Известия вузов. Математика *http://www.ksu.ru/journals/izv_vuz/*
2010, №2, с. 66–76 *Гос. номер статьи по НТЦ "Информрегистр" 0421000123\0014*

С.Я. СЕРОВАЙСКИЙ

ДИФФЕРЕНЦИРОВАНИЕ ОПЕРАТОРОВ И УСЛОВИЯ ЭКСТРЕМУМА С КАТЕГОРНОЙ ИНТЕРПРЕТАЦИЕЙ

Аннотация. Общая теория экстремума в значительной степени опирается на свойства операторных производных. В качестве примера рассматривается система, описываемая нелинейным уравнением эллиптического типа. При больших значениях показателя нелинейности и размерности области решение краевой задачи оказывается недифференцируемым в смысле Гато по управлению. Это не позволяет непосредственно найти производную критерия оптимальности и получить условия оптимальности стандартным методом. Однако указанная зависимость является расширенно дифференцируемой, что позволяет получить необходимые условия экстремума без ограничений на параметры задачи. В заключительной части работы показывается, что условия экстремума, связанные с классическим и расширенным дифференцированием, могут быть интерпретированы с позиции теории категорий.

Ключевые слова: условия экстремума, производные операторов, нелинейное эллиптическое уравнение, категории.

УДК: 571.988

Abstract. The general extremum theory essentially uses properties of operator derivatives. As an example we consider a system described by a nonlinear elliptic equation. In this system with large values of the nonlinearity parameter and large dimension of the domain the control-state mapping is not Gateaux differentiable. For this reason one cannot immediately differentiate the optimality criterion and establish the necessary optimality conditions by classical methods. However the mentioned mapping is extendedly differentiable. This allows one to obtain optimality conditions without constraints on parameters of the system. Concluding the paper, we interpret the optimality conditions with classical and extended derivatives within the theory of categories.

Keywords: optimality conditions, operator derivatives, nonlinear elliptic equation, categories.

1. Классическое дифференцирование

Результаты общей теории экстремума и общие схемы получения необходимых условий оптимальности в значительной степени опираются на свойства операторных производных (например, [1]–[5]). Пусть, в частности, задано уравнение состояния управляемой системы

$$Ay = v, \tag{1}$$

где $A : Y \to V$ есть непрерывно дифференцируемый оператор состояния, Y, V — банаховы пространства, v — управление, y — функция состояния системы. Предполагается, что для любого $v \in V$ уравнение (1) имеет единственное решение $y = y(v)$ из пространства Y.

Поступили: 22.10.2007, окончательный вариант 27.10.2008

Определен критерий оптимальности $I : V \to \mathbb{R}$ в соответствии с равенством $I(v) = J(v) + K[y(v)]$, где $J : V \to \mathbb{R}$, $K : Y \to \mathbb{R}$ — непрерывно дифференцируемые функционалы. Ставится следующая задача оптимального управления:

Задача P. *Найти точку $v \in V$, минимизирующую функционал I в пространстве V.*

Нетрудно устанавливается

Теорема 1. *Если производная оператора A в точке $y = y(v)$ обратима, где v есть решение задачи P, то справедливо равенство*

$$J'(v) + p = 0, \tag{2}$$

p — решение уравнения

$$[A'(y)]^* p = K'(y). \tag{3}$$

Для доказательства достаточно найти производную минимизируемого функционала с помощью теоремы об обратной функции ([6], с. 40) и приравнять ее нулю.

Для нахождения неизвестных функций v, y, p строится система, включающая в себя уравнение (1) на оптимальном управлении, условие стационарности (2) и сопряженное уравнение (3). Мы ограничились рассмотрением задачи на безусловный экстремум исключительно из соображений простоты. Если функционал минимизируется на выпуклом подмножестве пространства управлений, то условие стационарности заменяется вариационным неравенством (например, [7], с. 18, теорема 1.3). При наличии более сложных ограничений можно воспользоваться соответствующими результатами общей теории экстремума [1]–[5].

Рассмотрим пример. В открытой ограниченной области Ω из $\mathbb{R}^n$ дано уравнение

$$-\Delta y + |y|^\rho y = v \tag{4}$$

с однородными граничными условиями (задача Дирихле), где $\rho > 0$. Определим пространства $Y = H_0^1(\Omega) \cap L_q(\Omega)$, $V = Y^* = H^{-1}(\Omega) + L_{q'}(\Omega)$, где $q = \rho + 2$, $1/q + 1/q' = 1$. С помощью теории монотонных операторов (например, [8], с. 182–185) можно установить, что для любого $v \in V$ уравнение (1) имеет единственное решение $y = y(v)$ из пространства Y. Задается функционал

$$I(v) = \frac{\chi}{2}\|v\|^2_{L_2(\Omega)} + \frac{1}{2}\|y(v) - y_d\|^2_{H_0^1(\Omega)},$$

где $\chi > 0$, $y_d \in H_0^1(\Omega)$ — известная функция. Ставится оптимизационная

Задача Q. *Найти управление, минимизирующее функционал I в пространстве V.*

С помощью стандартной методики (например, [7], с. 13, теорема 1.1) нетрудно доказать (в частности, [9], с. 46, теорема 6.1), что рассматриваемая задача разрешима. Определим оператор $A : Y \to V$, выбирая в качестве Ay левую часть равенства (4). Тогда краевая задача сводится к операторному уравнению (1). Выбираем в качестве J и K соответственно первое и второе слагаемые в правой части последнего равенства. Тогда задача Q сводится к задаче P. Учитывая дифференцируемость оператора A и функционалов J и K, заключаем, что для применения теоремы 1 в данном случае достаточно установить обратимость производной $A'(y)$.

