УДК 519.651

ПРИМЕНЕНИЕ ПОЛУПРОИЗВОДНОЙ В ЧИСЛЕННОМ АНАЛИЗЕ

© 2008 г. С. Е. Михеев

(198504 СПб, Университетский пр-т, 35, СПб. гос. ун-т)

e-mail: him2@mail.ru

Поступила в редакцию 27.09.2006 г.
Переработанный вариант 09.06.2007 г.

На основе некоторого обобщения понятия константы Липшица и разделенной разности предлагается аппарат для исследования численных методов. Его применением получены новые результаты по оцениванию удаленности решения нелинейного уравнения в банаховом пространстве от данной точки. Библ. 12.

Ключевые слова: полупроизводная, итерация, итеративный процесс, оценка удаленности, анализ сходимости, нелинейное уравнение.

1. ВВЕДЕНИЕ

В анализе сходимости одноточечных итеративных методов, в оценивании удаленности решения нелинейного уравнения от заданной точки удобным оказалось исчисление полупроизводных, развиваемое в разд. 2. Оно позволяет полноценно использовать метод дифференцирования по итерации (разд. 3) для оценок погрешностей приближений в итеративных метода. Иллюстрация к применению исчисления полупроизводных дана в разд. 4 в доказательствах теорем 1 и 2. Там уточняются и обобщаются оценки удаленности для уравнений, подпадающих под теоремы Канторовича, Мысовских и Гавурина о методе Ньютона. Эти теоремы содержат оценки удаленности, которые и можно сравнивать с оценкой теоремы 1. Для уравнений типа П, левые части которых являются отображениями класса П (см. ниже определение 4) – класса, не содержащегося в классе $C^{1,1}$ и не содержащим класса $C^{1,1}$, к которому применяются упомянутые выше теоремы. С помощью исчисления полупроизводных получены оценки удаленности решения. Примером применения метода дифференцирования по итерации служат теоремы 3, 4 о сходимости метода Ньютона для уравнений типа П. Другое применение метода дифференцирования по итерации можно найти в [1], где дано обоснование итеративной минимизации с помощью квадратичных пошаговых аппроксимаций при нежестких ограничениях (см. [2,], [3]).

Пусть U, W – банаховы пространства (B-пространства), $M \subset U$ и имеется отображение A: $M \longrightarrow W$.

Определение 1. Пусть x – не изолированная точка множества M и величина

$$L_A(x) := \overline{\lim_{\Delta \to 0}} \|A(x+\Delta) - A(x)\|_W / \|\Delta\|_U \tag{1}$$

конечна. Назовем тогда ее *полупроизводной отображения A в точке x*. Здесь пределе берется по таким Δ, что $x + \Delta \in M$.

Если M не имеет изолированных точек и полупроизводная $L_A(x)$ существует для всех $x \in M$, то формула (1) определяет на M функцию, которую будем именовать *полупроизводной отображения A на множестве M*.

Далее индекс нормы, указывающий на пространство, в котором она определена, будет опускаться, поскольку его всегда можно выяснить из контекста.

Согласно определению, полупроизводную отображения A в точке x можно было бы еще назвать лучшей (минимальной) константой Липшица отображения A в точке x. Но, с с одной стороны, возникли бы трудности при именовании отображения $L_A(x)$ (т.е. получилась бы “переменная” константа). С другой стороны, известно понятие локальной липшицевости как липшицевости в некоторой окрестности, к которому можно было бы применить предельный переход при стягивании этой окрестности в точку x с выбором лучшей константы Липшица. Но такое построение приводит к иному, в общем случае возможно большему, чем $L_A(x)$, числу.

2. СВОЙСТВА ПОЛУПРОИЗВОДНОЙ

Очевидно, что если в точке x существует полупроизводная $L_A(x)$, то справедлива оценка

$$\|A(x+\Delta)-A(x)\| \le L_A(x)\|\Delta\| + o(\|\Delta\|, x),$$

т.е. A непрерывно в x (o – бесконечно малая относительно первого аргумента). Отсюда же нетрудно установить, что если отображение A дифференцируемо во внутренней точке x из M по Фреше, то его полупроизводная есть $\|A'(x)\|$.

Если A – вещественная функция вещественного аргумента – имеет в точке x конечные производные числа, то ее полупроизводная $L_A(x)$ есть максимум из их абсолютных значений (правое верхнее производное число есть правый верхний предел разделенной разности $[A(x+\delta)-A(x)]/\delta$ при $\delta \searrow 0$, аналогично определяются остальные три производные числа, (см. [4]). Если отображение A имеет в M – множестве без изолированных точек – константу Липшица L, то оно имеет в M полупроизводную, ограниченную сверху числом L. Для выпуклого множества верно и обратное.

Лемма 1. *Пусть полупроизводная отображения A ограничена константой L в выпуклой области задания M: $L_A(x) \le L$ $\forall x \in M$. Тогда отображение A липшицево в т с константой L.*

Доказательство. Пусть утверждение леммы неверно. Тогда найдутся x, y из M и $\varepsilon > 0$ такие, что $\|A(y)-A(x)\| > (L+\varepsilon)t'$, где $t' = \|y-x\|$. Параметризуем отрезок, соединяющий x и y: $y(t) = x + t(y-x)/t'$, $t \in [0, t']$. Пусть $N(t) := \|A(y(t))-A(x)\|$.

(Здесь и далее то, что стоит справа от знака :=, есть обозначение выражения. Стоящего слева от этого знака, а для знака =: – наоборот. Такие удобные асимметричные обозначения с линейной записью в настоящее время все чаще заменяют давно используемые симметричные двухэтажные.)

С одной стороны, $N(t)/t \longrightarrow L_A(x) \le L$ при $t \longrightarrow 0$, с другой – верно $N(t')/t' > L + \varepsilon$. Поэтому в силу непрерывности $A(y(t))$ по t существуют момент $T \in [0, \|y-x\|)$ такой, что $N(T) = (L+\varepsilon)T$, и монотонно убывающая до T последовательность $\{t_i\}_1^\infty$, на которой верно неравенство $N(t_i) > (L + \varepsilon)t_i$. Согласно обозначениям, получаем

$$\|A(y(t_i))-A(y(T))\| \ge \|A(y(t_i))-A(x)\| - \|A(x)-A(y(T))\| = N(t_i)-N(T) > (L+\varepsilon)(t_i-T).$$

Отсюда следует оценка $L_A(y(T)) \ge L + \varepsilon$, противоречащая условию леммы.

Когда значения отображения суть линейные ограниченные операторы из одного банахового пространства в другое, будем именовать его семейством операторов. Таким образом, для семейства операторов имеет смысл понятие полупроизводной. Для краткости записи оператора, обратного к оператору $A(y)$ из семейства $\{A(y)\}$, будет использоваться $A^{-1}(y)$ вместо более корректного $[A(y)]^{-1}$. Через $S(x, \delta)$ обозначим здесь замкнутый шар радиуса δ с центром в x.

Лемма 2. *Пусть Z, U, W суть* B-*пространства и множество M лежит в Z. Если определенное на множестве M семейство линейных ограниченных операторов $A(y)$: $U \longrightarrow W$, $y \in M$, имеет полупроизводную $L_A(x)$ в точке x из M и оператор $A(x)$ имеет ограниченный обратный $A^{-1}(x)$: $W \longrightarrow U$, то существует $\delta > 0$ такое, что для всех y из окрестности $V := S(x, \delta) \cap M$ операторы $A(y)$ имеют обратные, а семейство обратных операторов $A^{-1}(y)$: $W \longrightarrow U$, $y \in V$, равномерно ограничено, в точке x тоже имеет полупроизводную $L_{A^{-1}}(x)$ и*

$$L_{A^{-1}}(x) \le \|A^{-1}(x)\|^2 L_A(x).$$

Доказательство. Из существования полупроизводной в точке x следует непрерывность исходного семейства в этой точке. Следовательно, существует число $\delta > 0$ такое, что для всех y из окрестности $V = S(x, \delta) \cap M$ верно неравенство $\|A(y)-A(x)\| \le \|A^{-1}(x)\|^{-1}/2$. А тогда, согласно теореме Банаха и следствий из нее (см. [6, гл. V, теорема 4.4]), существуют обратные операторы $A^{-1}(y)$ в V, которые равномерно ограничены:

$$\|A^{-1}(y)\| \le \frac{\|A^{-1}(x)\|}{1-\|A^{-1}(x)\|\|A(y)-A(x)\|} \le 2\|A^{-1}(x)\| \quad \forall y \in V.$$

Пусть $y \in V$, тогда, принимая во внимание

$$A^{-1}(y) - A^{-1}(x) = A^{-1}(x)[A(x) - A(y)]A^{-1}(y),$$

имеем

$$\|A^{-1}(y) - A^{-1}(x)\| \le L_A(x)\|A^{-1}(x)\|\|y - x\|\|A^{-1}(y)\| + o(\|y - x\|, x). \quad (2)$$

В силу ограниченности $\|A^{-1}(y)\|$, из оценки (2) следует непрерывность A^{-1} в x.

