстремума здесь неудобны, так как трудно установить сильную положительность или сильную отрицательность второй производной функционала f.

Для вариационных задач разработаны специальные методы исследования функционала на экстремум. Некоторые их них изложены, например, в книге [5] (гл. 6, §9). Там же дано полное исследование задачи, которую мы сейчас рассматриваем.

Если точка минимума существует, то она имеет вид (18).

Предполагаемая форма трубы имеет вид (см. рис.1).

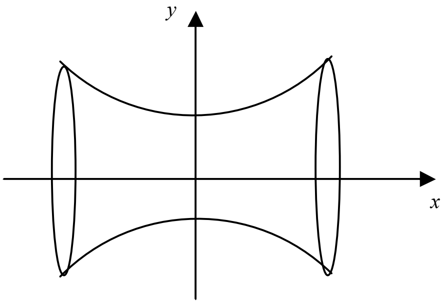

Рис. 1

Преобразуем уравнение (19).

Сделаем замену переменной $\tau = \dfrac{a}{c}$. Тогда $c = \dfrac{a}{\tau}$. Подставим в уравнение (19):

$$\frac{a}{\tau}\operatorname{ch}\tau = A,$$

$$\operatorname{ch}\tau = \frac{A}{a}\tau . \tag{20}$$

Обозначим

$$k = \frac{A}{a} . \tag{21}$$

Тогда

$$\operatorname{ch}\tau = k\tau .$$

Рассмотрим графики функций в левой и правой частях: $y = \operatorname{ch}\tau$ и $y = k\tau$.

Существует число k_0, при котором прямая $y = k_0\tau$ касается графика функции $y = \operatorname{ch}\tau$ (см. рис. 2).

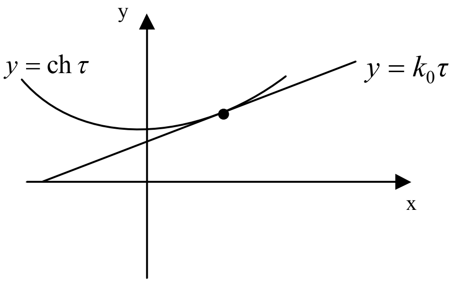

Рис. 2

Тогда из графика видно, что уравнение (21) не имеет решений при $0 < k < k_0$ (при сдвиге прямой $y = k_0\tau$ вниз она не пересекает график функции $y = \operatorname{ch}\tau$), имеет единственное решение при $k = k_0$ (касание) и два решения при $k > k_0$ (при сдвиге прямой $y = k_0\tau$ вверх она пересекает график функции $y = \operatorname{ch}\tau$ в двух точках).

Таким образом, возможны следующие случаи:

- нет локального экстремума;
- одна стационарная точка (экстремаль);
- две стационарные точки (две экстремали).

Используя приближенные методы, можно вычислить, что $k_0 \approx 1{,}5$.

В уже упомянутой книге [5] (гл. 6, §9) показано, что наименьшее значение функции f получается в точке x, определенной равенством (18), при значении c, равном наибольшему корню уравнения (19).

Таким образом, задача имеет решение при соотношении $\frac{A}{a} \geq k_0 \approx 1{,}5$, где A – радиус сечения, $d = 2A$ – диаметр сечения, $l = 2a$ – длина трубы.

$$\frac{d}{l} = \frac{A}{a} = k \geq k_0 \approx 1{,}5 \ .$$

Таким образом, задача имеет решение для относительно коротких труб с отношением длины трубы к диаметру сечения $\frac{d}{l} = k \geq k_0 \approx 1{,}5$.

В некоторых случаях задачу можно решить без использования достаточных условий экстремума при помощи сравнения значения функционала в стационарной точке и в остальных точках множества, на котором ищем точку экстремума. Рассмотрим соответствующий пример.

Пример. Дан функционал $f(x) = \int_0^1 (tx' + x'^2)dt$ и граничные условия $x(0) = 0$, $x(1) = \frac{3}{4}$.

Требуется сначала найти стационарные точки (экстремали) функционала f в пространстве $C^1[a;b]$, затем установить, есть ли среди экстремалей точки экстремума.

Решение. Запишем уравнения Эйлера:

$$\frac{\partial \varphi}{\partial x} - \frac{d\varphi}{dt}\frac{\partial \varphi}{\partial x'} = 0 \ .$$

Можно заметить, что $\varphi = tx' + x'^2$ явно не зависит от x (случай 1). Тогда используем уравнение

$$\frac{\partial \varphi}{\partial x'} = c \ .$$

Найдем частную производную $\frac{\partial \varphi}{\partial x'} = t + 2x'$ и подставим ее в уравнение выше:

$$t + 2x' = c . \tag{22}$$

Уравнение (22) представляет собой дифференциальное уравнение в разделяющихся переменных; решим его:

$$2\frac{dx}{dt} = c - t ,$$

$$2dx = (c - t)dt ,$$

$$2x = ct - \frac{t^2}{2} + c_0 .$$

Получим решение

$$x = -\frac{t^2}{4} + c_1 t + c_2 , \quad \text{где } c_1 = \frac{c}{2}, \quad c_2 = \frac{c_0}{2} .$$

Подставим в решение граничные условия:

$$x(0) = 0 \Rightarrow c_2 = 0 ,$$

$$x(1) = \frac{3}{4} \Rightarrow c_1 = 1 .$$

Получается стационарная точка $x = -\dfrac{t^2}{4} + t$.

Установим теперь, является ли эта стационарная точка точкой экстремума. Для удобства дальнейших рассуждений обозначим эту стационарную точку так:

$$x_0(t) = -\frac{t^2}{4} + t . \tag{23}$$

Пусть теперь x − произвольная точка (функция) пространства $C^1[0;1]$, удовлетворяющая граничным условиям

$$x(0) = 0, \quad x(1) = \frac{3}{4} . \tag{24}$$

Рассмотрим разность $\Delta f = f(x) - f(x_0)$. Введем обозначение $h = x - x_0$. Так как все функции x, а также x_0, удовлетворяют условиям (24), то

$$h(0)=x(0)-x_0(0)=0-0=0;\ h(1)=x(1)-x_0(1)=\frac{3}{4}-\frac{3}{4}=0\,.$$

Итак,

$$h(0)=h(1)=0\,. \tag{25}$$

Так как $h=x-x_0$, то $x=x_0+h$. Тогда

$$\Delta f=f(x)-f(x_0)=$$

$$f(x_0+h)-f(x_0)=\int_0^1(t(\cancel{x}_0'(t)-h'(t))+(x_0'(t)+h'(t))^2-tx_0'(t)-(x_0'(t))^2)dt=$$

$$=\int_0^1(th'(t)+2x_0'(t)h'(t)+(h'(t)^2)dt=\int_0^1(t+2x_0'(t))h'(t)+(h'(t)^2)dt\,.$$

Так как x_0 удовлетворяет условию (22), то

$$\Delta f=\int_0^1(ch'(t)+(h'(t))^2)dt==c(h(1)-h(0))+\int_0^1(h'(t))^2\,dt\,.$$

Так как выполнены условия (25), то $\Delta f=\int_0^1(h'(t))^2\,dt$.