Лемма 1. *При $n = 2$ или $\rho \le 4n/(n-2)$ для $n \ge 3$ производная $A'(y)$ обратима.*

Доказательство. Требуется установить, что для любого $h \in V$ однородная задача Дирихле для линеаризованного уравнения

$$-\Delta z + (\rho + 1)|y|^\rho z = h \tag{5}$$

однозначно разрешима в пространстве Y. Умножая равенство (5) на функцию z и интегрируя результат формально по области Ω с учетом формулы Грина, имеем

$$\int_\Omega \sum_{i=1}^n \left(\frac{\partial z}{\partial x_i}\right)^2 dx + (\rho+1)\int_\Omega |y|^\rho z^2 dx = \int_\Omega hz\,dx. \tag{6}$$

В условиях леммы в силу теоремы Соболева справедливо вложение $H_0^1(\Omega) \subset L_q(\Omega)$, откуда вытекают равенства $Y = H_0^1(\Omega)$, $V = H^{-1}(\Omega)$. Тогда из соотношения (6) следует, что для любого $h \in V$ уравнение (5) допускает априорную оценку решения в пространстве Y. Для завершения доказательства достаточно воспользоваться классической теорией линейных эллиптических уравнений (например, [7], с. 53, теорема 1.1). □

Пользуясь теоремой 1 с учетом леммы 1, получаем

Следствие. При выполнении условий леммы 1 решение задачи Q удовлетворяет равенству

$$\chi v + p = 0, \tag{7}$$

где p есть решение однородной задачи Дирихле для уравнения

$$-\Delta p + (\rho+1)|y|^\rho p = \Delta y_d - \Delta y. \tag{8}$$

Итак, для нахождения оптимального управления в задаче Q требуется решить систему условий оптимальности (4), (7), (8). Для ее практического решения выражаем управление из равенства (7) и подставляем его в уравнение (4). В результате получается однородная задача Дирихле для системы, включающей в себя уравнение (8), а также $-\Delta y + |y|^\rho y = -\chi^{-1}p$. Определив ее решение, можно найти искомое управление по формуле $v = -\chi^{-1}p$.

Приведенные выше результаты не выходят за рамки общих схем построения необходимых условий оптимальности (например, [1]–[5]) и хорошо известных методов решения задач оптимального управления для систем, описываемых нелинейными уравнениями эллиптического типа (например, [9]–[15]). Так, следствие является частным случаем теоремы 7.1 из [9], с. 127. Однако возникает вопрос: можно ли получить аналогичные результаты при нарушении условий леммы 1?

Для достаточно больших значений показателя нелинейности ρ и размерности области n в отсутствии вложения $H_0^1(\Omega) \subset L_q(\Omega)$ соотношение (7) уже не обеспечивает априорную оценку решения уравнения (6) в пространстве Y для любого $v \in V$. Это обстоятельство не позволяет воспользоваться теоремой об обратной функции для получения желаемых результатов. Более того, справедлива

Лемма 2 ([16]). *В отсутствии вложения $H_0^1(\Omega) \subset L_q(\Omega)$ оператор A^{-1} не дифференцируем по Гато.*

Доказательство. Для любой непрерывной функции y из Y левая часть равенства (6) для всех $z \in Y$ принадлежит пространству $V_* = H^{-1}(\Omega)$, являющемуся собственным подмножеством V. Если оператор A^{-1} дифференцируем по Гато в точке $v = Ay$, то для любого $h \in V$ при $\sigma \to 0$ имеем сходимость $[y(v+\sigma h) - y]/\sigma \to y'(v)h$ в Y. Вычитая из равенства (6) на управлении $v + \sigma h$ то же равенство на управлении v, после деления на σ и перехода к пределу получаем

$$-\Delta y'(v)h + (\rho+1)|y|^\rho y'(v)h = h.$$

Сравнивая это выражение с (6), заключаем, что для любого $h \in V$ линеаризованное уравнение имеет решение $z = y'(v)h$ из класса Y. Однако ранее было установлено, что это невозможно, по крайней мере, при $h \in V \setminus V_*$. □

