УДК 517.977.52, 517.977.58, 517.983.54, 519.85

# ПАРАМЕТРИЧЕСКАЯ ДВОЙСТВЕННАЯ РЕГУЛЯРИЗАЦИЯ В ОПТИМИЗАЦИИ, ОПТИМАЛЬНОМ УПРАВЛЕНИИ И ОБРАТНЫХ ЗАДАЧАХ

© **М. И. Сумин**

Ключевые слова: Линейно выпуклое математическое программирование; нелинейное математическое программирование; параметрическая задача; минимизирующая последовательность; двойственность; регуляризация; метод возмущений, принцип Лагранжа; оптимизация; оптимальное управление; обратные задачи.

Работа посвящена применению метода возмущений в теории двойственной регуляризации как для линейно выпуклой, так и для нелинейной задачи математического программирования в гильбертовом пространстве. Основное внимание в ней уделяется изучению качественных свойств метода двойственной регуляризации в зависимости от дифференциальных свойств функции значений ($S$-функции) оптимизационной задачи. Устанавливается теснейшая связь свойств сходимости метода с принципом Лагранжа. Показывается, что схема двойственной регуляризации дает новый способ доказательства принципа Лагранжа и приводит к его полезным уточнениям. Обсуждается так называемый регуляризованный принцип Лагранжа. Обсуждается также возможность применения метода двойственной регуляризации в параметрических задачах оптимизации, оптимального управления и в параметрических обратных задачах.

## Введение

Параметрическая двойственная регуляризация или, другими словами, параметрическая регуляризация на основе теории двойственности заключается в исследовании (решении) двойственной к исходной параметрической (то есть зависящей от параметров) задаче оптимизации с привлечением методов регуляризации. Метод параметрической двойственной регуляризации служит основой как для исследований в области теории оптимизации, оптимального управления, обратных задач, так и для конструирования на основе его идеологии новых алгоритмов решения таких задач.

Основное отличие метода параметрической двойственной регуляризации от других подходов в теории регуляризации оптимизационных задач заключается в том, что в нем: 1) самым существенным и непосредственным образом используется классическая идея «снятия» ограничений, заложенная в принципе Лагранжа; 2) непосредственно используется идеология метода возмущений; 3) по-видимому, наиболее полно используются преимущества оптимизационной техники, развитой для задач с операторными ограничениями, в основе которой лежат разработанные в последние десятилетия методы негладкого (нелинейного) анализа.

Двойственные алгоритмы традиционно являются одними из наиболее популярных и эффективных при решении оптимизационных задач с ограничениями [1, 2]. Целенаправленное изучение и применение метода двойственности в теории алгоритмов решения задач математического программирования началось, по-видимому, с работы Х.Удзавы [3]. К основным проблемам, связанным с классическим алгоритмом Удзавы, как отмечено в [4], можно отнести его неустойчивость к ошибкам исходных данных и потребность в существовании седловой точки функции Лагранжа оптимизационной задачи при доказательстве сходимости (соответствующие теоремы сходимости могут быть найдены, например, в [5–7]). Подход к преодолению указанных трудностей применения метода двойственности для решения оптимизационных задач с ограничениями был предложен в работах [8–11] на пути его объединения с методами регуляризации решения некорректных задач [12]. В дальнейшем этот подход был развит в работах [4, 13–15] применительно к линейно выпуклой задаче математического программирования общего вида с сильно выпуклым целевым функционалом. В то же время, как известно (подробности см., например, в [13–16]), двойственность неразрывно связана и с самим методом регуляризации Тихонова [12].

Рассмотрим для иллюстрации простейшую линейно выпуклую задачу математического программирования в гильбертовом пространстве

$$f(z) \to \min, \quad Az = h, \quad z \in \mathcal{D} \subset Z, \tag{1}$$

где $f: Z \to R^1$ – сильно выпуклый функционал, $A: Z \to H$ – линейный непрерывный оператор, $h \in H$ – фиксированный элемент, $\mathcal{D}$ – выпуклое замкнутое множество, $Z, H$ – гильбертовы пространства. Пусть решение задачи (1) существует. Обозначим это единственное решение через $z^0$.