Очевидно, $\Delta f\geq 0$.

Тогда $\Delta f=f(x)-f(x_0)\geq 0$. Это неравенство выполнено для всех $x\in C^1[0;1]$, удовлетворяющих условиям (24). Отсюда следует, что x_0 – точка минимума функционала f.

Задачи для самостоятельного решения

Найти точки экстремума интегрального функционала f в просранстве $C^1[1;2]$ при заданных граничных условиях.

1. $f(x)=\int\limits_1^2 x'(t)(1+t^{-3}x'(t))dt,\;\; x(1)=4,\; x(2)=19$.

2. $f(x)=\int\limits_1^2 x'(t)(1-t^{3}x'(t))dt,\;\; x(1)=1,\; x(2)=-\frac{1}{2}$.

3. $f(x)=\int\limits_1^2 x'(t)(1+t^{4}x'(t))dt,\;\; x(1)=7,\; x(2)=0$.

4. $f(x)=\int\limits_1^2 x'(t)(1-t^{-2}x'(t))dt,\;\; x(1)=1,\; x(2)=15$.

2.15. Принцип сжимающих отображений и его обобщения

Пусть X – некоторое множество, A – оператор, $A: X \to X$.

Определение. Точка $x^* \in X$ называется ***неподвижной точкой*** оператора A, если $x^* = Ax^*$.

Рассмотрим операторное уравнение

$$x = Ax. \tag{1}$$

Утверждение. Точка x^* является неподвижной точкой оператора A тогда и только тогда, когда x^* является решением уравнения (1).

Пусть $< X, d >$ – метрическое пространство, $A: X \to X$. Одним из способов приближенного решения уравнения (1) является метод последовательных приближений, который заключается в том, что берется некоторое начальное приближение $x_0 \in X$, затем строятся последовательные приближения по правилу $x_{n+1} = Ax_n,\;\; n = 0,1,2,...$ Если последовательность $\{x_n\}$ сходится, то при определенных условиях предел этой последовательности есть решение уравнения (1). А в качестве приближенного решения берут одно из последовательных приближений.

Достоинство метода – его простота. Но он применим далеко не всегда, так как последовательность приближений может расходиться. Для сходимости метода требуется выполнение дополнительных условий. Далее сформулируем вариант условий, при которых существует единственное решение уравнения (1) и его можно найти методом последовательных приближений. Для этого понадобится понятие оператора сжатия.

Определение. Оператор $A: X \to X$ называется ***оператором сжатия*** (***оператором*** α ***-сжатия***), если существует такая константа $\alpha \in [0;1)$, что для любых $x_1, x_2 \in X$ выполняется условие

$$d(Ax_1, Ax_2) \le \alpha d(x_1, x_2). \tag{2}$$

Теорема 1 (основной вариант принципа сжимающих отображений).

Пусть метрическое пространство $< X, d >$ полно, оператор $A: X \to X$ является оператором α -сжатия, $0 \le \alpha < 1$.

Тогда

1) неподвижная точка оператора A (решение уравнения (1)) в X существует;

2) неподвижная точка оператора A единственна в X;

3) решение уравнения (1) (неподвижную точку) можно найти методом последовательных приближений, начиная с любого начального приближения $x_0 \in X$;

4) справедлива оценка погрешности

$$d(x_n, x^*) \le \frac{\alpha^n}{1-\alpha} d(Ax_0, x_0) \ \ (n = 1,2,\ldots). \tag{3}$$

Здесь x^* – неподвижная точка, x_0 – начальное приближение, x_n – n-е приближение.

Доказательство есть в учебниках по функциональному анализу, напимер, в книге А.Н.Колмогорова и С.В.Фомина [1].

Возникает вопрос, можно ли доказать существование решения уравнения (1), если оператор A является оператором сжатия не на всем множестве X.

Известна теорема, которая позволяет это сделать – обобщение теоремы 1.

Теорема 2 (аналог теоремы 1 для оператора сжатия на части пространства).

Пусть выполнены условия:

а) $< X, d >$ – полное метрическое пространство, оператор $A : X \to X$;

б) оператор A является оператором α-сжатия на множестве $X_0 \subset X$, $\alpha \in [0;1)$;

в) множество X_0 замкнуто, оператор $A : X_0 \to X_0$.

Тогда

1) существует неподвижная точка оператора A в пространстве X_0;

2) неподвижная точка единственна в X_0;

3) неподвижную точку можно найти методом последовательных приближений, начиная с любого $x_0 \in X_0$;

4) справедлива оценка погрешности (3).

Доказательство. Будем рассматривать множество X_0 как самостоятельное метрическое пространство с той же метрикой *d.* Тогда оператор A преобразует множество X_0 в себя и является на множестве X_0 оператором сжатия.

Из полноты пространства X и замкнутости X_0 следует полнота пространства X_0. Итак, выполнены все условия теоремы 1 для пространства X_0. Поэтому выполняются все утверждения теоремы 2. ***Теорема доказана.***

Заметим, что единственность решения гарантируется лишь на множестве X_0.

Возникает вопрос, как проверить условие $A : X_0 \to X_0$.

Существует метод для случая, когда X_0 – замкнутый шар.

Теорема 3. Пусть $< X, d >$ – метрическое пространство, $A : X \to X$ является оператором α-сжатия на X, $\alpha \in [0;1)$; $X_0 \subset X$; существует такое $r > 0$, что

$$d(Ax_0, x_0) \le (1-\alpha)r . \tag{4}$$

Тогда $A:\overline{B}_r \to \overline{B}_r$, где $\overline{B}_r = \{x \in X : d(x, x_0) \le r\}$ – замкнутый шар с центром в точке x_0 и радиусом r.