Отметим, что обратимость производной оператора, выражающего бесконечномерные ограничения типа равенства, или связанное с этим требование замкнутости образа этой производной в соответствующем пространстве восходит к стандартным линеаризационным теоремам (теорема об обратной функции, теорема о неявной функции [17], с. 660, теорема Люстерника об аппроксимации гладкого многообразия касательной плоскостью [5]), лежащим в основе известных результатов общей теории экстремума (например, [1]–[5]). Вследствие этого указанные методы оказываются неприменимыми для анализа задачи Q в отсутствии ограничений на показатель нелинейности и размерность области. Этим обстоятельством объясняется тот факт, что подобные ограничения неизменно присутствуют в работах, посвященных задачам оптимального управления системами, описываемыми нелинейными уравнениями эллиптического типа (например, [10], с. 127; [18], с. 260). Аналогичные требования используются и в случае нестационарных уравнений (например, [10], с. 137; [18], с. 24; [19], с. 184). Указанные трудности можно обойти, если искать решение краевой задачи в пространстве L_∞ [15] или на некотором гильбертовом пространстве [16]. Однако для справедливости указанных свойств потребуются дополнительные ограничения на параметры задачи. Исследование задачи без дополнительных ограничений в наиболее естественных функциональных пространствах (соответствующих простейшим априорным оценкам) достигается с помощью иного математического аппарата.

2. Расширенное дифференцирование

Может создаться впечатление, что трудности, вызванные отсутствием дифференцируемости решения краевой задачи по свободному члену, удастся преодолеть с помощью методов негладкого анализа (субдифференцирование, производная Кларка, например, [20]) или гладкой аппроксимации системы [16], [19]. В данном случае указанное отрицательное свойство обусловлено не наличием в постановке задачи негладких членов типа модуля, максимума двух функций, вследствие чего способы решения негладких экстремальных задач также оказываются неэффективными. Невозможность применения теоремы 1 в данном случае объясняется отсутствием желаемого набора априорных оценок соответствующего линеаризованного уравнения. Однако для уравнений (5), (6) положительные результаты могут быть установлены и в общем случае.

Определим множества $Y(v) = H_0^1(\Omega) \cap \{z \mid |y(v)|^{\rho/2} z \in L_2(\Omega)\}$, $V(v) = Y(v)^*$. Пространство $Y(v)$ является гильбертовым со скалярным произведением

$$(y, z) = \int_\Omega \nabla y \nabla z \, dx + (\rho + 1) \int_\Omega |y(v)|^\rho yz \, dx.$$

Пространство $V(v)$ гильбертово с нормой сопряженного пространства. Тогда из равенства (7) выводится оценка $\|z\|_{Y(v)} \le \|h\|_{V(v)}$, откуда следует, что для любого h из пространства $V(v)$, более узкого, чем V, уравнение (6) однозначно разрешимо в классе $Y(v)$, более широком по сравнению с Y. Это свойство можно трактовать как ослабленную форму обратимости производной $A'(v)$. Можно предположить, что вследствие этого оператор A^{-1} будет обладать некоторыми положительными свойствами, хотя и более слабыми, чем классическая дифференцируемость. Рассматривается оператор $L : V \to Y$, где V, Y — банаховы пространства.

Определение. Оператор L назовем *расширенно дифференцируемым* (более точно, $(V_0, Y_0; V_*, Y_*)$-расширенно дифференцируемым) *по Гато* в точке $v \in V$, если существуют такие линейные нормированные пространства V_0, Y_0, V_*, Y_*, удовлетворяющие непрерывным вложениям $V_* \subset V_0 \subset V$, $Y \subset Y_0 \subset Y_*$, и линейный непрерывный оператор

$L'(v) : V_0 \to Y_0$, что при $\sigma \to 0$ имеет место сходимость $[L(v + \sigma h) - Lv]/\sigma \to L'(v)h$ в Y_* для всех $h \in V_*$. Оператор называется *расширенно дифференцируемым по Фреше* в точке v, если справедливо равенство $L(v + h) = Lv + L'(v)h + \eta(h)$, где $\|\eta(h)\|_{Y_*} = o(\|h\|_{V_*})$. Оператор *расширенно непрерывно дифференцируем* в этой точке, если он обладает $(V_u, Y_u; V_*, Y_*)$-расширенной производной Гато $L'(u)h$ в каждой точке u из некоторой V_*-окрестности v, причем при $(u - v) \to 0$ в V_* имеет место сходимость $L'(u)h \to L'(v)h$ в Y_* для всех $h \in V_*$.

Очевидно, $(V, Y; V, Y)$-расширенные производные совпадают с соответствующими классическими аналогами. Соотношение между различными типами расширенных производных (с одним и тем же сочетанием пространств) остается таким же, как и для их классических аналогов (для обоснования соответствующих результатов достаточно воспользоваться доказательствами классических аналогов с естественной коррекцией функциональных пространств). Понятие расширенной производной оператора (без непрерывной дифференцируемости) вводится в [17], [21]–[23]. Можно показать ([17], теорема 4), что рассматриваемый обратный оператор в каждой точке $v \in V$ является $(V(v), Y(v); V_*, Y_*)$-расширенно дифференцируемым, где $Y_* = H_0^1(\Omega)$, а все остальные пространства были определены выше.