Напомним, что классический двойственный алгоритм Удзавы [3] применительно к решению задачи (1) представляет собой непосредственное решение на основе градиентной итерационной процедуры задачи, двойственной к исходной оптимизационной задаче (1)

$$\lambda^{k+1} = \lambda^k + \beta_k(Az^k - h), \quad k = 1, 2, \ldots, \ \lambda^1 \in H,$$

$$z^k = argmin\,\{L(z, \lambda^k),\, z \in \mathcal{D}\}, \quad L(z,\lambda) \equiv f(z) + \langle \lambda, Az - h\rangle.$$

При некоторых достаточно сильных предположениях он сходится одновременно по двойственной и прямой переменным и выполняются предельные соотношения (напомним, что классические теоремы сходимости алгоритма Удзавы можно найти в [5–7])

$$\|\lambda^k - \lambda^0\| \to 0, \quad \|z^k - z^0\| \to 0, \quad k \to \infty,$$

где $\lambda^0$ – решение двойственной задачи $\min\limits_{z\in\mathcal{D}} L(z,\lambda) \to \max, \quad \lambda \in H$.

Как уже отмечалось выше, проблемы классического алгоритма Удзавы заключаются в его неустойчивости к ошибкам исходных данных и в требовании при обосновании сходимости существования седловой точки функции Лагранжа оптимизационной задачи.

Неустойчивость к ошибкам исходных данных иллюстрирует

П р и м е р 1. Пусть имеется задача минимизации сильно выпуклой квадратичной функции двух переменных на множестве, задаваемом аффинным ограничением типа равенства, эквивалентная задаче поиска нормального решения линейной алгебраической системы двух уравнений с двумя неизвестными

$$|x|^2 \to \min, \ Ax = y, \ x \in R^2,$$

$$A \equiv \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \ y \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \ x^* = (0,5; 0,5) \ - \text{нормальное решение.}$$

Двойственная задача:

$$V(\lambda) \equiv L(x(\lambda), \lambda) = -\frac{1}{4}\langle AA^*\lambda, \lambda\rangle - \langle y, \lambda\rangle \to \max, \ \lambda \in R^2,$$

где $L(x,\lambda) \equiv |x|^2 + \langle \lambda, Ax - y\rangle$, $x(\lambda) \equiv argmin\{L(x,\lambda) : x \in R^2\} = -\frac{1}{2}A^*\lambda$. Ее решение $(-1, \alpha)\, \forall\, \alpha \in \in R^1$.

Возмущенная задача:

$$|x|^2 \to \min, \ A^\delta x = y^\delta, \ x \in R^2, \quad A^\delta \equiv \begin{pmatrix} 1 & 1 \\ 0 & \delta^2 \end{pmatrix}, \ y^\delta \equiv \begin{pmatrix} 1 \\ \delta \end{pmatrix}, \quad \delta > 0.$$

Возмущенная двойственная задача:

$$V^\delta(\lambda) \equiv L^\delta(x^\delta(\lambda),\lambda) \to \max, \ \ \lambda \in R^2.$$

Ее решение $\lambda^\delta = (\frac{2-2\delta}{\delta}, \frac{2\delta-4}{\delta^3})$. Вектор $x^\delta(\lambda^\delta) \equiv argmin\{L^\delta(x,\lambda^\delta) : x \in R^2\} = (1-\frac{1}{\delta}, \frac{1}{\delta})$ есть в соответствии с классическим алгоритмом Удзавы «приближенное» решение исходной задачи, но $x^\delta(\lambda^\delta) \nrightarrow x^*$.

Далее следуют два примера, в которых для функций Лагранжа справедливо равенство

$$\sup_{\lambda\in H}\inf_{z\in\mathcal{D}} L(z,\lambda) = \inf_{z\in\mathcal{D}}\sup_{\lambda\in H} L(z,\lambda),$$

но внешний экстремум в левой части не достигается и которые иллюстрируют то обстоятельство, что требование существования седловой точки у функции Лагранжа оптимизационной задачи является весьма жестким.