Теорема 4. Пусть X – нормированное пространство, M – выпуклое множество в пространстве X, оператор $A: M \to X$ имеет производную Гато на множестве M и выполняется условие

$$\left\| A_c'(x) \right\| \le \alpha < 1 \quad \forall x \in M . \tag{5}$$

Тогда оператор A является оператором α-сжатия на множестве M.

Доказательство. Утверждение теоремы следует из второй формулы конечных приращений (раздел 2.4).

Рассмотрим еще одно важное обобщение принципа сжимающих отображений. Для этого потребуется следующая теорема.

Теорема 5. Пусть X – некоторое множество; пусть выполнены условия:

1) $A, B: X \to X$;

2) оператор B имеет единственную неподвижную точку на множестве X;

3) оператор B коммутирует с оператором A ($AB = BA$).

Тогда существует неподвижная точка оператора A на множестве X.

Доказательство. Пусть x^* − единственная неподвижная точка оператора B на X, т.е. $x^* = Bx^*$, отсюда следует равенство $Ax^* = ABx^*$.

Так как A коммутирует с B, то

$$Ax^* = BAx^* \iff Ax^* = B(Ax^*).$$

Отюда следует, что Ax^* − неподвижная точка оператора B.

Так как x^* − тоже неподвижная точка оператора B, а по условию неподвижная точка оператора B единственна, то $x^* = Ax^*$, откуда следует, что x^* − неподвижная точка оператора A.

Итак, неподвижная точка оператора A существует. ***Теорема доказана.***

Комментарий. В этой теореме единственность неподвижной точки оператора A не гарантируется. Из доказательства теоремы следует, что одна из неподвижных точек оператора A совпадает с неподвижной точкой оператора B.

Теорема 6 (обобщение принципа сжимающих отображений).

Пусть X – полное метрическое пространство, $A: X \to X$, существует такое число $n \in \mathbf{N}$, что A^n является оператором сжатия. Тогда существует единственная неподвижная точка оператора A в X, которую можно найти методом последовательных приближений, начиная с любого начального приближения $x_0 \in X$.

Идея доказательства. Из принципа сжимающих отображений следует существование и единственность неподвижной точки оператора A^n. Но оператор A коммутирует с оператором A^n, так как $AA^n = A^{n+1} = A^n A$ (ассоциативность композиции).

По теореме 5 существует неподвижная точка оператора A. Остальные пункты утверждения теоремы 6 доказываются довольно легко. Полное доказательство есть в книге А.Н.Колмогорова и С.В.Фомина [1].

2.16. Приложения принципа сжимающих отображений

Известны различные приложения принципа сжимающих отображений. Рассмотрим некоторые из них.

1. Уравнение $x = \varphi(x)$**,** $\varphi: \mathbf{R} \to \mathbf{R}$ (1).

Теорема 1. Если функция φ дифференцируема на $\mathbf{R}$, для любого $x \in \mathbf{R}$ $|\varphi'(x)| \le q < 1$, то уравнение (1) имеет единственное решение в $\mathbf{R}$, которое можно найти методом последовательных приближений.

Доказательство. По теореме 4 раздела 2.14 оператор φ - оператор сжатия. Существование, единственность решения, сходимость метода последовательных приближений следуют из принципа сжимающих отображений. ***Теорема доказана.***

2. Задача Коши для обыкновенного дифференциального уравнения.

$$\begin{cases} \dfrac{dx}{dt} = f(t, x(t)) \\ x(t_0) = x_0 \end{cases}. \tag{2}$$

Здесь f – заданная функция, x_0, t_0 – заданные числа.

Это задача Коши для обыкновенного дифференциального уравнения первого порядка.

Теорема 2. Пусть выполнены условия:

1) функция f непрерывна на $[t_0; t_1] \times \mathbf{R}$, $t_1 > t_0$;

2) функция f удовлетворяет условию Липшица по второму аргументу, т.е существует такая константа $L \geq 0$, что для любых $x_1, x_2 \in \mathbf{R}$ и для любого $t \in [t_0; t_1]$ выполняется условие

$$\left| f(t, x_1) - f(t, x_2) \right| \leq L \left| x_1 - x_2 \right|;$$

3) $0 \leq t_1 - t_0 < \dfrac{1}{L}$.

Тогда существует единственное решение задачи Коши (7)-(8), определенное на отрезке $[t_0; t_1]$.

Идея доказательства. Сначала доказывается, что задача Коши (2) эквивалентна интегральному уравнению

$$x(t) = x_0 + \int_{t_0}^{t} f(s, x(s)) ds \qquad (t \in [t_0; t]),$$

которое можно рассматривать в качестве операторного уравнения $x = Ax$ в пространстве $C[t_0; t_1]$. В этом уравнении оператор A определяется равенством

$$(Ax)(t) = x_0 + \int_{t_0}^{t} f(s, x(s)) ds \quad (t \in [t_0; t_1]). \tag{3}$$

Далее легко доказывается, что оператор A есть оператор сжатия. Существование и единственность решения вытекает из принципа сжимающих отображений.

Комментарий. Это локальная теорема существования, т.к. длина отрезка $[t_0;t_1]$ ограничена сверху.

С помощью обобщения принципа сжимающих отображений (теорема 6 раздела 2.14) можно доказать нелокальную теорему существования и единственности решения задачи Коши.

Теорема 3. Пусть выполнены условия:

1) функция f непрерывна на $\mathbf{R}\times\mathbf{R}$;

2) функция f удовлетворяет условию Липшица по второму аргументу, т.е существует такая константа $L \ge 0$, что для любых $t, x_1, x_2 \in \mathbf{R}$ выполняется условие

$$\left|f(t,x_1)-f(t,x_2)\right| \le L\left|x_1-x_2\right|.$$

Тогда существует единственное решение задачи Коши (2), определенное на любом отрезке $[a,b]$.

Идея доказательства. Если рассматривать отрезок произвольной длины, на котором определяется решение задачи Коши, то оператор *A,* определенный равенством (3), может не быть оператором сжатия.

Но здесь оказывается полезной теорема 6 раздела 2.14. Несложно доказывается, что при любом таком заданном отрезке найдется такое натуральное число n, что оператор A^n будет операторм сжатия. Поэтому по теореме 6 раздела 2.14 решение задачи Коши, определенное на заданном отрезке, существует и единственно.