Лемма 3. *Оператор A^{-1} для уравнения* (4) *в каждой точке $v \in V$ имеет $(V(v), Y(v); V_*, Y_*)$-расширенную производную Гато D, характеризуемую соотношением*

$$\int_\Omega \mu Dh\, dx = \int_\Omega h p_\mu(v) dx \quad \forall h \in V(v), \quad \mu \in Y(v)^*,$$

где $p_\mu(v)$ есть решение однородной задачи Дирихле для уравнения

$$-\Delta p_\mu(v) + (\rho + 1)|y(v)|^\rho p_\mu(v) = \mu.$$

Согласно теории классического дифференцирования (леммы 1 и 2) с ростом показателя нелинейности и размерности области свойства уравнения (4) меняются скачком: дифференцируемость решения краевой задачи по свободному члену уравнения была и исчезла. Согласно лемме 3 указанная зависимость остается расширенно дифференцируемой для всех значений указанных параметров. Просто по мере их роста пространства, входящие в определение расширенной производной, все сильнее отличаются от пространств, в которых действует исходный оператор. При этом в случае выполнения условий леммы 1 расширенная производная обратного оператора совпадает с его классической производной. Тем самым теория расширенной дифференцируемости позволяет установить более точные и естественные свойства зависимости решения краевой задачи от свободного члена. Указанными результатами можно воспользоваться для анализа задачи Q без ограничений на параметры задачи.

Теорема 2. *Утверждения следствия* (*на с.* 68) *остаются в силе и без условий леммы* 1.

Действительно, если v есть решение задачи, то справедливо неравенство $I(v+\sigma h)-I(v)\geq 0$ для всех $h \in V_*$ и $\sigma > 0$. Разделив это неравенство на σ и переходя к пределу при $\sigma \to 0$ с учетом леммы 3, после элементарных преобразований получаем (7) и (8). Таким образом, использование аппарата расширенной дифференцируемости позволяет провести анализ экстремальных задач в условиях, когда применение известных методов (например, [1]–[5], [8], [10]–[20]) не представляется возможным. Другие примеры применения указанной методики в задачах оптимального управления нелинейными бесконечномерными системами приводятся в [23].

3. Приложение. Категорная интерпретация

Как известно, теория категорий находит широкое применение практически во всех разделах математики [24], [25]. Среди ее нетрадиционных приложений можно отметить и теорию управляемых систем [26], [27]. Определим категории, связанные с процедурой дифференцирования операторов, которые дают интересную интерпретацию необходимых условий экстремума в нелинейных бесконечномерных системах.

Рассматривается совокупность всевозможных пар (X, x), где X есть банахово пространство, а x — некоторая фиксированная точка пространства X, т. е. множество банаховых пространств с выделенными точками. Для двух подобных пар (X, x) и (Y, y) задается такой оператор $L : X \to Y$, определенный в окрестности точки x и дифференцируемый по Фреше в этой точке, что справедливо равенство $Lx = y$. Пусть пары (Y, y) и (Z, z) связаны оператором M с аналогичным набором свойств. Тогда композиция ML действует из пространства X в Z, причем справедливо равенство $MLx = z$. В соответствии с теоремой о сложной функции оператор ML дифференцируем по Фреше в точке x, причем справедливо равенство $(ML)'(x) = M'(y)L'(x)$. Тем самым каждый такой оператор L однозначно задает некоторый морфизм L_x с началом (X, x) и концом (Y, y), причем композиции последовательных морфизмов $L_x \circ M_y$ соответствует оператор ML. Естественно, два оператора, совпадающие в окрестности данной точки, определяют один и тот же морфизм. Определяемую в результате категорию обозначим через Γ. Отметим, что на любом множестве параллельных морфизмов определена естественная операция сложения.

Рассмотрим некоторый морфизм L_x категории Γ, обладающий тем свойством, что соответствующий ему оператор L непрерывно дифференцируем в некоторой окрестности O точки x, причем производная $L'(x)$ обратима. Тогда согласно теореме об обратной функции ([7], с. 40) существует обратный оператор L^{-1}, определенный и непрерывно дифференцируемый в некоторой окрестности O' точки $y = Lx$, причем справедливо равенство $(L^{-1})'(\xi) = \{L'[L^{-1}(\xi)]\}^{-1}$ для всех $\xi \in O'$. Таким образом, обратный оператор определяет морфизм $(L^{-1})_y$ категории Γ, оказывающийся обратным к L_x, т. е. $(L^{-1})_y = (L_x)^{-1}$. Тем самым множество банаховых пространств с выделенными точками и морфизмами указанного класса образуют подкатегорию категории Γ, в которой все морфизмы являются изоморфизмами.