П р и м е р 2. Рассмотрим задачу минимизации

$$\|z\|^2 \to \min, \ \ A(z)(x) \equiv \int_a^b K(x,s)z(s)\,ds = q(x), \tag{2}$$

с $z \in L_2(a,b)$, $q \in L_2(a,b)$ и с замкнутым симметрическим непрерывным на квадрате $\Pi \equiv [a,b] \times \times [a,b]$ ядром $K$ интегрального оператора. Задача (2) эквивалентна задаче поиска нормального решения интегрального уравнения Фредгольма первого рода

$$A(z)(x) \equiv \int_a^b K(x,s)z(s)\,ds = q(x), \ \ a \leqslant x \leqslant b.$$

Пусть $z_0 \in L_2(a,b)$ – разрывная функция (соответствующий класс эквивалентности не содержит непрерывной функции), $q_0(x) \equiv A(z_0)(x)$, $x \in [a,b]$. Уравнение $A(z)(x) = q_0(x)$, $a \leqslant x \leqslant b$ имеет в силу замкнутости ядра единственное решение $z_0$. Соответствующая эквивалентная задача минимизации

$$\|z\|^2 \to \min, \ \ Az = q_0.$$

Пусть

$$V_{q_0}(\lambda) = \min_{z\in L_2(a,b)}\{\|z\|^2 + \langle\lambda, Az - q_0\rangle\}.$$

Двойственная задача

$$V_{q_0}(\lambda) = -1/4\langle AA\lambda,\lambda\rangle - \langle q_0,\lambda\rangle \to \max, \ \ \lambda \in L_2(a,b) \tag{3}$$

решения не имеет.

Действительно, если бы некоторая точка $\lambda \in L_2(a,b)$ была решением двойственной задачи, то она удовлетворяла бы равенству $-1/2AA\lambda = q_0$. Но тогда, в силу замкнутости ядра, мы должны были бы иметь равенство $-1/2A\lambda = z_0$, которое противоречиво, так как $z_0$ – функция, не являющаяся непрерывной, а $A\lambda$ – непрерывная функция (ядро $K$ непрерывно).

Интересно, что последними рассуждениями, по сути дела, доказано также, что в задаче минимизации (2), эквивалентной исходному уравнению при $q = q_0$, принцип Лагранжа не выполняется. Действительно, если бы это было не так, то существовала бы пара множителей Лагранжа $(\mu,\lambda) \in R^1_+ \times L_2(a,b)$, $(\mu,\lambda) \neq 0$, $R^1_+ \equiv \{x \in R^1 : x \geqslant 0\}$, такая, что

$$2\mu z_0 + A\lambda = 0, \quad Az_0 = q_0.$$

Тогда в случае $\mu > 0$ мы бы имели равенство $\partial V_{q_0}(\lambda/\mu) = -\frac{1}{2}AA(\lambda/\mu) - q_0 = 0$, то есть элемент $\lambda/\mu$ доставлял бы максимальное значение в двойственной задаче (3), что противоречит доказанной выше ее неразрешимости. В случае же $\mu = 0$ мы имеем равенство $A\lambda = 0$, из которого в силу инъективности оператора $A$ получаем $\lambda = 0$. Таким образом, предположение о том, что принцип Лагранжа в задаче (2) при $q = q_0$ выполняется, приводит к противоречию, заключающемуся в вырожденности соответствующей пары множителей Лагранжа. Одновременно можно заметить также, что точки вида $z_0$ лежат всюду плотно в $L_2(a,b)$ и, как следствие, всюду плотно во множестве всех тех $q$, для которых разрешимо равенство $Az = q,\ z \in L_2(a,b)$, лежат точки вида $q_0$, для которых принцип Лагранжа в задаче $\|z\|^2 \to \min,\ \ Az = q_0$ не выполняется.

П р и м е р 3. Рассмотрим обратную задачу финального наблюдения, в которой требуется найти начальное условие $v(x),\ x \in (0,1)$ по известному в финальный момент времени $T$ решению $z[v](\cdot,T) = q \in L_2(0,1)$ третьей краевой задачи для уравнения теплопроводности

$$\partial z/\partial t - \partial^2 z/\partial x^2 = 0,\ \ z(x,0) = v(x),\ x \in (0,1),$$

$$\partial z(0,t)/\partial x - z(0,t) = 0,\ \ \partial z(1,t)/\partial x + z(1,t) = 0,\quad t \in [0,T],\ \ v(x) \in V \equiv [-1,1].$$

Эквивалентная задача минимизации имеет здесь вид

$$\int_0^1 v^2(x)\,dx \to \inf,\quad z[v](\cdot,T) = q,\ v \in \mathcal{D} \equiv \{v \in L_2(0,1) : v(x) \in [-1,1]\,\text{п.в. на}\,(0,1)\}. \qquad (4)$$