3. Решение уравнения вида *f(x)=0* методом последовательных приближений

Принцип сжимающих отображений может оказаться полезным не только для уравнений вида

$$x=\varphi(x),\ \varphi:\mathbf{R}\to\mathbf{R},$$

но и для уравнений вида

$$f(x)=0\,,\ f:\mathbf{R}\to\mathbf{R}\,. \qquad (4)$$

Рассмотрим такое уравнение.

Преобразуем уравнение (1) к виду

$$x=\varphi(x)\,. \qquad (5)$$

Рассмотрим уравнение (1) при условии, что функция f дифференцируема на $\mathbf{R}$, и для любого $x\in\mathbf{R}$ выполняется условие

$$0<m\le f'(x)\le M\ . \qquad (6)$$

Один из способов преобразований имеет вид

$$\varphi(x)=x-\lambda f(x)\,, \qquad (7)$$

где λ – константа, $\lambda>0$.

Выясним, какое λ нужно взять, чтобы выполнить преобразование (4).

Утверждение. Если функция f удовлетворяет условию (7) и

$$\lambda=\frac{2}{M+m} \qquad (8),$$

то φ является оператором α -сжатия, где

$$\alpha=\frac{M-m}{M+m} \qquad (9).$$

Пример

Дано уравнение $f(x)=0$, где $f(x)=x^3+x^2+2x-0{,}1$.

Требуется:

1) преобразовать уравнение к виду (5);

2) взять начальное приближение x_0 и доказать, что последовательные приближения сходятся к решению;

3) вычислить x_1;

4) предсказать количество приближений, достаточных для достижения точности $\varepsilon = 0{,}01$.

Этапы решения

1. Найти отрезок $[a;b]$, на котором уравнение имеет корень.

2. Найти $m = \min\limits_{[a;b]} f'(x)$ и $M = \max\limits_{[a;b]} f'(x)$.

3. Найти λ по формуле (8) и преобразовать уравнение (4) к виду (5).

4. Найти α, которое определяет сжатие, по формуле (9).

5. Проверить выполнение условия $\varphi:[a;b] \to [a;b]$ (по теореме 3 раздела 2.14). Если условие не выполнено, нужно уменьшить отрезок $[a;b]$.

6. Взять начальное приближение x_0 и вычислить x_1.

7. Предсказать количество приближений, достаточных для достижения точности $\varepsilon = 0{,}01$, используя условие

$$d(x_n, x^*) < \frac{\alpha^n}{1-\alpha} d(x_n, x_0)\,. \qquad (10)$$

Решение

В качестве отрезка $[a;b]$ возьмем отрезок $[0;1]$.

Найдем производную $f'(x) = 3x^2 + 2x + 2$.

Для данной производной $m = \min\limits_{[0;1]} f'(x) = 2$, $M = \max\limits_{[0;1]} f'(x) = 7$.

По формуле (8) $\lambda = \dfrac{2}{M+m} = \dfrac{2}{9}$.

По формуле (9) $\alpha = \dfrac{M-m}{M+m} = \dfrac{5}{9}$.

Тогда преобразование (7) имеет вид

$$\varphi(x) = x - \frac{2}{9}(x^3 + x^2 + 2x - 0{,}1)\,.$$

Отрезок $[0;1]$ в пространстве $\mathbf{R}$ представляет собой замкнутый шар с центром $x_0 = \dfrac{1}{2}$ и радиусом $r = \dfrac{1}{2}$. Тогда согласно теореме 3 раздела 2.14 $\varphi:[0;1] \to [0;1]$, если

$$d(\varphi(x_0), x_0) \le (1-\alpha)r\ . \qquad (11)$$

Проверим это условие.

$$\varphi(x_0) = \frac{1}{2} - \frac{2}{9}(\frac{1}{8} + \frac{1}{4} + 1 - 0{,}1) = \frac{13}{60}\ .$$

$$d(\varphi(x_0), x_0) = \left|\varphi(x_0) - x_0\right| = \left|\frac{13}{60} - \frac{1}{2}\right| = \frac{17}{60}\ .$$

$$(1-\alpha)r = (1 - \frac{5}{9})\frac{1}{2} = \frac{2}{9}\ .$$

По условию (11) должно быть $\frac{17}{60} \le \frac{2}{9}$, но это неверно. Поэтому не удалось доказать, что $\varphi : [0;1] \to [0;1]$. Следовательно, требуется уменьшить отрезок.

Возьмем отрезок $[0; \frac{1}{2}]$. На этом отрезке $m = 2$, $M = \frac{15}{4}$, $\lambda = \frac{8}{23}$, $\alpha = \frac{7}{23}$.

Тогда преобразование (7) имеет вид

$$\varphi(x) = x - \frac{8}{23}(x^3 + x^2 + 2x - 0{,}1)\ .$$

Снова проверим условие (11). Отрезок $[0; \frac{1}{2}]$ в пространстве **R** представляет собой замкнутый шар с центром $x_0 = \frac{1}{4}$ и радиусом $r = \frac{1}{4}$.

$$\varphi(x_0) = \varphi(\frac{1}{4}) = \frac{1}{4} - \frac{8}{23}(\frac{1}{64} + \frac{1}{16} + \frac{1}{2} - 0{,}1) \approx 0{,}09\ .$$

$$d(\varphi(x_0),x_0)=\left|\varphi(x_0)-x_0\right|\approx\left|0{,}25-0{,}09\right|=0{,}16\,.$$

$$(1-\alpha)r=(1-\frac{7}{23})\frac{1}{4}\approx 0{,}174\,.$$

По условию (11) должно быть $0{,}16\le 0{,}174$, это верно. Тогда действительно $\varphi:[0;\frac{1}{2}]\to[0;\frac{1}{2}]$.

Возьмем начальное приближение $x_0=\frac{1}{4}$. Тогда $x_1=\varphi(x_0)\approx 0{,}09$.

Для оценки погрешности используем формулу (11).

$$d(x_n,x^*)=\left|x_n-x^*\right|\le\frac{\left(\frac{7}{23}\right)^n}{1-\frac{7}{23}}\cdot 0{,}16<0{,}01\,.$$

Это условие выполнено, начиная с $n=3$.

Таким образом, для достижения точности $\varepsilon=0{,}01$ необходимо взять 3 приближения.