Определим преобразование D из Γ в категорию Σ банаховых пространств с линейными непрерывными операторами. Для любого объекта (X, x) и морфизма L_x с началом (X, x) и концом (Y, y) категории Γ определяем $D(X, x) = X$ и $D(L_x) = L'(x)$. Пусть M_y есть морфизм категории Γ с началом (Y, y). В соответствии с теоремой о сложной функции для композиции морфизмов справедливо равенство $D(L_x \circ M_y) = (ML)'(x) = D(L_x) \circ D(M_y)$. Тем самым D оказывается функтором из Γ в категорию Σ. Он допускает сужение на указанную подкатегорию, для которого сохраняем исходное обозначение. При этом, если производная $L'(x)$ обратима, то из равенства $(L^{-1})'(y) = [L'(x)]^{-1}$ следует

$$D[(L^{-1})_y] = [D(L_x)]^{-1}.$$

Покажем, что введенные понятия дают некоторую интерпретацию теоремы 1. Необходимым условием безусловного экстремума в точке v гладкого функционала I является условие стационарности $I'(v) = 0$. Данному функционалу сопоставляется морфизм I_v. Тогда условие стационарности принимает вид $D(I_v) = 0$. Пусть v есть решение задачи P, а y — соответствующее ему решение уравнения (1). Если производная $A'(y)$ обратима, то оператор A^{-1} непрерывно дифференцируем в точке v. Рассмотрим морфизмы A_y, J_v, K_y и $(A^{-1})_v$ категории Γ. Справедливо равенство $I_v = J_v + (A^{-1})_v \circ K_y$. Учитывая линейность производной, находим

$$D(I_v) = D(J_v) + D[(A^{-1})_v \circ K_y] = D(J_v) + D[(A^{-1})_v] \circ D(K_y).$$

Пользуясь представлением функтора от обратного морфизма, приходим к равенству

$$D(I_v) = D(J_v) + [D(A_y)]^{-1} \circ D(K_y).$$

Обозначим через H основной кофунктор из Σ в категорию множеств, определяемый объектом $\mathbb{R}$. Для любого объекта X из Γ под $H(X)$ понимается множество всех морфизмов из X в $\mathbb{R}$, а для любого морфизма L с началом X и концом Y в качестве $H(L)$ выбирается такое отображение из $H(Y)$ в $H(X)$, что для любого $a \in H(Y)$ справедливо равенство $H(L)a = L \circ a$. Тем самым банахову пространству X и линейному непрерывному оператору L функтором H сопоставляются соответствующие сопряженное пространство X^* и сопряженный оператор L^*. В результате предшествующее равенство принимает вид

$$D(I_v) = D(J_v) + H\{[D(A_y)]^{-1}\}D(K_y) = D(J_v) + \{H[D(A_y)]\}^{-1}D(K_y).$$

Таким образом, справедлива

Теорема 3. *Если производная оператора A в точке $y = y(v)$ обратима, где v есть решение задачи P, то справедливо равенство*

$$D(J_v) + \{H[D(A_y)]\}^{-1}D(K_y) = 0. \tag{9}$$

Из теоремы 3 несложно вывести условие оптимальности в стандартной форме. Действительно, соотношение (9) можно записать в виде $J'(v) + \{[A'(y)]^*\}^{-1}K'(y) = 0$. Если функция p является решением уравнения (3), то последнее равенство принимает вид (2). Таким образом, теорема 1 оказывается простым следствием теоремы 3.

Возникает вопрос: можно ли применить аппарат теории категорий для расширенно дифференцируемых операторов и установить соответствующий аналог теоремы 1? Для этого потребуются аналоги теорем об обратной и сложной функций для расширенных производных. Пусть заданы банаховы пространства Y, V, дифференцируемый по Гато обратимый оператор $A : Y \to V$ и точки $y \in Y$, $v = Ay$. Дано банахово пространство V_*, являющееся подпространством V, с некоторой окрестностью нуля O_*. Обозначив обратный оператор A^{-1} через L, рассмотрим равенство $ALv_\sigma = v + \sigma h$ для произвольной функции $h \in V_*$ и числа $\sigma > 0$, столь малого, чтобы имело место включение $v_\sigma \in O$, где $v_\sigma = v + \sigma h$. Отсюда следует равенство $ALv_\sigma - ALv = \sigma h$. С помощью теоремы о среднем [28] получаем соотношение

$$\langle \lambda, G(v_\sigma)(Lv_\sigma - Lv)/\sigma \rangle = \langle \lambda, h \rangle \quad \forall \lambda \in (V_*)^*, \quad h \in V_*,$$

где линейный непрерывный оператор $G(u) : Y \to V$ для любого $u \in O$ определим по формуле

$$G(u)z = \left\{ \int_0^1 A'[Lv + \theta(Lu - Lv)]d\theta \right\} z \quad \forall z \in Y.$$

Для любого $z \in O$ рассматриваются такие банаховы пространства $V(v)$ и $Y(v)$, что Y, V_* и $V(v)$ непрерывно и плотно вложены в $Y(v)$, $V(v)$ и V соответственно. Пусть выполнено

Условие 1. *Для любого $u \in O$ существует такое непрерывное продолжение $\overline{G}(u)$ оператора $G(u)$ на $Y(u)$, что множество его значений является подмножеством $V(u)$.*

Тогда из предшествующего соотношения получаем

$$\langle [\overline{G}(u)]^* \lambda, (Lv_\sigma - Lv)/\sigma \rangle = \langle \lambda, h \rangle \quad \forall \lambda \in [V(v_\sigma)]^*, \quad h \in V_*. \tag{10}$$

Определим линейное уравнение
$$[\overline{G}(u)]^* p_\mu(u) = \mu. \tag{11}$$
При $u = v$ оно принимает вид $[\overline{A}'(y)]^* p_\mu(v) = \mu$, где $[\overline{A}'(y)]^* = [\overline{G}(v)]^*$ — продолжение оператора $[A'(y)]^* = [G(v)]^*$ на множество $Y(v)$.