Поэтому, разделив на $\|y - x\|$ неравенство (2) и перейдя к верхнему пределу при $y \longrightarrow x$, получим утверждение леммы.

Лемма 3. *Пусть для каждой точки s интервала* $[0, \bar{s}\,]$ *имеет место оценка полупроизводной вещественной функции f вещественного аргумента* $L_f(x) \le z(s)$ *и* $z(s)$ *конечна, неотрицательна и суммируема по Лебегу. Тогда справедлива оценка*

$$|f(s) - f(0)| \le \int_0^s z(t)dt \quad \forall s \in [o, \bar{s}]. \quad (3)$$

Лемма 3 – элементарное следствие леммы 3.8 из [1].

Лемма 4. *Пусть отображение A имеет в точке x полупроизводную* $L_A(x)$, *а скалярная функция f определена в некоторой окрестности точки* $y := \|A(x)\|$ *и существует полупроизводная* $L_f(y)$. *Тогда суперпозиция* $G(x) := f(\|A(x)\|)$ *имеет в x полупроизводную, которая удовлетворяет оценке* $L_G(x) \le L_f(y)L_A(x)$.

Доказательство. Оценим приращение суперпозиции G при изменении аргумента на Δ столь малого, что $x + \Delta$ принадлежит области задания отображения A, а $\|A(x + \Delta)\|$ – области задания функции f:

$$|G(x + \Delta) - G(x)| = |f(\|A(x + \Delta)\|) - f(y)| \le L_f(y)|\|A(x + \Delta)\| - y| + o(|\|A(x + \Delta)\| - y|, y) \le$$
$$\le L_f(y)\|A(x + \Delta) - A(x)\| + o_1(\|\Delta\|, y) \le L_f(y)L_A(x)\|\Delta\| + o_1(\|\Delta\|, y) + o_2(\|\Delta\|, A(x)).$$

Поделив на $\|\Delta\|$ и перейдя к верхнему пределу, получим утверждение леммы.

Лемма 5. *Пусть Z, U, W суть* B-*пространства и ограниченное выпуклое множество M лежит в Z. Если определенное на M семейство линейных ограниченных операторов* $A(x)$: $U \longrightarrow W$, $x \in M$, *на M имеет полупроизводную* $L_A(x)$ *и ограниченные обратные* $A^{-1}(x)$ *такие, что* $\|A^{-1}(x)\|L_A(x) \le \sigma$ $\forall x \in M$, *то семейство обратных операторов* $\{A^{-1}(x)\}_{x \in M}$ *ограничено снизу по норме положительной величиной, а семейства операторов* $\{A(x)\}_{x \in M}$ *и* $\{A^{-1}(x)\}_{x \in M}$ *липшицевы на M*.

Доказательство. Согласно условию и лемме 2 имеем

$$L_{A^{-1}}(x) \le \|A^{-1}(x)\|^2 L_A(x) \le \sigma\|A^{-1}(x)\| \quad \forall x \in M.$$

Принимая во внимание лемму 4, получаем оценку $L_{\ln\|A^{-1}\|}(x) \le \sigma$.

Отсюда и на основании леммы 1 имеем $|\ln\|A^{-1}(y)\| - \ln\|A^{-1}(x)\|| \le \sigma\|y - x\|$.

Потенцируя и обозначая через R диаметр множества M, получаем неравенства

$$\|A^{-1}(y)\| \ge \|A^{-1}(x)\|e^{-\sigma\|y - x\|} \ge \|A^{-1}(x)\|e^{-\sigma R} > 0.$$

Таким образом, доказано первое утверждение леммы 5. Второе следует из леммы 1 и того, что $L_A(y) \le \sigma/\|A^{-1}(y)\| \le \sigma e^{\sigma R}/\|A^{-1}(x)\|$ $\forall y \in M$.

Наличие константы Липшица для A^{-1} на M вытекает из лемм 2 и 1.

3. АНАЛИЗ СХОДИМОСТИ С ПОМОЩЬЮ ДИФФЕРЕНЦИРОВАНИЯ ПО ИТЕРАЦИИ

Прием, именуемый далее дифференцированием по итерации, можно использовать в исследовании методов, использующих недифференцируемые функции, но обладающих полупроизводной. Он будет применен в следующем разделе при доказательстве теорем 1 и 2.

Суждения о сходимости к решению α итераций $x^0, x^1, \ldots$ можно выводить из оценки $\|x^{k+1} - \alpha\| \le \le \|\alpha - x^k\|^p$. Ее, в свою очередь, часто удобно получать следующим образом.

Пусть поиск решения $\alpha \in U$ некоторой исходной задачи сведен к поиску решения уравнения $G(0, \alpha) = 0$ итерациями $\{x^k\} \subset U$, определяемыми рекуррентно из уравнений

$$G(x^{k+1} - x^k, x^k) = 0, \quad k = 0, 1, \ldots, \tag{4}$$

где $G : U \times U \longrightarrow W$, U, W суть B-пространства. Предполагается, естественно, что существует достаточно простой (сравнительно с исходной задачей) алгоритм разрешения уравнения $G(x, y) = 0$ относительно первого аргумента.

Соединим отрезком текущую итерацию x^k с решением α исходной задачи и параметризуем его:

$$x^k(t) := \alpha + \bar{d}t, \quad \bar{d} := x^k - \alpha, \quad t \in [0, 1]. \tag{5}$$

После подстановки в (4) переменной $x^k(t)$ вместо постоянной x^k получится семейство уравнений с параметром t. Обозначим решение такого уравнения относительно x^{k+1} через $x(t)$ и введем переменную $y := x(t) - \alpha$. Поскольку $G(0, \alpha) = 0$, справедливо $y(0) = 0$.

Если G имеет производные по первому и второму аргументам (G'_1, G'_2), то тождество $G(y - \bar{d}t, x_k(t)) \equiv 0$ можно продифференцировать по t:

$$G'_1(y - \bar{d}t, x_k(t))(\dot{y} - \bar{d}) + G'_2(y - \bar{d}t, x_k(t))\bar{d} = 0. \tag{6}$$

Если $G'_1(y - \bar{d}t, x_k(t))$ не вырождена, то уравнение (6) можно разрешить относительно $\dot{y}$:

$$\dot{y} = [I - (G'_1(y - \bar{d}t, x_k(t))]^{-1} G'_2(y - \bar{d}t, x_k(t)))\bar{d}. \tag{7}$$

Отсюда получаем дифференциальное неравенство

$$\|\dot{y}\| \le \left\|I - [G'_1(y - \bar{d}t, x_k(t))]^{-1} G'_2(y - \bar{d}t, x_k(t))\right\| \|\bar{d}\|. \tag{8}$$

Если получить оценку

$$\left\|I - [G'_1(y - \bar{d}t, x_k(t))]^{-1} G'_2(y - \bar{d}t, x_k(t))\right\| \le f(\|y\|, t),$$

то (8) совместно с очевидной оценкой $\|y\|'_t \le \|\dot{y}\|$ и с $y(0) = 0$ даст систему

$$(\|y\|)'_t \le f(\|y\|, t)\|\bar{d}\|, \quad \|y(0)\| = 0. \tag{9}$$

Если f непрерывна в некотором открытом множестве M, то (см. теорему III.4.1 из [5]) максимальное решение задачи Коши

$$\dot{z} = f(z, t)\|\bar{d}\|, \quad z(0) = \|y(0)\| = 0 \tag{10}$$

мажорирует любое решение системы (9) относительно скалярной функции $\|y\|$ на общем интервале существования $[0, a]$. Каким же должно быть эта a, чтобы извлечь из решения задачи Коши (10) оценку сверху на $\|x^{k+1} - \alpha\|$?