3.1. Метод Ньютона

Одним из известных методов решения уравнений вида

$$f(x)=0\,,\ f:\mathbf{R}\to\mathbf{R}$$

является так называемый метод Ньютона, который заключается в том, что по заданному начальному приближению $x_0\in\mathbf{R}$ вычисляются последовательные приближения

$$x_{n+1}=x_n-\frac{f(x_n)}{f'(x_n)}\,,\ n=0,1,\ldots$$

При определенных условиях последовательность таких приближений сходится к решению уравнения. Особенностью этого метода является очень быстрая сходимость, в чем и заключается его особая привлекательность. Но в таком варианте у метода есть очевидный недостаток, поскольку он предназначен для решения уравнений в одномерном пространстве.

Для преодоления этого недостатка в середине XX в. было найдено обобщение метода на случай уравнений в нормированных пространствах любой размерности и в бесконечномерных пространствах. При этом сохранено основное достоинство метода – его быстрая сходимость.

.

Пусть X, Y – нормированные пространства, оператор $F: U \to Y$, U – некоторое множество в X.

Рассмотрим уравнение $Fx = \theta$. (1)

Метод Ньютона для уравнения (1) заключается в следующем:

1) берем какое-нибудь начальное приближение $x_0 \in U$;

2) затем находим последовательные приближения, которые определяются формулой

$$x_{n+1} = x_n - \left(F'(x_n)\right)^{-1} Fx_n,\ n = 0,1,\ldots, \qquad (2)$$

где $F'(x)$ – производная Фреше.

Теорема 1 (о сходимости метода Ньютона для уравнения (1)).

Пусть выполнены условия:

1) X, Y – нормированные пространства, $x^* \in X$, U – окрестность точки x^*, оператор $F: U \to Y$, x^* – решение уравнения (1);

2) оператор F дифференцируем по Фреше в U;

3) все операторы $F'(x)$ обратимы при $x \in U$ и операторы $\left(F'(x)\right)^{-1}$ непрерывны;

4) оператор $\left(F'(x)\right)^{-1}: U \to N(Y, X)$, где $N(Y, X)$ – пространство линейных непрерывных операторов, действующих из Y в X, непрерывен в точке x^*.

Тогда существует такая окрестность U_0 точки x^*, что для любого начального приближения $x_0 \in U_0$ метод Ньютона (2) сходится к решению x^*.

Идея доказательства. Рассматривается оператор $A: U \to X$, определенный формулой

$$Ax = x - \left(F'(x)\right)^{-1} Fx \,, \tag{3}$$

и рассматривается уравнение

$$x = Ax \,. \tag{4}$$

Это уравнение эквивалентно исходному уравнению (1). Тогда метод простой итерации для уравнения (4) есть метод Ньютона (2). Поэтому надо доказать, что при условиях теоремы метод простой итерации для уравнения (4) сходится к решению при некотором начальном приближении.

Можно доказать, что для любого числа $\alpha \in (0;1)$ найдется такая окрестность U_0 точки x^*, что

$$\left\|Ax - x^*\right\| \le \alpha \left\|x - x^*\right\| \quad \forall x \in U_0 \;.$$

Тогда для последоваельных приближений выполняется неравенство

$$\left\|Ax_n - x^*\right\| \le \alpha \left\|x_n - x^*\right\| \qquad (n = 0,1,2,\ldots) \,,$$

если начальное приближение взять из U_0. Это неравенство эквивалентно такому неравенству:

$$\left\|x_{n+1} - x^*\right\| \le \alpha \left\|x_n - x^*\right\| \qquad (n = 0,1,2,\ldots) \,. \tag{5}$$

Из неравенства (5) следует такое неравенство:

$$\left\|x_n - x^*\right\| \le \alpha^n \left\|x_0 - x^*\right\| \qquad (n = 0,1,2,\ldots) \,, \tag{6}$$

из которого следует сходимость $x_n \to x^*$ при $n \to \infty$.

Комментарий. Из доказательства теоремы следует, что можно взять любое $\alpha \in (0;1)$, при таком α можно найти начальное приближение, при котором выполняется неравенство (6). Из этого неравенства следует, что чем меньше $\alpha \in (0;1)$, тем быстрее сходимость. Поэтому удачным выбором начального приближении можно обеспечить как угодно быструю сходимость.

Определение. Говорят, что итерационный процесс для уравнения (4) ***сходится как геометрическая прогрессия со знаменателем*** α, если все последовательные приближения удовлетворяют условию (5).

При решении уравнения методом Ньютона чем ближе x_n к решению x^*, тем меньше α, которым характеризуется скорость сходимости, причем $\alpha \to 0$ при $n \to \infty$. Это следует из предыдущих рассуждений. Поэтому говорят, что скорость сходимости метода Ньютона выше, чем скорость сходимости геометрической прогрессии.

Пример. Дана система уравнеий

$$\begin{cases} x^3 + y^2 - 2 = 0, \\ x^2 y + y^2 - 2 = 0. \end{cases}$$

Эту систему можно рассматривать как операторное уравнение

$$Fz = \theta ,$$

где $F : \mathbf{R}^2 \to \mathbf{R}^2$, $Fz = \begin{pmatrix} x^3 + y^2 - 2 \\ x^2 y + y^2 - 2 \end{pmatrix}$, $z = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbf{R}^2$.

Очевидно,

$$F'(z) = \begin{pmatrix} 3x^2 & 2y \\ 2xy & x^2 + 2y \end{pmatrix}, \quad (F'(z))^{-1} = \frac{1}{\Delta(z)} \begin{pmatrix} x^2 + 2xy & -2xy \\ -2y & 3x^2 \end{pmatrix},$$

где $\Delta(z) = \det F'(z) = 3x^2\left(x^2 + 2y\right) - 4xy^2$.

Последовательные приближения вида (2) по методу Ньютона будут иметь вид

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} x_n \\ y_n \end{pmatrix} - \frac{1}{\Delta(z_n)} \begin{pmatrix} x_n^2 + 2x_n y_n & -2x_n y_n \\ -2y_n & 3x_n^2 \end{pmatrix} \begin{pmatrix} x_n^3 + y_n^2 - 2 \\ x_n^2 y_n + y_n^2 - 2 \end{pmatrix},$$

$n = 0,1,2,...$

Возьмем начальное приближение $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$. Вычислим несколько последовательных приближений. Получились такие результаты:

n	x	y
0	3	4
1	1,888	2,629
2	1,233	1,747
3	1,013	1,196
4	1,006	1,009
5	1,000	1,000

Последовательные приближения довольно быстро сходятся к точному решению $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Последовательность расстояний от последовательных приближений до точного решения, вычисленных по норме пространства $\mathbf{R}^2$, такова:

$$3{,}606;\ \ 1{,}855;\ \ 0{,}782;\ \ 0{,}196;\ \ 0{,}009;\ 0.$$

Последовательность отношений соседних расстояний:

$$0{,}617;\ \ 0{,}421;\ 0{,}251;\ \ 0{,}046;\ 0.$$

То, что эти отношения стремятся к 0, свидетельствует о том, что здесь скорость сходимости выше, чем скорость сходимости геометрической прогрессии.