Рассмотрим некоторое банахово пространство Y_*, в которое непрерывно и плотно вложено пространство $Y(u)$ для любого $u \in O$. Пусть справедливо

Условие 2. *Для любых значений $u \in O$, $\mu \in [Y(u)]^*$ уравнение* (11) *имеет единственное решение $p_\mu(u) \in [V(u)]^*$, причем для любых значений $h \in V_*$ при $\sigma \to 0$ имеет место сходимость $p_\mu(v_\sigma) \to p_\mu(v)$ $*$-слабо в $(V_*)^*$, равномерно по μ из замкнутого единичного шара M с центром в нуле пространства $(Y_*)^*$.*

Определим в соотношении (10) $\lambda = p_\mu(v_\sigma)$ при достаточно малых σ. Получаем
$$\langle \mu, [L(v + \sigma h) - Lv]/\sigma \rangle = \langle p_\mu(v_\sigma), h \rangle \ \ \forall \mu \in [Y(v_\sigma)]^*, \quad h \in V_*.$$

Справедлива теорема о расширенной дифференцируемости обратной функции.

Теорема 4. *При выполнении условий* 1 *и* 2 *оператор A^{-1} в точке v имеет $(V(v), Y(v); V_*, Y_*)$-расширенную производную Гато D, определяемую равенством*
$$\langle \mu, Dh \rangle = \langle p_\mu(v), h \rangle \ \ \forall \mu \in [Y(v)]^*, \quad h \in V(v).$$
Она оказывается $(V(v), Y(v); W, Y_*)$ -расширенной производной Фреше, если W есть банахово пространство, компактно и плотно вложенное в V_*.

Доказательство. Справедливо соотношение
$$\langle \mu, [L(v + \sigma h) - Lv]/\sigma - Dh \rangle = \langle p_\mu(v_\sigma) - p_\mu(v), h \rangle \ \ \forall \mu \in M, \quad h \in V_*,$$
откуда следует равенство
$$\|[L(v + \sigma h) - Lv]/\sigma - Dh\|_{Y_*} = \sup_{\mu \in M} |\langle p_\mu(v_\sigma) - p_\mu(v), h \rangle|.$$
При $\sigma \to 0$ в силу условия 2 для любого $h \in V_*$ имеет место сходимость $p_\mu(v_\sigma) \to p_\mu(v)$ $*$-слабо в $(V_*)^*$ равномерно по $\mu \in M$. Тогда в результате перехода к пределу в последнем равенстве получаем, что оператор D действительно оказывается расширенной производной Гато оператора $L = A^{-1}$.

По аналогии с предшествующим равенством установим соотношение
$$\|[L(v + h) - Lv] - Dh\|_{Y_*} = \sup_{\mu \in M} |\langle p_\mu(v + h) - p_\mu(v), h \rangle| \le \sup_{\mu \in M} \|p_\mu(v + h) - p_\mu(v)\|_{W^*} \|h\|_W.$$
Ввиду компактности вложения пространства $(V_*)^*$ в W^* при $h \to 0$ в W заключаем, что $p_\mu(v + h) \to p_\mu(v)$ сильно в W^* равномерно по $\mu \in M$. Тогда из последнего неравенства следует, что D есть соответствующая расширенная производная Фреше. □

Другие варианты теоремы о расширенной дифференцируемости обратного оператора приводятся в [18], [21]–[23]. В [23] доказывается соответствующий аналог теоремы о неявной функции. По аналогии с леммой 2 из [18] показывается, что при выполнении условий теоремы об обратной функции справедливы условия 1 и 2. Естественно, для оператора A, соответствующего уравнению (5), указанные ограничения также имеют место, вследствие чего утверждения леммы 3 можно вывести из теоремы 3. Для получения здесь расширенной дифференцируемости по Фреше достаточно определить $W = L_2(\Omega)$.

Рассмотрим теперь банаховы пространства X_*, X_0, X; Y_1, Y^0, Y_*, Y_0, Y; Z_1, Z^0, Z_*, удовлетворяющие непрерывным вложениям $X_* \subset X_0 \subset X$, $Y_1 \subset Y^0 \subset Y_* \subset Y_0 \subset Y$, $Z_1 \subset Z^0 \subset Z_*$.