Согласно построению, $\|x_{k+1} - \alpha\| = \|y(1)\|$. Следовательно, нужно использовать оценку $\|y(1)\| \le z(1)$. Поэтому общий интервал существования должен содержать $[0, 1]$, т.е. должно быть обеспечено неравенство $a \ge 1$. Пусть к этому указывает следствие из теоремы Пеано.

Лемма 6. *Определим для положительных b прямоугольник $R_b = [-b, b] \times [0, 1]$. Пусть существует b такое, что*

$$\sup_{(z,t) \in R_b} |f(z, t)| \le b.$$

Пусть f непрерывна в R_b. Тогда задача Коши (10) *имеет на сегменте* $[0, 1]$ *хотя бы одно решение, мажорирующее любое решение системы* (9) *на общем интервале существования.*

Согласно построению и допущениям относительно G, функция y определена на $[0, 1]$ и удовлетворяет условиям (9). Поэтому когда решение задачи Коши (10) существует на сегменте

$[0, a')$ и $a' \geq 1$ (что будет, например, если выполнены условия последней леммы), тогда справедлива оценка

$$\|x^{k+1} - \alpha\| \leq z(1) =: \psi(\|\bar{d}\|) \equiv \psi(\|x^k - \alpha\|), \tag{11}$$

чего часто бывает достаточно для анализа сходимости и скорости сходимости.

В тех случаях, когда функция G всего лишь липшицева и не дифференцируема, существенное удобство в выкладках предоставляет несколько иной взгляд на понятие разностного отношения $h(t, \tau) := \dfrac{h(t+\tau) - h(t)}{\tau}$ функции $h(\eta)$ скалярного аргумента η из сегмента D. Обычно первый аргумент разностного отношения считается фиксированным, а в оценках, которые предстоит сделать, он будет переменным. В описании такой смысловой нагрузки на параметр удобно.

Определение 2. Разностной производной в точке $t \in D$, заданной на сегменте $D \subset R^1$ вектор-функции $h : D \longrightarrow U$ (U есть В-пространство), будем называть зависящую от параметра t вектор-функцию $h^\nabla(t, \cdot)$, заданную в $\hat{D}(t, D) := (D - t)\backslash\{0\}$ формулой

$$h^\nabla(t, \varepsilon) = \frac{h(t+\varepsilon) - h(t)}{\varepsilon} \quad \forall \varepsilon \in \hat{D}(t, D). \tag{12}$$

Величину ε будем называть *отклонением* (*разностным*), а операцию нахождения разностной производной – *разностным дифференцированием*.

Допуская некоторую вольность, будем опускать второй параметр разностной производной, если это не исказит смысл формулы.

У разностной производной есть важное достоинство по сравнению с обычной производной: она существует и конечна всегда, когда исходная функция всего лишь определена и принимает конечные значения. Вместе с тем она обладает свойствами, сходными со свойствами обычной производной.

Свойство 1. $(f \pm g)^\nabla = f^\nabla \pm g^\nabla$.

Свойство 2. $(\tau f)^\nabla = \tau f^\nabla$, τ – число.

Свойство 3. Если f дифференцируема в точке t, то $f^\nabla(t, \varepsilon) = f'(t) + \omega(\varepsilon) \cong f'(t)$, где $\lim\limits_{\varepsilon \to 0} \omega(\varepsilon) = 0$.

Здесь и далее под знаками $\cong$ и $\lesssim$ будет пониматься, соответственно, равенство и неравенство, приближенные с точностью до величин, стремящихся к $0 \in U$, когда к 0 стремится отклонение в разностной производной.

Свойство 4. Пусть U, V, W суть В-пространства и на $U \times V$ задана бинарная операция, сопоставляющая паре (x, χ), $x \in U$, $\chi \in V$, элемент из W. Пусть эта операция записывается мультипликативно и удовлетворяет двум аксиомам:

1) $(\forall x, y \in U)(\forall \chi \in V)\ (x + y)\chi = x\chi + y\chi$;

2) $(\forall x \in U)(\forall \chi, \eta \in V)\ x(\chi + \eta) = x\chi + x\eta$.

(Такому построению удовлетворяют, например, в качестве В-пространств U, V, W множества матриц размерностью $(m \times n)$, $(n \times k)$, $(m \times k)$ соответственно и взятое в качестве бинарной операции стандартное произведение матриц).

Пусть имеются вектор-функции $f : D \longrightarrow U$ и $g : D \longrightarrow V$ (D – сегмент из R^1). Тогда fg можно трактовать как вектор-функцию $fg : D \longrightarrow W$, разностную производную которой интересно бы выразить через разностные производные f и q.

Пусть g имеет полупроизводную в t. Тогда справедливо

$$g(t+\varepsilon) \equiv g^\nabla(t, \varepsilon)\varepsilon + g(t) \cong g(t). \tag{13}$$

С другой стороны,

$$(fg)^\nabla(t, \varepsilon) = [f(t+\varepsilon)g(t+\varepsilon) - f(t)g(t)]/\varepsilon =$$

$$= \{[f(t+\varepsilon) - f(t)]g(t+\varepsilon) + f(t)[g(t+\varepsilon) - g(t)]\}/\varepsilon = f^\nabla(t)g(t+\varepsilon) + f(t)g^\nabla(t),$$

что совместно с (13) дает

$$(fg)^\nabla(t) \cong f^\nabla(t)g(t) + f(t)g^\nabla(t).$$

Заметим, что это свойство имеет место и при существовании в t полупроизводной вектор-функции f вместо полупроизводной вектор-функции g.

Свойство 5. Пусть на сегменте $D_g \subset R^1$ заданы вектор-функция $g : D_g \longrightarrow D_f \subset U$ и отображение $f : D_f \longrightarrow V$. Тогда определена суперпозиция $f(g) : D_g \longrightarrow V$. И если f дифференцируемо в точке $g(t)$ $(t \in D_g)$, то

$$(f(g))^{\nabla}(t, \varepsilon) = [f(g(t+\varepsilon) - f(g(t)))]/\varepsilon = f'(g(t))g^{\nabla}(t) + o(g(t+\varepsilon) - g(t))/\varepsilon \cong f'(g(t))g^{\nabla}(t).$$

Свойство 6. Пусть вектор-функция $g : D_g \longrightarrow D_f \subset U$ и отображение $f : D_f \longrightarrow V$ имеют полупроизводные L_g и L_f в точках t и $g(t)$ соответственно. Тогда

$$\|(f(g))^{\nabla}(t)\| \lessapprox L_f\|g^{\nabla}(t)\| \lessapprox L_f L_g.$$

Действительно, $\|(f(g))^{\nabla}(t, \varepsilon)\| \le L_f[\|g(t+\varepsilon) - g(t)\| - o(g(t+\varepsilon) - g(t))]/\varepsilon$ и при $\varepsilon \longrightarrow 0$

$$\frac{o(g(t+\varepsilon) - g(t))}{\varepsilon} = \frac{o(g(t+\varepsilon) - g(t))}{\|g(t+\varepsilon) - g(t)\|}\|g^{\nabla}(t)\| \longrightarrow 0.$$

Свойство 7. Пусть имеются вектор-функции $g : D \longrightarrow U$, $h : D \longrightarrow V$, которые определяют множество $D_f := \{(x, y) | (\exists t \in D) x = g(t) \wedge y = h(t)\}$. Пусть верно следующее:

1) g и h имеют полупроизводные L_g и L_h в точке $t \in D$;

2) $z := (g(t), h(t)) \in D_f \backslash \partial D_f$;

3) в точке z отображение f обладает производными Фреше по первому и второму аргументам: f'_g и f'_h.

Тогда

$$(f(g, h))_t^{\nabla}(t) \cong f'_g(g(t), h(t))g^{\nabla}(t) + f'_h(g(t), h(t))h^{\nabla}(t).$$

Это, очевидно, следует из разложения

$$(f(g, h))^{\nabla}(t, \varepsilon) = (f'_g(g(t), h(t)), f'_h(g(t), h(t)))\begin{pmatrix} g(t+\varepsilon) - g(t) \\ h(t+\varepsilon) - h(t) \end{pmatrix}\varepsilon^{-1} +$$
$$+ o(g(t+\varepsilon) - g(t), h(t+\varepsilon) - h(t))\varepsilon^{-1}.$$

Свойство 8. Пусть теперь из предположений свойства 7 изъято 2) и ослаблено 3): 3') в точке $z := (g(t), h(t))$ отображение f имеет полупроизводные по первому и второму аргументам: L_{fg} и L_{fh}.