Пусть

$$L_q(v,\lambda) \equiv \int_0^1 v^2(x)\,dx + \int_0^1 \lambda(x)(z[v](x,T) - q(x))\,dx,\ v \in \mathcal{D},\ \lambda \in L_2(0,1),$$

$$V_q(\lambda) \equiv \min_{v \in \mathcal{D}} L_q(v,\lambda).$$

Пусть $\bar{v} \in \mathcal{D}$ разрывная кусочно-непрерывная функция, $\bar{v}(x) \in (-1,1)$ и $q = q_{\bar{v}} = z[\bar{v}](\cdot,T)$. При $q = q_{\bar{v}}$ двойственная к (4) задача $V_{q_{\bar{v}}}(\lambda) \to \max,\ \ \lambda \in L_2(0,1)$ решения не имеет (подробности см. в [4]).

Очевидно, функции $\bar{v}$ указанного вида лежат всюду плотно в $\mathcal{D}$ и, как следствие, всюду плотно во множестве всех тех $q$, для которых разрешимо равенство $z[v](\cdot,T) = q,\ v \in \mathcal{D}$, лежат точки вида $q_{\bar{v}}$, для которых функционал $V_{q_{\bar{v}}}$ не достигает максимального значения на $L_2(0,1)$.

Задачи вида (1) возникают во многих приложениях и, в частности, в приложениях, связанных в том числе и с нелинейными обратными задачами наблюдения для дифференциальных уравнений, в которых в результате измерений имеется информация о траектории системы «в целом» (см., в частности, [17]).

П р и м е р 4. Пусть имеется управляемая нелинейная система

$$\dot{y} = f(t,y)u(t) + g(t,y),\ \ 0 \leqslant t \leqslant T,\ \ y(0) = y_0,\ \ u(t) \in P\ \text{ п.в. на } (0,T),$$

где для определенности $f : R^1 \times R^n \to R^{n\times r}$, $g : R^1 \times R^n \to R^n$ – функции, удовлетворяющие глобальному условию Липшица по совокупности переменных, $P \subset R^r$ – компакт. Если в нашем распоряжении имеется наблюдаемая траектория $y^0(t)$, $0 \leqslant t \leqslant T$, и нужно определить «вызывающее» эту траекторию минимальное по норме управление $u^0 \in \mathcal{D} \equiv \{u \in L_2(0,T) : u(t) \in$ $\in P$ п.в. на $(0,T)\} \subset L_2(0,T)$, то невозмущеннная линейно выпуклая задача (1) в этом случае имеет вид

$$f_0(u) \equiv \|u\|^2 \to \min,$$

$$A[u](t) \equiv \int_0^t f(s,y^0(s))u(s)ds = h(t) \equiv y^0(t) - \int_0^t g(s,y^0(s))ds - y_0,\ t \in [0,T],\ u \in \mathcal{D}.$$

Соответственно возмущенная линейно выпуклая задача в случае приближенно наблюдаемой траектории $y^\delta(t)$, $|y^\delta(t) - y^0(t)| \leqslant \delta,\ 0 \leqslant t \leqslant T$, записывается в том же виде, но с возмущенным аффинным ограничением

$$A^\delta[u](t) \equiv \int_0^t f(s,y^\delta(s))u(s)ds = h^\delta(t) \equiv y^\delta(t) - \int_0^t g(s,y^\delta(s))ds - y_0,\ t \in [0,T].$$

В настоящей работе метод параметрической двойственной регуляризации излагается как в случае линейно выпуклой, так и в случае нелинейной задачи математического программирования. Основное внимание при этом мы уделяем изучению его связи с теорией двойственности, принципом Лагранжа, а также изучению зависимости свойств сходимости метода (процесса конструирования минимизирующей последовательности) от дифференциальных свойств функции значений ($S$-функции) задачи математического программирования. Как в линейно выпуклом, так и в нелинейном случаях приводятся возможные постановки задач оптимального управления и обратных задач, которые могут быть сведены к рассматриваемым в работе задачам математического программирования.