Определены операторы $L : X \to Y_1$ и $M : Y \to Z_1$, а также точки $x \in X$, $y = Lx$. По аналогии с классической теоремой о дифференцировании сложной функции ([17], с. 637) устанавливается

Теорема 5. *Если оператор L имеет $(X_0, Y^0; X_*, Y_*)$-расширенную производную Гато (соответственно Фреше) D_L в точке x, а оператор M — $(Y_0, Z^0; Y_*, Z_*)$-расширенную производную Фреше D_M в точке y, то композиция ML имеет $(X_0, Z^0; X_*, Z_*)$-расширенную производную Гато (соответственно Фреше) D_{ML} в точке x, равную $D_{ML} = D_M D_L$.*

Перейдем к определению категорий, связанных с расширенно дифференцируемыми операторами. В качестве класса объектов категории Γ_E выбираем семейство всевозможных троек (X, x, X_*), где X, X_* — банаховы пространства, удовлетворяющие непрерывному вложению $X \subset X_*$, а x — фиксированная точка из X. Рассмотрим две таких тройки (X, x, X_*) и (Y, y, Y_*). Пусть существуют банаховы пространства X_0, Y^0, Y_1, удовлетворяющие непрерывным вложениям $X_* \subset X_0 \subset X$, $Y_1 \subset Y^0 \subset Y_* \subset Y$, и такой оператор $L : X \to Y_1$, определенный в X_*-окрестности точки x и $(X_0, Y^0; X_*, Y_*)$-расширенно дифференцируемый по Фреше в точке x, что $Lx = y$. Композицию $L_x = \iota(Y_1, Y)L$, где $\iota(Y_1, Y)$ есть каноническое вложение пространства Y_1 в Y, будем называть морфизмом категории Γ_E, связывающим указанные выше объекты. Рассмотрим морфизм M_y с началом (Y, y, Y_*) и концом (Z, z, Z_*). Тогда существуют пространства Y_0, Z^0, Z_1, удовлетворяющие непрерывным вложениям $Y_* \subset Y_0 \subset Y$, $Z_1 \subset Z^0 \subset Z_* \subset Z$, и такой оператор $M : Y \to Z_1$, определенный в Y_*-окрестности точки y и $(Y_0, Z^0; Y_*, Z_*)$-расширенно дифференцируемый по Фреше в точке z, что $My = z$, причем $M_y = \iota(Z_1, Z)M$. Очевидно, композиция операторов ML определена в X_*-окрестности точки x и действует из пространства X в Z_1, причем $MLx = z$. В соответствии с теоремой 4 она является $(X_0, Z^0; X_*, Z_*)$-расширенно дифференцируемой по Фреше в точке x. Тогда $(ML)_z = \iota(Z_1, Z)ML$ оказывается морфизмом категории Γ_E, связывающим объекты (X, x, X_*) и (Z, z, Z_*). Его можно понимать как композицию $L_x \circ M_y$ морфизмов L_x и M_y. Тем самым Γ_E действительно образует категорию. Теорема 4 позволяет охарактеризовать ее изоморфизмы.

Определим преобразование $D_E : \Gamma_E \to \Sigma$. Для любого объекта (X, x, X_*) полагаем $D_E(X, x, X_*) = X_*$. Для любого морфизма L_x с началом (X, x, X_*) и концом (Y, y, Y_*) в качестве $D_E(L_x)$ выбираем расширенную производную Фреше D_L оператора L, ассоциированного с морфизмом L_x, в точке x. Пусть задан морфизм M_y категории Γ_E с началом (Y, y, Y_*) и концом (Z, z, Z_*). Тогда $D_E(M_y) = D_M$, где D_M — расширенная производная Фреше оператора M, ассоциированного с морфизмом M_y, в точке y. Согласно теореме 5 выполняется равенство $D_{ML} = D_M D_L$, где D_{ML} есть расширенная производная Фреше композиции операторов ML, определяющей морфизм $(ML)_z = L_x \circ M_y$. Тем самым справедливо равенство $D_E(L_x \circ M_y) = D_E(L_x) \circ D_E(M_y)$. Таким образом, отображение D_E действительно является функтором, действующим в рассматриваемых категориях.

Полученные результаты могут быть использованы для исследования задачи P. Предположим, что функционалы J и K для задачи P дифференцируемы по Фреше в пространствах W и Y_* соответственно. Тогда, заменив в теореме 3 обычные производные на расширенные с учетом теорем 4 и 5, приходим к следующему заключению.

Теорема 6. *Если оператор A относительно решения v задачи P удовлетворяет условиям теоремы 3, то справедливо равенство*

$$D_E(J_v) + \{H[D_E(A_y)]\}^{-1} D_E(K_y) = 0.$$

Поскольку теорема 6 связана с более слабыми ограничениями на оператор A, чем теорема 3, то она имеет более широкую область применения. В частности, для уравнения

(4) зависимость решения от свободного члена в общем случае (в отсутствии ограничения на малость показателя нелинейности и размерности области) является расширенно дифференцируемым, но не дифференцируемым по Гато. Функционалы J и K для задачи Q дифференцируемы по Фреше в пространствах W и Y_* соответственно, входящих в определение расширенной производной соответствующего обратного оператора. Тогда теорема 2 может быть получена как следствие из теоремы 6.

Категорная интерпретация позволяет взглянуть с иных позиций на необходимые условия экстремума. С ее помощью можно также заключить, что для наших целей эффективными оказываются такие обобщения операторных производных, которые будут обладать запасом свойств, достаточных для определения некоторой категории (аналоги теорем о сложной функции, об обратной функции).

Литература

[1] Дубовицкий А.Д., Милютин А.А. *Задачи на экстремум при наличии ограничений* // Журн. вычисл. матем. и матем. физ. – 1965. – Т. 5. – № 3. – С. 395–453.