Тогда $\|(f(q, h))_t^{\nabla}(t, \varepsilon)\| \lessapprox L_{fg}L_g + L_{fh}L_h$.

Действительно, справедлива цепь соотношений

$$\|(f(g, h))_t^{\nabla}(t, \varepsilon)\| \le L_{fg}\|g^{\nabla}(t, \varepsilon)\| + L_{fh}\|h^{\nabla}(t, \varepsilon)\| + o(g(t+\varepsilon) - g(t), h(t+\varepsilon) - h(t))/\varepsilon \cong$$
$$\cong L_{fg}\|(g^{\nabla}(t)\| + L_{fh}\|h^{\nabla}(t)\| \cong L_{fg}L_g + L_{fh}L_h.$$

Теорема III.4.1. из [5] позволяет применить метод дифференцирования по итерации при отсутствии G'_2 – производной функции G по второму аргументу и уйти от вопроса о существовании производной у $x(t)$. Для этого следует произвести разностное дифференцирование по t тождества $G(y - \bar{d}t, x_k(t)) \equiv 0$ и начать исследовать вместо (8) приближенное дифференциальное неравенство

$$\|y^{\nabla}\| \lessapprox \|I - [G'_1(y - \bar{d}t, x_k(t))]^{-1}G_2^{\nabla}(y - \bar{d}t, x_k(t))\|\|\bar{d}\| \tag{14}$$

с использованием приведенных выше свойств разностной производной.

Если удается оценить сверху величиной $f(\|y\|, t)\|\bar{d}\|$ правую часть (14), то вектор-функция $y \equiv x(t) - \alpha$ должна удовлетворять системе

$$\|y\|_t^{\nabla} \lessapprox f(\|y\|, t)\|\bar{d}\|, \quad \|y(0)\| = 0, \tag{15}$$

любое решение которой также мажорируется решением задачи Коши (10).

Когда f в системе (9) или в системе (15) получается не зависящей от первого аргумента $f(z, t)\|\bar{d}\| =: \varphi(t)$, то справедливо более сильное, чем лемма 6, утверждение.

Лемма 7. *Достаточно измеримости по Лебегу и конечности функции* φ *на* [0, 1] *для того, чтобы интеграл* $\mathcal{I}(t) := \int_0^t \varphi(\tau)d\tau$, $t \in [0, 1]$, *мажорировал* (*см.* [1]) *все решения систем*

$$\|y\|_t^{\nabla} \lessapprox \varphi(t), \quad \|y(0)\| = 0,$$

$$\|y\|_t' \le \varphi(t), \quad \|y(0)\| = 0.$$

Замечание 1. Интеграл $\mathcal{I}(t)$ является решением задачи Коши (10): $z(t) = \mathcal{I}(t)$ для всех t из [0, 1] (см. [7, п. 11.8.2.]).

4. ПРИМЕНЕНИЕ ПОЛУПРОИЗВОДНОЙ В ОЦЕНКЕ УДАЛЕННОСТИ РЕШЕНИЯ НЕЛИНЕЙНОГО УРАВНЕНИЯ

Рассматривается классическая задача о существовании и локализации решения уравнения

$$g(x) = 0, \quad x \in D \subset U, \tag{16}$$

для нелинейного отображения $g: D \longrightarrow W$ в паре банаховых пространств U, W (B-пространств). Для вынесения какого-либо суждения о существовании решения и об оценке его удаленности от некоторой заданной точки необходима дополнительная информация от отображении g и области задания D. Такие суждения как часть результатов по обоснованию метода Ньютона для (16) в работах Канторовича и Мысовских (см., например, [6] и [8]) получены, когда скорость изменения производной Фреше $J := g'$ в некотором замкнутом шаре $\Omega_0 := S(x_0, R) := \{x \| \|x - x_0\| \le R\}$ (x_0 – начальная точка метода), принадлежащем открытой области задания D, была стеснена либо условием

$$\sup_{x \in \Omega_0} \left\|J^{-1}(x_0)J'(x)\right\| \le K, \tag{17}$$

либо условием $\|J'(x)\| \le L$ $\forall x \in \Omega_0$. Во втором варианте непосредственно из доказательств в указанных трудах была видна возможность легкого расширения результатов и на случай, когда J всего лишь липшицева:

$$\|J(x + \Delta) - J(x)\| \le L\|\Delta\| \quad \forall x, x + \Delta \in \Omega_0, \tag{18}$$

что и было вскоре сделано другими авторами (см., например, [9]).

Дополнительное ограничение вида $\|J^{-1}(x)\| \le r_M$ $\forall x \in \Omega_0$ в теореме Мысовских (ТМ) о методе Ньютона (см. [8, теорема 1]) приводит к несколько иным результатам, которые и в оценке удаленности отличаются от теоремы Канторовича (ТК) (см. [6, теорема 6, § 1, гл. XVIII]). С этим же ограничением, исследуя непрерывный метод Ньютона, Гаврин получил точную оценку удаленности (см. [10]), которая, когда r_M совпадает с известной оценкой $r_0 \ge \|J^{-1}(x_0)\|$, строго меньше оценки в ТМ.

Весьма эффективным при решении рассматриваемой классической задачи оказалась полупроизводная отображения. С ее помощью в теореме 1 произведено уточнение оценки удаленности в условиях, близких к ТМ. Частный случай теоремы 1, когда r_0 совпадает с r_M, есть небольшое усиление теоремы Гаврина (ТГ) (см. [10, теорема 1]) в части, касающейся существования решения и оценки его удаленности.

Однотипным образом в теореме 2 исследуется существование и удаленность решения уравнения (16) при условиях, часть из которых отличается от ранее известных. При существовании J' – производной от J – одно из них принимает вид

$$\|J'(x)\|\left\|J^{-1}(x)\right\| \le \sigma \quad \forall x \in D\backslash\partial D.$$

Здесь под нормой линейного оператора: $A : U \longrightarrow W$ подразумевается норма, подчиненная нормам пространств U и W:

$$\|A\| := \sup_{x \ne 0}\|Ax\|_W/\|x\|_U.$$

Отметим, что в конечномерном B-пространстве можно после выбора базиса трактовать производную функции g как матрицу Якоби J (отсюда и обозначение).

В конечномерном пространстве для липшицевых J гарантию существования и оценку решения уравнения (16) предоставляет ТК о методе Ньютона с доказательством Ортеги в [11]. Там же приведена обширная библиография по данной тематике.

Далее будем полагать область задания D отображения g представимой в виде объединения открытой области D' и подмножества M' ее границы $\partial D'$, т.е. $D = D' \cup M'$, а отображение g – непрерывным в D и имеющим в D' невырожденную производную Фреше $J := q'$. Замкнутый шар $\{y|\|y - x\| \le \delta\}$ обозначается здесь через $S(x, \delta)$.

Для краткости записи оператора, обратного к линейному оператору $J(y) : U \longrightarrow W$, будет использоваться $J^{-1}(y)$ вместо более корректного $[J(y)]^{-1}$.

Получим точную и глобальную оценку удаленности решения в условиях ТМ с использованием липшицевости J вместо ограниченности его производной.

Теорема 1. *Пусть* U, W *суть* B-*пространства,* $D \subset U$, $g : D \longrightarrow W$ *и выполняются следующие условия:*

1) $\|g(x_0)\| \le \bar{g}_0$, $\|J^{-1}(x)\| \le r_0$;

2) J *липшицева в* D' *с константой* L;

3) $\|J^{-1}(x)\| \le r_M$ $\forall x \in D'$; *если оценка сверху на* $\|J^{-1}(x)\|$ *неизвестна, то полагается* $r_M = +\infty$;

4) *замкнутый шар* $S(x_0, d_1)$ *лежит в* D, *где*

$$d_1 := \begin{cases} d_K := \dfrac{1 - \sqrt{1 - 2P_L}}{Lr_0}, & P_L \le 1/2 \wedge T_M \ge d_K, \\ r_M\bar{g}_0 - \dfrac{(r_M/r_0 - 1)^2}{2r_M L}, & (r_M < \infty \wedge P_L \ge 1/2) \vee (P_L < 1/2 \wedge T_M \le d_K), \end{cases} \tag{19}$$

$P_L := Lr_0^2\bar{g}_0$, *если* $r_M < \infty$, *то* $T_M := \dfrac{1 - r_0/r_M}{Lr_0}$, *иначе* $T_M := \dfrac{1}{Lr_0}$.