3.2. Неявные функции

1. Декартово произведение нормированных пространств

Пусть X, Y – нормированные пространства, $\|\cdot\|_X$ и $\|\cdot\|_Y$ – нормы на X и Y соответственно.

Декартово произведение пространств X и Y определяется формулой

$$X \times Y = \{< x, y >: x \in X, y \in Y\}.$$

Норму $\|\cdot\|_{X\times Y}$ пространства $X \times Y$ можно определить разными способами через $\|\cdot\|_X$ и $\|\cdot\|_Y$, например так:

1) $\|< x, y >\|_{X\times Y} = \max\{\|x\|_X, \|y\|_Y\}$,

2) $\|< x, y >\|_{X\times Y} = \|x\|_X + \|y\|_Y$,

3) $\|< x, y >\|_{X\times Y} = \sqrt{\|x\|_X^2 + \|y\|_Y^2}$.

Эти нормы эквивалентны в следующем смысле: из сходимости в пространстве по одной норме следует сходимость в этом же пространстве по другой норме.

Рассмотрим $\|< x, y >\|_{X\times Y} = \max\{\|x\|_X, \|y\|_Y\}$.

Окрестность точки $< x_0, y_0 >$ в пространстве $X \times Y$ представляет собой шар

$$\{< x, y >: \|< x, y > - < x_0, y_0 >\| < r\}.$$

Преобразуем неравенство из определения шара.

$$\|< x, y > - < x_0, y_0 >\| < r \Leftrightarrow \|< x - x_0, y - y_0\| < r \Leftrightarrow$$
$$\Leftrightarrow \max\{\|x - x_0\|_X, \|y - y_0\|_Y\} < r.$$

Получим шар, имеющий вид

$$B_{< x_0, y_0 >, r} = \{< x, y >: \|x - x_0\| < r, \|y - y_0\| < r\}.$$

В пространстве X имеем

$$B_{x_0, r} = \{x : \|x - x_0\| < r\}.$$

В пространстве Y имеем

$$B_{y_0, r} = \{y : \|y - y_0\| < r\}.$$

Тогда шар $B_{< x_0, y_0 >, r}$ представляет собой декартово произведение шаров в пространствах X и Y:

$$B_{< x_0, y_0 >, r} = B_{x_0, r} \times B_{y_0, r}.$$

2. Неявные функции

Пусть X, Y, Z – нормированные пространства.

Рассмотрим уравнение

$$F(x, y) = \theta \quad (1),$$

где $F : W \leftarrow Z$, W – окрестность точки $< x_0, y_0 > \in X \times Y$.

Пусть точка $< x_0, y_0 >$ – решение уравнения (1).

Выясним, существует ли для каждого x из некоторой окрестности точки x_0 значение y из некоторой окрестности точки y_0, при котором $F(x, y) = \theta$ (другое решение при x, близких к точке x_0).

Определение. Если для любого x из некоторой окрестности U_0 точки x_0 существует такой единственный элемент y из некоторой окрестности точки $y_0 \in Y$, что $F(x, y) = \theta$, то говорят, что существует ***неявная функция***, определенная на U_0, со значениями в пространстве Y, которая задается уравнением (1).

Введем обозначение $f : U_0 \to Y$ – неявная функция, определенная уравнением (1).

Тогда $F(x, f(x)) = \theta \quad \forall x \in U_0$.

Теорема о неявных функциях

Пусть выполнены условия:

1) X, Y, Z – нормированные пространства, пространство Y полно, точка $< x_0, y_0 > \in X \times Y$, W – окрестность точки $< x_0, y_0 >$, оператор $F(\cdot, \cdot) : W \to Z$, $F(x_0, y_0) = \theta$ ($< x_0, y_0 >$ – решение уравнения (1));

2) оператор F непрерывен в точке $< x_0, y_0 >$;

3) оператор F дифференцируем по Фреше по второму аргументу (по y) при $< x, y > \in W$, оператор $F'_y(x, y)$ непрерывен в точке $< x_0, y_0 >$ как отображение $F'_y(\cdot, \cdot) : W \to N(Y, Z)$, где $N(Y, Z)$ – пространство линейных непрерывных операторов, действующих из Y в Z;

4) оператор $F'_y(x_0, y_0)$ обратим, линейный оператор $\left(F'_y(x_0, y_0)\right)^{-1}$ непрерывен.

Тогда существует такая окрестность U_0 точки $x_0 \in X$, что на U_0 определена неявная функция $f : U_0 \to Y$, заданная уравнением (1); при этом неявная функция f непрерывна в точке x_0.

Идея доказательства. Надо доказать существование такой окрестности точки x_0, что для любых x из этой окрестности существует решение y уравнения (1), единственное для каждого такого x. Для этого вместо уравнения (1), т.е. уравнения $F(x, y) = \theta$, рассматривается уравнение

$$y = A_{(x)} y, \tag{2}$$

где $A_{(x)} y = y - (F'_y(x_0, y_0))^{-1} F(x, y)$.

Уравнение (2) эквивалентно уравнению (1). Поэтому надо доказать существование такой окрестности точки x_0, что для любых x из этой окрестности существует решение y уравнения (2), единственное для каждого такого x.

Для такого доказательства используется принцип сжимающих отображений. Но доказать сжимаемость оператора $A_{(x)}$ на всем пространстве Y совершенно невозможно, поэтому остается использовать обобщение принципа сжимающих отображений, в котором предполагается сжатие на некоторой части пространства Y (раздел 2.14, теорема 2).

А для использования этой теоремы показано, что для любых x из некоторой окрестности U точки x_0 найдется такой замкнутый шар $\|y - y_0\| \le \varepsilon$, который оператор $A_{(x)}$ преобразует в себя и является на этом шаре оператором сжатия. Для этого используются теоремы 3 и 4 из раздела 2.14.