[2] Neustadt L.W. *An abstract variational theory with application to a broad class of optimization problems* I // SIAM J. Control and Optim. – 1966. – № 4. – P. 505–527.

[3] Neustadt L.W. *An abstract variational theory with application to a broad class of optimization problems* II // SIAM J. Control and Optim. – 1967. – № 1. – P. 90–137.

[4] Иоффе А.Д., Тихомиров В.М. *Теория экстремальных задач.* – М.: Наука, 1974. – 480 с.

[5] Дмитрук А.В., Милютин А.А., Осмоловский Н.П. *Теорема Люстерника и теория экстремума* // УМН. – 1980. – Т. 35. – № 6. – С. 11–46.

[6] Обен Ж.-П., Экланд И. *Прикладной нелинейный анализ.* – М.: Мир, 1988. – 512 с.

[7] Лионс Ж.-Л. *Оптимальное управление системами, описываемыми уравнениями с частными производными.* – М.: Мир, 1972. – 416 с.

[8] Лионс Ж.-Л. *Некоторые методы решения нелинейных краевых задач.* – М.: Мир, 1972. – 596 с.

[9] Фурсиков А.В. *Оптимальное управление распределенными системами. Теория и приложения.* – Новосибирск: Научная книга, 1999. – 352 с.

[10] Иваненко В.И., Мельник В.С. *Вариационные методы в задачах управления для систем с распределенными параметрами.* – Киев: Наук. Думка, 1988. – 284 с.

[11] Райтум У.Е. *Задачи оптимального управления для эллиптических уравнений.* – Рига: Зинатне, 1989. – 280 с.

[12] Leblebicioglu M.K., Celebi A.O. *An optimal control problem with nonlinear elliptic state equations* // J. Math. Anal. Appl. – 1992. – V. 163. – № 1. – P. 178–205.

[13] Alibert J.J., Raymond J.P. *Boundary control of semilinear elliptic equations with discontinuous leading coefficients and unbounded control* // Num. Funct. Anal. Optim. – 1997. – V. 18. – № 3–4. – P. 235–250.

[14] Casas E., Troltzsch F., Unger A. *Second order sufficient optimality conditions for some state-constrainted control problems of semilinear elliptic equations* // SIAM J. Contr. And Optim. – 2000. – V. 38. – № 5. – P. 1369–1381.

[15] Лионс Ж.-Л. *Управление сингулярными распределенными системами.* – М.: Наука, 1987. – 368 с.

[16] Tiba D. *Optimal control of nonsmooth distributed parameter systems.* – Lecture Notes in Mathematics. – V. 1459. – Berlin: Springer Verlag, 1990. – 159 p.

[17] Канторович Л.В., Акилов Г.П. *Функциональный анализ.* – М.: Наука, 1977. – 744 с.

[18] Serovaiskii S.Ya. *Calculation of functional gradients and extended differentiation of operators* // J. of Inverse and Ill-Posed Problems. – 2005. – V. 13. – № 4. – P. 383–396.

[19] Neittaanmaki P., Tiba D. *Optimal control of nonlinear parabolic systems. Theory, algorithms, and applications.* – New York: Marcel Dekker, 1994. – 400 p.

[20] Кларк Ф. *Оптимизация и негладкий анализ.* – М.: Наука, 1988. – 280 с.

[21] Серовайский С.Я. *Дифференцирование обратной функции в ненормированных пространствах* // Функц. анализ и прилож. – 1993. – Т. 27. – № 4. – С. 84–87.

[22] Серовайский С.Я. *Теорема об обратной функции и расширенная дифференцируемость в банаховых пространствах* // Изв. вузов. Математика. – 1995. – № 8. – С. 39–49.

[23] Серовайский С.Я. *Оптимизация и дифференцирование.* Т. 1. *Минимизация функционалов. Стационарные системы.* – Алматы: Print-S, 2006. – 390 с.

[24] Букур И., Деляну А. *Введение в теорию категорий и функторов.* – М.: Мир, 1972. – 260 с.
[25] *Общая алгебра.* Т. 2 / Под общ. ред. Л.А. Скорнякова. – М.: Наука, 1977. – 480 с.
[26] Кухтенко А.И. *О теории категорий и топосов в задачах управления* // Слож. системы управл., Киев. – 1989. – С. 4–15.
[27] Елкин В.И. *О категориях и основах теории нелинейных управляемых динамических систем* // Дифференц. уравнения. – 2002. – Т. 38. – № 11. – С. 1467–1482.
[28] Авербух В.И., Смолянов О.Г. *Теория дифференцирования в линейных топологических пространствах* // УМН. – 1967. – Т. 22. – № 6. – С. 201–260.

С.Я. Серовайский

профессор, кафедра вычислительной математики,
Казахский национальный университет,
050012, Алматы, ул. Масанчи, д. 39/47,

e-mail: serovajskys@mail.ru

S.Ya. Serovaiskii

Professor, Chair Of Computational Mathematics,
Kazakh National University,
39/47 Masanchi str., Almaty, 050012 Republic of Kazakhstan,

e-mail: serovajskys@mail.ru