Тогда верно следующее:

а) *в шаре* $S(x_0, d_1)$ *уравнение* (16) *имеет решение* α,

б) *являющееся предельной точкой фазовой траектории решения задачи Коши*

$$\dot{x} = -J^{-1}(x)g(x)/\left\|J^{-1}(x)g(x)\right\| =: F(x), \quad x(0) = x_0, \tag{20}$$

а именно существует $T \in [0, d_1]$ *такое, что решение этой задачи определено на* $[0, T)$, *непродолжимо вправо и* $\lim\limits_{t \to T - 0} x(t) = \alpha$, *причем* $\|g(x(t))\|$ *монотонно убывает до* 0 *на* $[0, T)$.

Кроме того, если $r_M < \infty$, *а оценка* r_0 *огрубляется до* r_M, *то условие* 2) *можно ослабить до локальной липшицевости* J *в* D'.

Комментарии к формулировке теоремы. Условия 1), 2), $P_L \le 1/2$ и $S(x_0, d_K) \subset D$ суть основные условия ТК в варианте, предлагаемом в [9] (дополнительные условия ТК обеспечивают единственность решения).

Условие 3) имеет два варианта: $r_M < \infty$ (так полагается в ТМ), а когда оценка сверху на $\|J^{-1}(x)\|$ неизвестна (так полагается в ТК), параметр r_M считается равным $+\infty$. В последнем случае $T_M := 1/Lr_0$ и, очевидно, $T_M \ge d_K$ при выполнении условия $P_K \le 1/2$. Таким образом, тогда $d_1 = d_K$ и условие 4) равносильно условию из ТК $S(x_0, d_K) \subset D$. Т.е. условия теоремы 1, когда неизвестна оценка $\|J^{-1}(x)\|$, совпадают с основными условиями ТК. Совпадают и заключения теорем об удаленности решения α от точки x_0.

Добавление информации в условии 3) о величине r_M (конечной) позволяет дать оценку удаленности, когда ТК ее не дает ($P_K > 1/2$), и улучшить оценку удаленности решения относительно таковой в ТК, когда $P_K \le 1/2$ (d_1 вместо d_K и $d_1 < d_K$).

Отметим, что эта оценка всегда лучше оценки согласно ТМ:

$$d_1 < d_M := L\bar{g}_0 \sum_0^{\infty} (P_M/2)^{2^k - 1},$$

которая имеет место еще и при дополнительном условии $P_M < 2$. Здесь $P_M := Lr_M^2\bar{g}_0$.

Интегрирование задачи Коши $\dot{x} = -J^{-1}(x)g(x)$, $x(0) = x_0$, Гаврурин в [10] называл непрерывным методом Ньютона. Для него решение уравнения (16) достигается при $t \longrightarrow +\infty$.

Доказательство теоремы 1. Сведем уравнение (16) к абстрактной задаче Коши (20) и исследуем ее с помощью полупроизводной отображения.

Известные теоремы о существовании решения задачи Коши для абстрактной функции скалярного аргумента вида (20) требуют хотя бы локальной липшицевости правой части уравнения в некоторой области $M \ni x_0$, т.е. существования для любого $x \in M$ окрестности, принадлежащей M, такой, что в ней F липшицева.

Оказывается, отображение F обладает локальной липшицевостью в открытом множестве $M := (S(x_0, d_1)\backslash\partial S(x_0, d_1))\backslash\{\alpha|g(\alpha) = 0\}$. Действительно, из условий теоремы следует, что в любом шаре из M операторы $J(x)$ и (см. леммы 2 и 1) $J^{-1}(x)$ липшицевы, ограничены, поэтому таково же в нем будет и $g(x)$.

Произвольно фиксируем $x \in M$ и рассмотрим шар $S(x, \nu)$ столь малого радиуса ν, что $\|J^{-1}(y)g(y)\| \ge \|J^{-1}(x)g(x)\|/2$ $\forall y \in S(x, \nu)$. Поскольку числитель и знаменатель дроби F липшицевы в $S(x, \nu)$, а знаменатель отделен от нуля, то, как нетрудно убедиться, и F липшицева в $S(x, \nu)$. Ввиду произвольности x это означает локальную липшицевость F в M.

Поэтому, согласно, например, [12, теорема 1.8.3], задача (20) имеет на некотором интервале $[0, T)$ непродолжимое вправо решение $x(t)$. Так как $\dot{x}(t)$ непрерывна и $\|F(x)\| = 1$ для любого $x \in M$, то для $x(t)$ справедлива оценка

$$\|x(\tau) - x(t)\| = \left\|\int_t^{\tau} \dot{x}(s)ds\right\| \le \left|\int_t^{\tau} \|\dot{x}(s)\|ds\right| \le |\tau - t| \quad \forall\tau, t \in [0, T). \tag{21}$$

Следовательно, если $T < \infty$ и $t_i \longrightarrow T$ при $i \longrightarrow \infty$, то последовательность $\{x(t_i)\}_0^{\infty}$ фундаментальна и существует предел $\lim\limits_{t \to T} x(t)$, который, согласно критерию продолжимости (см. [12, п. 1.8]), не может быть внутренней точкой множества M. Обозначим этот предел через $x(T)$.

Если $T < d_1$, то, положив $t = 0$ в (21) и устремив τ к T, получим $\|x(T) - x_0\| < d_1$ и $x(T) \in S(x_0, d_1)\backslash\partial S(x_0, d_1)$. Следовательно, попадание на границу M эквивалентно равенству $g(x(T)) = 0$, т.е. утверждения а) и б) теоремы справедливы.

Пусть теперь $T \ge d_1$. Рассмотрим поведение функции $\|g(x)\|$ на решении задачи (20). Пусть t, $t + \delta \in (0, T)$ и δ столь мало, что отрезок, соединяющий $x(t)$ и $x(t + \delta)$, лежит в M.

Обозначим $x(t + \delta) - x(t)$ через Δ и будем понимать под $\bar{o}(\delta, \cdot)$ и $\bar{o}_1(\delta, \cdot)$ отображения, которые при фиксированном втором аргументе удовлетворяют соотношениям

$$\lim_{\delta \to 0} \frac{\|\bar{o}(\delta, \cdot)\|}{\delta} = \lim_{\delta \to 0} \frac{\|\bar{o}_1(\delta, \cdot)\|}{\delta} = 0.$$

Тогда имеем

$$\|g(x(t + \delta))\| - \|g(x(t))\| = \|g(x(t)) + J(x(t))\Delta + \bar{o}(\delta, x(t))\| - \|g(x(t))\| =$$

$$= \left\|g(x(t)) - \frac{g(x(t))\delta}{\|J^{-1}(x(t))g(x(t))\|} + J(x(t))\bar{o}_1(\delta, t) + \bar{o}(\delta, x(t))\right\| - \|g(x(t))\| =$$

$$= \frac{-\|g(x(t))\|\delta}{\|J^{-1}(x(t))g(x(t))\|} + o_2(\delta, x(t)).$$

Поделив на δ и перейдя к пределу, получим

$$\|g(x(t))\|'_t = \frac{-\|g(x(t))\|}{\|J^{-1}(x(t))g(x(t))\|} \le \frac{-1}{\|J^{-1}(x(t))\|}, \tag{22}$$

следовательно, суперпозиция $\|g(x(t))\|$ монотонно убывает, если $x(t)$ – решение задачи Коши (20), что, конечно, верно и для случая $T < d_1$.

Оценим полупроизводную функции $1/\rho(t)$, где $\rho(t) := \|J^{-1}(x(t))\|$. Используя лемму 4, затем лемму 2 и неравенство (21), получаем

$$L_{1/\rho}(t) \le \rho^{-2}(t) L_{J^{-1}} L_x \le \rho^{-2}(t)\|J^{-1}(x(t))\| L \cdot 1 = L.$$

Далее, по лемме 3, имеем $|1/\rho(t) - 1/\rho(0)| \le Lt$, откуда следует неравенство $\rho(t) \le \dfrac{r_0}{1 - Lr_0 t}$. Кроме этой оценки, согласно условию 3) должно быть $\rho(t) \le r_M$ $\forall t \in [0, T)$. Единственный момент совпадения оценок легко вычисляется из уравнения $r_M = \dfrac{r_0}{1 - Lr_0 t}$ относительно t: $T_M := \dfrac{1 - r_0/r_M}{Lr_0}$. При $r_M = \infty$ полагаем $T_M := 1/(Lr_0)$. Для более поздних моментов лучше априорной оценке r_M для ρ не существует. Таким образом, $\rho(t) \le \varphi(t)$ $\forall t \in [0, T)$, где

$$\varphi(t) := \begin{cases} \dfrac{r_0}{1 - Lr_0 t}, & t \in [0, T_M) \\ r_M, & t \ge T_M \end{cases}. \tag{23}$$

Ясно, что при $r_M = \infty$ оценка φ конечна, лишь когда $T \le T_M$.