По теореме 2 из того же раздела для любого x из упомянутой окрестности U точки x_0 существует единственное решение $y \in Y$ уравнения (2) и, следовательно, уравнения (1). А это означает, что в окрестности U определена неявная функция, заданная уравнением $F(x,y)=\theta$.

Заключение

В данной работе описана лишь часть материала, относящегося к нелинейному функциональному анализу. Рассмотрены понятия производных Фреше и Гато, а также их различные применения. Но при этом не затрагивались такие разделы, как топологические методы нелинейного функционального анализа, функциональный анализ в полуупорядоченных пространствах и другие области нелинейного анализа. Некоторые из этих разделов изложены, например, в книге [6].

4. Ответы к задачам для самостоятельного решения

Ответы к задачам раздела 2.1

1. $Ax = \begin{pmatrix} x_1^2 \\ x_1 - x_2 \end{pmatrix}, \quad Bx = \begin{pmatrix} x_2^2 \\ x_1 \end{pmatrix}.$

Решение первым способом.

$$ABx = A(Bx) = \begin{pmatrix} x_2^4 \\ x_2^2 - x_1 \end{pmatrix},\ (AB)'(x) = \begin{pmatrix} 0 & 4x_2^3 \\ -1 & 2x_2 \end{pmatrix}.$$

Решение вторым способом.

$$A'(x) = \begin{pmatrix} 2x_1 & 0 \\ 1 & -1 \end{pmatrix},\ A'(Bx) = \begin{pmatrix} 2x_2^2 & 0 \\ 1 & -1 \end{pmatrix},\ B'(x) = \begin{pmatrix} 0 & 2x_2 \\ 1 & 0 \end{pmatrix},$$

$$(AB)'(x) = A'(Bx)B'(x) = \begin{pmatrix} 2x_2^2 & 0 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 0 & 2x_2 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 4x_2^3 \\ -1 & 2x_2 \end{pmatrix}.$$

Результаты решения двумя способами совпадают.

2. $Ax = \begin{pmatrix} x_2^2 \\ x_1 - x_2 \end{pmatrix}, \quad Bx = \begin{pmatrix} x_2^2 \\ x_1 \end{pmatrix}.$

Решение первым способом.

$$ABx = A(Bx) = \begin{pmatrix} x_1^2 \\ x_2^2 - x_1 \end{pmatrix},\ (AB)'(x) = \begin{pmatrix} 2x_1 & 0 \\ -1 & 2x_2 \end{pmatrix}.$$

Решение вторым способом.

$$A'(x) = \begin{pmatrix} 0 & 2x_2 \\ 1 & -1 \end{pmatrix},\ A'(Bx) = \begin{pmatrix} 0 & 2x_1 \\ 1 & -1 \end{pmatrix},\ B'(x) = \begin{pmatrix} 0 & 2x_2 \\ 1 & 0 \end{pmatrix},$$

$$(AB)'(x) = A'(Bx)B'(x) = \begin{pmatrix} 0 & 2x_1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 0 & 2x_2 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2x_1 & 0 \\ -1 & 2x_2 \end{pmatrix}.$$

Результаты решения двумя способами совпадают.

3. $Ax = \begin{pmatrix} x_1 + x_2 \\ x_1 - x_2 \end{pmatrix}, \quad Bx = \begin{pmatrix} x_2^2 \\ x_1 \end{pmatrix}.$

Решение первым способом.

$$ABx = A(Bx) = \begin{pmatrix} x_2^2 + x_1 \\ x_2^2 - x_1 \end{pmatrix}, \ (AB)'(x) = \begin{pmatrix} 1 & 2x_2 \\ -1 & 2x_2 \end{pmatrix}.$$

Решение вторым способом.

$$A'(x) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \ A'(Bx) = \begin{pmatrix} 0 & 2x_1 \\ 1 & -1 \end{pmatrix}, \ B'(x) = \begin{pmatrix} 0 & 2x_2 \\ 1 & 0 \end{pmatrix},$$

$$(AB)'(x) = A'(Bx)B'(x) = \begin{pmatrix} 0 & 2x_1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 0 & 2x_2 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2x_1 \\ -1 & 2x_2 \end{pmatrix}.$$

Результаты решения двумя способами совпадают.

4. $Ax = \begin{pmatrix} x_1 + x_2 \\ x_1^2 \end{pmatrix}, \quad Bx = \begin{pmatrix} x_2^2 \\ x_1 \end{pmatrix}.$

Решение первым способом.

$$ABx = A(Bx) = \begin{pmatrix} x_2^2 + x_1 \\ x_2^4 \end{pmatrix}, \ (AB)'(x) = \begin{pmatrix} 1 & 2x_2 \\ 0 & 4x_2^3 \end{pmatrix}.$$

Решение вторым способом.

$$A'(x)=\begin{pmatrix}1 & 1\\ 2x_1 & 0\end{pmatrix},\quad A'(Bx)=\begin{pmatrix}1 & 1\\ 2x_2^2 & 0\end{pmatrix},\quad B'(x)=\begin{pmatrix}0 & 2x_2\\ 1 & 0\end{pmatrix},$$

$$(AB)'(x)=A'(Bx)B'(x)=\begin{pmatrix}1 & 1\\ 2x_2^2 & 0\end{pmatrix}\begin{pmatrix}0 & 2x_2\\ 1 & 0\end{pmatrix}=\begin{pmatrix}1 & 2x_2\\ 0 & 4x_2^3\end{pmatrix}.$$

Результаты решения двумя способами совпадают.

Ответы к задачам раздела 2.9

1. $Ax=\begin{pmatrix}x_1^3+2x_1+3\\ x_1x_2-x_1+1\end{pmatrix},\ x^*=\begin{pmatrix}1\\ 0\end{pmatrix}.$

Ряд Тейлора оператора A в окрестности точки x^*:

$$A(x^*+h)=\begin{pmatrix}6\\ 0\end{pmatrix}+\begin{pmatrix}5h_1\\ -h_1+h_2\end{pmatrix}+\frac{1}{2!}\begin{pmatrix}6h_1^2\\ 2h_1h_2\end{pmatrix}+\frac{1}{3!}\begin{pmatrix}6h_1^3\\ 0\end{pmatrix}.$$

Значение оператора A в точке $\theta=\begin{pmatrix}0\\ 0\end{pmatrix}$, очевидно, равно $\begin{pmatrix}3\\ 1\end{pmatrix}$.

То же через разложение в ряд Тейлора:

$$A\theta=A\left(x^*+\begin{pmatrix}-1\\ 0\end{pmatrix}\right)=\begin{pmatrix}6\\ 0\end{pmatrix}+\begin{pmatrix}-5\\ 1\end{pmatrix}+\begin{pmatrix}3\\ 0\end{pmatrix}+\begin{pmatrix}-1\\ 0\end{pmatrix}=\begin{pmatrix}3\\ 1\end{pmatrix}.$$

2. $Ax=\begin{pmatrix}x_1^2x_2-2x_1x_2+1\\ x_1^2+3x_1+2\end{pmatrix},\ x^*=\begin{pmatrix}1\\ 1\end{pmatrix}.$

Ряд Тейлора оператора A в окрестности точки x^*:

$$A(x^*+h)=\begin{pmatrix}0\\ 6\end{pmatrix}+\begin{pmatrix}h_2\\ 5h_1\end{pmatrix}+\frac{1}{2!}\begin{pmatrix}2h_1^2\\ 2h_1h_2\end{pmatrix}+\frac{1}{3!}\begin{pmatrix}6h_1^2h_2\\ 0\end{pmatrix}.$$

Значение оператора A в точке $\theta = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, очевидно, равно $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

То же через разложение в ряд Тейлора:

$$A\theta = A\left(x^* + \begin{pmatrix} -1 \\ -1 \end{pmatrix} \right) = \begin{pmatrix} 0 \\ 6 \end{pmatrix} + \begin{pmatrix} 1 \\ -5 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

3. $Ax = \begin{pmatrix} x_1^3 + x_1x_2 + 3 \\ x_1x_2 - 3x_1 + 2 \end{pmatrix}$, $x^* = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Ряд Тейлора оператора A в окрестности точки x^*:

$$A(x^* + h) = \begin{pmatrix} 5 \\ 0 \end{pmatrix} + \begin{pmatrix} 4h_1 + h_2 \\ -2h_1 + h_2 \end{pmatrix} + \frac{1}{2!}\begin{pmatrix} 6h_1^2 + 2h_1h_2 \\ 2h_1h_2 \end{pmatrix} + \frac{1}{3!}\begin{pmatrix} 6h_1^3 \\ 0 \end{pmatrix}.$$

Значение оператора A в точке $\theta = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, очевидно, равно $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$.

То же через разложение в ряд Тейлора:

$$A\theta = A\left(x^* + \begin{pmatrix} -1 \\ -1 \end{pmatrix} \right) = \begin{pmatrix} 5 \\ 0 \end{pmatrix} + \begin{pmatrix} -5 \\ 1 \end{pmatrix} + \begin{pmatrix} 4 \\ 1 \end{pmatrix} + \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

4. $Ax = \begin{pmatrix} x_1x_2^2 + 4x_2 + 2 \\ x_1^2 - x_1x_2 + 3 \end{pmatrix}$, $x^* = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$.

Ряд Тейлора оператора A в окрестности точки x^*:

$$A(x^* + h) = \begin{pmatrix} 2 \\ 4 \end{pmatrix} + \begin{pmatrix} 4h_2 \\ -2h_1 + h_2 \end{pmatrix} + \frac{1}{2!}\begin{pmatrix} -2h_2^2 \\ 2h_1^2 - 2h_1h_2 \end{pmatrix} + \frac{1}{3!}\begin{pmatrix} 6h_1h_2^2 \\ 0 \end{pmatrix}.$$

Значение оператора A в точке $\theta = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, очевидно, равно $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$.

То же через разложение в ряд Тейлора:

$$A\theta = A\left(x^* + \begin{pmatrix}1\\0\end{pmatrix}\right) = \begin{pmatrix}2\\4\end{pmatrix} + \begin{pmatrix}0\\-2\end{pmatrix} + \begin{pmatrix}0\\1\end{pmatrix} + \begin{pmatrix}0\\0\end{pmatrix} = \begin{pmatrix}2\\3\end{pmatrix}.$$

Ответы к задачам раздела 2.12

1. $\left(-\frac{7}{3}, \frac{28}{3}\right)$ – точка максимума.

2. $\left(\frac{1}{4}, -\frac{1}{8}\right)$ – точка минимума.

3. Нет точек экстремума.

4. (0; 1) – точка минимума.

В этой задаче все последовательные главные миноры матрицы Гессе в стационарной точке неотрицательны, но один из них равен нулю.

Очевидно, существует такая окрестность стационарной точки, в которой все главные миноры матрицы Гессе неотрицательны, отсюда следует, что стационарная точка – точка минимума.

Ответы к задачам раздела 2.13

1. $x_0(t) = t^4 + 3$ – точка минимума.
2. $x_0(t) = \frac{2}{t^2} - 1$ – точка максимума.
3. $x_0(t) = \frac{8}{t^3} - 1$ – точка минимума.
4. $x_0(t) = 2t^3 - 1$ – точка максимума.

Библиографический список

1. Колмогоров А.Н., Фомин С.В. Элементы теории функций и функционального анализа. М.: Наука, 1976.
2. Люстерник Л.А., Соболев В.И. Краткий курс функционального анализа. М.: Высшая школа, 1982.
3. Галлеев Э.М., Тихомиров В.М. Оптимизация. М.: Наука, 2000.
4. Васильев Ф.П. Методы решения экстремальных задач. М.: Наука, 1981.
5. Федорюк М.В. Обыкновенные дифференциальные уравнения. М.: Наука, 1980.
6. Мисюркеев И.В. Введение в нелинейный функциональный анализ. Пермь, 1968.

Содержание

Учебное издание

Еленская Елизавета Юрьевна
Еленский Юрий Наполеонович

Основы нелинейного функционального анализа

Учебно-методическое пособие

Редактор *Л. Л. Савенкова*
Корректор *Л. Л. Соболева*
Компьютерная верстка: *Е. Ю. Еленская*

Объем данных 1 Мб
Подписано к использованию 15.07.2022

Размещено в открытом доступе
на сайте www.psu.ru
в разделе НАУКА / Электронные публикации
и в электронной мультимедийной библиотеке ELiS

Издательский центр
Пермского государственного
национального исследовательского университета
614990, г. Пермь, ул. Букирева, 15