Усилим (22) с помощью (23): $(\|g\|)'_t \le -1/\varphi(t)$. Отсюда получаем

$$(\forall t \in [0, d_1)) \quad \|g(x(t))\| \le -\int_0^t \frac{d\tau}{\varphi(\tau)} + \bar{g}_0 =: \psi(t). \tag{24}$$

Если $r_M < \infty$ то подынтегральная функция ограничена снизу величиной $1/r_M > 0$, поэтому $\psi(t)$ монотонно стремится к $-\infty$, когда $t \longrightarrow +\infty$, и, следовательно. На $(0, +\infty)$ имеет единственный корень d. Нетрудно убедиться, что $d = d_1$.

Действительно, пусть $d \in (0, T_M]$. Тогда d должен удовлетворять системе

$$\psi(d) \equiv Ld^2/2 - d/r_0 + \bar{g}_0 = 0 \wedge (\forall d' \in (0, d))\psi(d') > 0. \tag{25}$$

Решение системы (25) существует, если и только если $P_L \le 1/2$. Тогда $d = d_K$, а $d_K = d_1$. (В этом случае знание оценки r_M не дает выигрыша в оценке удаленности сравнительно с ТК, которая не использует r_M. Это объясняется тем, что, в силу непрерывности, $\|J^{-1}(x)\|$ не достигает своего порогового значения в малой окрестности точки x_0). Отметим, что при $r_M = \infty$ и $P_L \le 1/2$ будет $T_M > d_K$, а ψ имеет корень $d = d_K = d_1$.

Пусть $d \ge T_M$. Тогда d определится из равенства

$$(Lt^2/2 - t/r_0)\Big|_0^{T_M} - \frac{d - T_M}{r} + \bar{g}_0 = 0.$$

Отсюда следует, что

$$d = T_M + r_M\left(\bar{g}_0 + \frac{LT_M^2}{2} - \frac{T_M}{r_0}\right) = r_M\bar{g}_0 + T_M\frac{r_0 - r_M}{r_0} + \frac{r_M L T_M^2}{2} =$$

$$= r_M\bar{g}_0 - r_L L T_M^2/2 = r_M\bar{g}_0 - (r_M/r_0 - 1)^2/(2r_M L) = d_1.$$

Так же, как для случая $T < d_1$, показывается, что существует $\lim\limits_{t \to d_1 - 0} x(t) =: \alpha$ и $\|x_0 - \alpha\| \le d_1$.

Переходя в неравенстве (24) к пределу при $t \longrightarrow d_1 - 0$ и учитывая непрерывность q в $S(x_0, d_1)$, получаем $\|g(\alpha)\| \le 0$. Поскольку $g(x(t)) \neq 0\ \forall t \in (0, d_1)$ (иначе было бы $T < d_1$, так как через корень уравнения (16) решение задачи Коши (20) непродолжимо), то элемент α – корень отображения g. Из тех же соображений следует, что $T = d_1$.

Как видно из предыдущего, если r_0 огрублено до r_M, т.е. информация о $J^{-1}(x_0)$ не используется, то $d_1 = r_M\bar{g}_0$ независимо от величины L. В этом случае. Как явствует из предыдущих рассуждений, условие 2) можно ослабить, потребовав лишь локальной липшицевости J в D'.

Замечания. 2. При $r_M = \infty$ теорема 1 совпадает с ТК в части, касающейся оценки удаленности корня отображения g от начальной точки.

3. Теорема 1 точна в том смысле, что для любого набора входных параметров можно указать функцию, на которой достигается оценка удаленности d_1. Действительно. Пусть имеются конкретные входные параметры – положительные числа $\bar{g}_0$, r_0, L, r_M ($r_M \ge r_0$). Согласно им вычисляем, как в теореме 1, параметр T_M и задаем скалярную функцию f следующим образом:

$$f(t) = \begin{cases} \bar{g}_0 - t/r_0 + Lt^2/2, & t \in [-\infty, T_M], \\ \dfrac{T_M - t}{r_M} + \bar{g}_0 - T_M/r_0 + LT_M^2/2, & t > T_M. \end{cases} \tag{26}$$

Очевидно, в любом сегменте из R^1 функция f липшицева с константой L и $|\dot{f}(t)|^{-1} \le r_M$. Несложно найти ближайший к $t_0 = 0$ корень f. Это как раз и есть d_1. Таким образом, для применения теоремы 1 к f необходимо, чтобы область задания D содержала сегмент $[-d_1, d_1] \equiv S(0, d_1)$. Поэтому функцию f из (26) будем именовать максимайзером теоремы 1 (удаленность ближайшего к x_0 корня).

Теорему 1, так же как и ТК, ТМ, ТГ, можно рассматривать как утверждения относительно не конкретного отображения, а класса отображений $\mathcal{K}_{\mathrm{N}}(p)$, где p – набор параметров, N – имя теоремы.

Уточним это высказывание. Пусть имеется обычная теорема-импликация N, на вход которой должны подаваться параметры p класса $\mathcal{K}_{\mathrm{N}}(p)$. Тогда ей можно придать формализованный вид

$$\Lambda(p) = 1 \Rightarrow (\forall g \in \mathcal{K}_{\mathrm{N}}(p))Y(p),$$

где Y – заключение теоремы, Λ – вещественнозначная функция от набора параметров p. Таким образом, N может трактоваться как утверждение относительно класса $\mathcal{K}_{\mathrm{N}}(p)$, но может быть и перефразирована относительно элемента этого класса:

$$((\exists p)g \in \mathcal{K}_{\mathrm{N}}(p) \wedge \Lambda(p) = 1) \Rightarrow Y(p).$$

Под применимостью теоремы N к элементу g обычно подразумевается не только $g \in \mathcal{K}_{\mathrm{N}}(p)$ для некоторого p, но и $\Lambda(p) = 1$; также и применимость теоремы к классу трактуется не только как соответствие его языку теоремы, но и как истинность $\Lambda(p) = 1$.

Будем именовать одно лишь такое языковое соответствие лингвистической применимостью теоремы к классу, а $(\exists p)g \in \mathcal{K}_{\mathrm{N}}(p)$ – лингвистической применимостью теоремы к g.

Определение 3. Назовем классом $C^{1,1}(D, L, x_0, \bar{g}_0, r_0)$, где L, $\bar{g}_0$, r_0 – положительные числа, а $x_0 \in D'$, класс отображений g, удовлетворяющих таким условиям:

1) g непрерывно в D;

2) линейный оператор J – производная (Фреше) отображения g – в области D' не вырожден и удовлетворяет условию Липшица с константой L;

3) $\bar{g}_0 \ge \|g(x_0)\|$, $r_0 \ge r(x_0)$.

Замечание 4. Из липшицевости J следует ее непрерывность. Поэтому производные Гато и Фреше совпадают.

ТК лингвистически применима к классам отображений $C^{1,1}(D, L, x_0, \bar{g}_0, r_0)$, а $\mathcal{K}_{\mathrm{K}} \equiv C^{1,1}$. Теорема 1 лингвистически применима к классам отображений

$$C^{1,1}(D, L, x_0, \bar{g}_0, r_0) \cap \{g | (\forall x \in D')\|J^{-1}(x)\| \le r_M\},$$

которые совпадают с $\mathcal{K}_{т1}(D, L, x_0, \bar{g}_0, r_0, r_M)$.

Для классов $\mathcal{C}(q) \neq \mathcal{K}_N(p)$ (q – набор параметров, вообще говоря, иной структуры, чем у набора p) можно предлагать отличные от теоремы N суждения, которые, будучи достаточно тонкими, могут оказаться лучше теоремы N, лингвистически применимой при некотором p к классу $\mathcal{K}_N(p) \supset \mathcal{C}(q)$.

И напротив, если заключение теоремы N относительно класса $\mathcal{K}_N(p)$ точно, то оно будет не хуже заключения любой другой теоремы M относительно любого класса $\mathcal{K}_M(q) \supset \mathcal{K}_N(p)$.

Определение 4. Назовем классом $\Pi(D, \sigma, x_0, \bar{g}_0, r_0)$, где σ, $\bar{g}_0$, r_0 – положительные числа и $x_0 \in D'$, класс отображений g, удовлетворяющих таким условиям:

1) g непрерывно в D;

2) линейный оператор J – производная (Фреше) отображения g – в области D' не вырожден и имеет полупроизводную L_J;

3) для всех x из D' выполняется $L_J(x)r(x) \le \sigma$, где $r(x) := \|J^{-1}(x)\|$;

4) $\bar{g}_0 \ge \|g(x_0)\|$, $r_0 \ge r(x_0)$.

Замечание 5. Лемма 1 при наличии полупроизводной гарантирует липшицевость J в любой окрестности внутренней точки области задания, следовательно, и непрерывность. Поэтому производные Гато и Фреше совпадают.

Объединение таких классов по какому либо параметру будем обозначать звездочкой на месте этого параметра либо просто опусканием параметра(ов), когда он(они) стоит(ят) в конце списка.

Теорема 2. *Пусть U, W суть B-пространства, $D \subset U$, $g : D \longrightarrow W$, g принадлежит* $\Pi(D, \sigma, x_0, \bar{g}_0, r_0)$ (*т.е.* $\mathcal{K}_{т.2} \equiv \Pi$) *и выполняются следующие условия:*

1) $P_\sigma := r_0\sigma\bar{g}_0 < 1$;

2) *шар $S(x_0, d_2)$ принадлежит области D, где $d_2 := -\sigma^{-1}\ln(1 - P_\sigma)$.*

Тогда в шаре $S(x_0, d_2)$ существует корень α уравнения (16), *который является правой предельной точкой фазовой траектории решения задачи Коши с автономным уравнением*

$$\dot{x} = -J^{-1}(x)g(x)\left\|J^{-1}(x)g(x)\right\|^{-1} =: F(x), \quad x(0) = x_0, \tag{27}$$

а именно: существует $T \in (0, d_1]$ такое, что решение этой задачи определено на $[0, T)$, *непродолжимо вправо и* $\lim\limits_{t \to T-0} x(t) = \alpha$. *Причем $\|g(x(t))\|$ монотонно убывает до* 0 *на* $[0, T)$.

Доказательство. Повторим доказательство теоремы 1 вплоть до введения функции $\rho(t) := r(x(t)) := \|J^{-1}(x(t))\|$, которое происходит при исследовании случая $T \ge d_1$. Здесь, соответственно, это будет случаем $T \ge d_2$.

Оценим полупроизводную функции $\rho(t)$. Используя лемму 4, затем лемму 2 и неравенство (21), получаем

$$L_\rho(t) \le \rho^2(t)L_J(x(t)) \le \sigma\rho(t). \tag{28}$$

Еще раз используя лемму 4, выясняем, что $L_{\ln\rho}(t) \le L_\rho(t)/\rho(t)$. Совместно с (28) это даст $L_{\ln\rho}(t) \le \sigma$. Отсюда, согласно лемме 3, вытекает неравенство

$$\ln\rho(t) - \ln\rho(0) \le \sigma t,$$

т.е. $\rho(t) \le r_0e^{\sigma t}$, $t \in [0, T)$.

Усилив неравенство в (22) подстановкой вместо $r(x(t))$ функции $r_0e^{\sigma t}$, получим $(\|g\|)'_t \le -e^{-\sigma t}/r_0$.

Интегрируем последнее неравенство и используем оценку $\|g(x(0))\| \le \bar{g}_0$:

$$\|g(x(t))\| \le \|g(x(0))\| + \frac{e^{-\sigma t} - 1}{r_0\sigma} \le \bar{g}_0 + \frac{e^{-\sigma t} - 1}{r_0\sigma} \quad \forall t \in [0, T). \tag{29}$$

Правая часть неравенства (29) стремится к нулю при стремлении t к определенному в условии теоремы d_2. Рассматривается случай $T \ge d_2$, т.е. скалярная функция $\|g(x(t))\|$ определена на $[0, d_2)$, поэтому найдется $\tau \in [0, d_2)$ такое, что $\|g(x(\tau))\| = 0$. Но тогда должно быть $T \le \tau$, следовательно,

$T = \tau = d_2$. Как уже говорилось ранее, существует $x(d_2) := \lim\limits_{t \to d_2} x(t)$ и $\|x(d_2) - x_0\| \le d_2$. А в силу непрерывности g верно $g(x(d_2)) = 0$, т.е. в шаре $S(x_0, d_2)$ существует корень отображения g.

Замечание 6. Условия 1) и 2) теоремы 2 точны в том смысле, что при их нарушении можно указать функцию из $\Pi(D, \sigma, x_0, \bar{g}_0, r_0)$, не имеющую корней в D.

Действительно, рассмотрим скалярную функцию вида

$$f(x) = \bar{g}_0 + (e^{-\sigma x} - 1)/(r_0\sigma), \tag{30}$$

где $x \in D \subset R^1$, $D' \ni x_0 = 0$, $\bar{g}_0$, $r_0 > 0$. Для нее $|f''(x)/f'(x)| = \sigma$ при всех x из D. Нетрудно проверить, что выполнены и остальные условия принадлежности f классу $\Pi(D, \sigma, x_0, \bar{g}_0, r_0)$. Поскольку f монотонно убывает и $f(0) = \bar{g}_0 > 0$, из существования корня у нее следует условие $0 > f(+\infty) = g_0 - 1/(r_0\sigma)$, или $P_\sigma < 1$. Это означает, что условие 1) теоремы 2 не может быть ослаблено.

Вместе с тем корень функции f, расширенной согласно формуле (30) на все R^1, есть $d_2 = -\sigma^{-1}\ln(1 - P_\sigma)$ и нарушение условия 2) теоремы 2 такое, что $d_2 \notin D$, означало бы, что в D нет корней.

Более того, удаленность корня функции f от x_0 равна d_1 – оценке, доставляемой теоремой 2, следовательно, d_2 – наименьшая оценка при имеющейся информации об отображении.

5. ПРИМЕНЕНИЕ ДИФФЕРЕНЦИРОВАНИЯ ПО ИТЕРАЦИИ В АНАЛИЗЕ СХОДИМОСТИ

При известной оценке d удаленности решения от начальной точки для метода Ньютона в $C^{1,1}$ были получены обоснования при условии $P_M := r_0Ld \le 2/3$ (см. [8]). Аналогичный результат в классе Π интересен появлением трансцендентной константы, причем она точна.

Теорема 3 (локальная сходимость метода Ньютона). *Пусть область D принадлежит* B-*пространству; пусть отображение g принадлежит* $\Pi(D, \sigma, x_0)$ (*см. определение* 4) *и некоторая оценка d удаленности решения* α *уравнения* (16) *от начальной точки x_0 удовлетворяет следующим условиям:*

1) $\sigma\hat{d} =: c_0 < \bar{c}$, *где $\bar{c}$ – решение уравнения*

$$C(c) := (e^c - 1 - c)/c = 1 \tag{31}$$

(*расчеты показывают, что $\bar{c}$ принадлежит интервалу* (1.25, 1.26));

2) $S(x_0, p\hat{d}) \subset D$, *где* $p = 1 - C(\sigma\hat{d})$.

Тогда верно следующее:

а) *метод Ньютона, начатый в точке x_0, корректно определен и сходится к решению* α, *все итерации лежат в шаре* $S(x_0, p\hat{d})$;

б) *скорость сходимости оценивается неравенством*

$$\|x_{k+1} - \alpha\| \le C(\sigma\|x_k - \alpha\|)\|x_k - \alpha\|,$$

причем при $s \le 1$ *справедливо* $C(\sigma s) \le C(\sigma)s$, *что соответствует квадратичной скорости сходимости*

$$\|x_{k+1} - \alpha\| \le C(\sigma)\|x_k - \alpha\|^2,$$

в) *решение* α *системы* $g(x) = 0$ *единственно в шаре* $S(x_0, \hat{d})$.

Доказательство. Корректность метода Ньютона на первом шаге (т.е. существование $J^{-1}(x_0)$) следует из того, что $g \in \Pi(D, \sigma, x_0)$, и из того, что $x_0 \in D$ (условие 2)).

Пусть метод Ньютона был корректен и не порождал итераций, удаленных от α более чем на $\hat{d}$ вплоть до k-го шага. Исследуем $(k + 1)$-й шаг с помощью метода дифференцирования по итерации.

Определим $d := \|x_k - \alpha\|$, $\bar{d} := (x_k - \alpha)/d$. Затем введем параметр t и зададим $x_k(t)$, как $\alpha + \bar{d}t$ (очевидно, $x_k(d) = x_k$).

Решение относительно x уравнения

$$J(x_k(t))[x - x_k(t)] + g(x_k(t)) = 0 \tag{32}$$

при $t = d$ есть x_{k+1}, при $t = 0$ есть α. Дифференцируя разностно (32) по t и заменяя g^{∇} на J, с точностью до погрешностей, возникших из-за этой замены, и бесконечно малых относительно разностного отклонения, получаем

$$(J)_t^{\nabla}(x - x_k) + J(x^{\nabla} - \bar{d}) + J\bar{d} \cong 0.$$

(Здесь опущены аргумент t у x_k и аргумент $x_k(t)$ у матриц J, J_t^{∇}). После преобразования получим $x^{\nabla} \cong J^{-1}J_t^{\nabla}(x_k - x)$. Оценивая полупроизводной норму разностной производной от J, с учетом $\|x_k^{\nabla}(t)\| = 1$ имеем

$$\|x^{\nabla}\| \lessapprox \|J^{-1}\|L_J(x)\|x_k - x\| \le \sigma\|x_k - x\|.$$

Произведем подстановку $x = \alpha + y$. Тогда

$$\|y^{\nabla}\| \lessapprox \sigma(t + \|y\|), \quad y(0) = 0. \tag{33}$$

Так как $\left|\|y\|^{\nabla}\right| \le \|y^{\nabla}\|$, то, согласно [5, теорема III.4.1], мажорирующая норму любого решения этой системы задача Коши имеет следующий вид:

$$\dot{z} = (t + z)\sigma, \quad z(0) = 0.$$

Решив ее, придем к оценке

$$\|x_{k+1} - \alpha\| = \|y(d)\| \le z(d) = \left.\frac{e^{\sigma t} - \sigma t - 1}{\sigma}\right|_{t = d} = d\frac{e^{\sigma d} - \sigma d - 1}{\sigma d} = \|x_k - \alpha\|C(\sigma d). \tag{34}$$

Нетрудно проверить разложением в ряд, что функция C монотонно возрастает:

$$C(c) = -\frac{1}{c} - 1 + \frac{1}{c} + 1 + \frac{c}{2!} + \frac{c^2}{3!} + \dots .$$

Первые 4 слагаемых взаимно уничтожаются, а последующие суть степени c с положительными коэффициентами.

По условию 1 и индуктивному предположению выполняется соотношение

$$\sigma d \le \sigma\hat{d} = c_0 < \bar{c}, \quad C(\bar{c}) = 1.$$

В совокупности с монотонностью функции C это дает соотношение

$$C(\sigma d) \le C(\sigma\hat{d}) = C(c_0) < C(\bar{c}) = 1,$$

с учетом (34) обеспечивающее выполнение неравенства $\|x_{k+1} - \alpha\| < \|x_k - \alpha\|$ и по индукции – справедливость (34) для всех k. Так как α находится не далее $\hat{d}$ от x_0, а итерации $x_1, x_2, \dots$ – не далее от α, чем $C(\sigma\hat{d})\|\hat{d}\|$, то все итерации лежат в шаре $S(x_0, \hat{d} + \hat{d}C(\sigma\hat{d}))$. Следовательно можно применить оценку (34) k раз: $\|x_k - \alpha\| \le C^k(c_0)\|x_0 - \alpha\|$, что означает сходимость итеративного процесса.

При наличии сходимости оценка (34) есть п. б) теоремы 3.

Единственность решения системы $g(x) = 0$ в шаре $S(x_0, \hat{d})$ следует из того, что при наличии там еще одного решения α' уже сходящаяся к α ньютоновская итеративная последовательность должна также сходиться и к α', что возможно только при $\alpha' = \alpha$.

Если $d < 1$, то из разложения C в ряд следует $C(\sigma d) < C(\sigma)d$. Поэтому если верны неравенства $\|x_k - \alpha\|C(\sigma) < 1$ и $\|x_k - \alpha\| < 1$, то справедлива следующая оценка скорости сходимости: $\|x_{k+1} - \alpha\| < C(\sigma)\|x_k - \alpha\|^2$.

Замечания 7. Теорема 3 не использует оценки $\bar{g}_0 \geq \|g(x_0)\|$ и $r_0 \geq \|J^{-1}(x_0)\|$.

8. Условия 1), 2) теоремы 3 являются также и необходимыми в том смысле, что при их нарушении можно найти функцию и начальную точку, для которых метод Ньютона не сходится, а для начальных точек, обеспечивающих сходимость, оценка скорости сходимости точна.

Простым следствием теорем 2 и 3 является

Теорема 4. *Пусть* U, W *суть* B-*пространства*, $D \subset U$, $g : D \longrightarrow W$, g *принадлежит* $\Pi(D, \sigma, x_0, \bar{g}_0, r_0)$ *и выполнены такие условия:*

1) $P_\sigma \equiv r_0\sigma\bar{g}_0 < \dfrac{2\bar{c}}{1+2\bar{c}}$, *где* $\bar{c}$ *– решение уравнения* (31);

2) $S(x_0, pd_0) \subset D$, *где* $p := 1 + C(\sigma d_0)$, $d_0 := -\dfrac{\ln(1-P_\sigma)}{\sigma}$.

Тогда верно следующее:

а) *существует решение* α *системы* $g(x) = 0$, *единственное в шаре* $S(x_0, d_0)$;

б) *метод Ньютона, начатый в точке* x_0, *корректно определен и сходится к конечному решению* α;

в) *скорость сходимости оценивается неравенством*

$$\|x_{k+1} - \alpha\| \leq C(\sigma\|x_k - \alpha\|)\|x_k - \alpha\|.$$

Для некоторых отображений, принимаемых за элементы класса $\Pi(D, \sigma, x_0, \bar{g}_0, r_0)$, эта теорема гарантирует сходимость метода Ньютона, в то время как зачисление этих отображений в $C^{1,1}(D, L, x_0, \bar{g}_0, r_0)$, сколь хорошо ни вычислялась бы константа Липшица L, дает параметр $P_L > 1/2$, и теорема Канторовича не гарантирует сходимости метода Ньютона.

6. ЗАКЛЮЧЕНИЕ

Разумеется, предложенный исследовательский инструмент (исчисление полупроизводных и дифференцирование по итерации) не является волшебной палочкой, мановением которой можно без труда получать результаты. По крайне мере, длиной доказательство теоремы 1 не уступает упомянутой теореме Канторовича. Однако новый подход после затрат нужного объема труда может приводить и к существенно новым результатам (см. теорему 2 и доказательство сходимости метода квадратичной оптимизации в [1]) и улучшать известные.

СПИСОК ЛИТЕРАТУРЫ

1. *Михеев С.Е.* Нелинейные методы в оптимизации. СПб: Изд-во СПбГУ, 2001.
2. *Пономарев В.М., Горичев Ю.В., Городецкий В.И.* Последовательная оптимизация нелинейного закона управления // Нелинейная оптимизация систем автоматич. управления. М.: Машиностр., 1970. С. 15–19.
3. *Поцелуев А.В., Майборода Л.А., Пономарев В.М. и др.* Статистический анализ и оптимизация следящих систем. М.: Машиностр., 1977.
4. *Шилов Г.Е.* Математический анализ: Специальный курс. М.: Физматгиз, 1961.
5. *Хартман Ф.* Обыкновенные дифференциальные уравнения. М.: Мир, 1970.
6. *Канторович Л.В., Акилов Г.П.* Функциональный анализ. М.: Наука, 1977.
7. *Титчмарш Е.* Теория функций. М.: Мир, 1982.
8. *Мысовских И.П.* О сходимости метода Л.В. Канторовича решения функциональных уравнений и его применениях // Докл. АН СССР. 1950. Т. 70. № 4. С. 565–568.
9. *Красносельский М.А. и др.* Приближенное решение операторных уравнений. М.: Наука, 1969.
10. *Гаврурин М.К.* Нелинейные функциональные уравнения и непрерывные аналоги итерационных методов // Изв. вузов, 1956. № 5(6). С. 18–31.
11. *Ортега Дж., Рейнболдт В.* Итерационные методы решения нелинейных систем уравнений со многими неизвестными. М.: Мир, 1975.
12. *Картан А.* Дифференциальное исчисление, дифференциальные формы. М.: Мир, 1971.