## 1. Параметрическая двойственная регуляризация в линейно выпуклой задаче математического программирования

В данном разделе мы изучаем линейно выпуклую задачу математического программирования с сильно выпуклым функционалом цели. Основной конструкцией при этом является классическая конструкция функции Лагранжа, а основным искомым объектом – минимизирующая последовательность. В заключение раздела кратко обсуждается возможность применения метода двойственной регуляризации в линейно выпуклой задаче математического программирования с выпуклым функционалом цели, а также в линейно выпуклой задаче оптимального управления с поточечными фазовыми ограничениями типа равенства и неравенства.

### 1.1. Постановка линейно выпуклой задачи математического программирования в случае сильно выпуклого целевого функционала

Рассмотрим параметрическую задачу минимизации

$$(P_{p,r}) \qquad f(z) \to \min, \;\; Az = h + p, \;\; g_i(z) \leqslant r, \; i = 1, \dots, m,$$

$$z \in \mathcal{D} \subset Z, \quad p \in H, \;\; r = (r_1, \dots, r_m)^* \in R^m \;\; - \;\; \text{параметры},$$

где $f : Z \to R^1$ – сильно выпуклый функционал, $A : Z \to H$ – линейный непрерывный оператор, $h \in H$ – заданный элемент, $g_i : Z \to R^1$, $i = 1, \dots, m$, – выпуклые функционалы, $g(z) \equiv (g_1(z), \dots, g_m(z))^*$, $\mathcal{D}$ – выпуклое замкнутое множество, $Z$, $H$ – гильбертовы пространства. Будем также считать, что

$$|f(z_1) - f(z_2)|, \; |g_i(z_1) - g_i(z_2)| \leqslant L_M \|z_1 - z_2\| \; \forall \, z_1, \, z_2 \in \mathcal{D} \cap S_M,$$

где $L_M > 0$ – постоянная, $S_M \equiv \{z \in Z : \|z\| \leqslant M\}$.

Обозначим единственное решение задачи $(P_{p,r})$, если оно существует, через $z_{p,r}$.

Важное значение при конструировании двойственного алгоритма в задаче $(P_{p,r})$ имеет понятие минимизирующего приближенного решения в смысле Дж. Варги [18]. Напомним, что под минимизирующим приближенным решением понимается такая последовательность $z^i$, $i = 1, 2, \dots$, для которой справедливы соотношения $f(z^i) \leqslant \beta(p,r) + \delta^i$, $z^i \in \mathcal{D}^{\varepsilon^i}_{p,r}$ для некоторых последовательностей сходящихся к нулю неотрицательных чисел $\delta^i$, $\varepsilon^i$, $i = 1, 2, \dots$. Здесь $\beta(p,r)$ – обобщенная нижняя грань – функция значений – S-функция задачи $(P_{p,r})$:

$$\beta(p,r) \equiv \beta_{+0}(p,r) \equiv \lim_{\varepsilon \to +0} \beta_\varepsilon(p,r), \quad \beta_\varepsilon(p,r) \equiv \inf_{z \in \mathcal{D}^{0\varepsilon}_{p,r}} f(z), \qquad \beta_\varepsilon(p,r) \equiv +\infty, \text{ если } \mathcal{D}^{0\varepsilon}_{p,r} = \emptyset,$$

$$\mathcal{D}^{0\varepsilon}_{p,r} \equiv \{z \in \mathcal{D} : \|Az - h - p\| \leqslant \varepsilon, \; g_i(z) \leqslant r_i + \varepsilon, \; i = 1, \dots, m\}, \; \varepsilon \geqslant 0.$$

Очевидно, в общей ситуации $\beta(p,r) \leqslant \beta_0(p,r)$, где $\beta_0(p,r)$ – классическое значение задачи. Но в случае поставленной выше задачи $(P_{p,r})$ имеет место равенство $\beta(p,r) = \beta_0(p,r)$.

Справедливы важные для дальнейшего изложения следующие леммы (подробности см., например, в [15]).

Л е м м а 1. *Функционал* $\beta: H \times R^m \to R^1 \cup \{+\infty\}$ *является выпуклым и полунепрерывным снизу.*

Л е м м а 2. *Если* $\beta(p,r) < +\infty$, *то нижняя грань* $\beta(p,r)$ *в задаче* $(P_{p,r})$ *достигается и справедливо равенство* $\beta_0(p,r) = \beta(p,r)$.

Отсюда и из определения минимизирующего приближенного решения $z^i$, $i = 1, 2, \ldots$, вытекает справедливость предельного соотношения

$$f(z^i) \to f(z_{p,r}) = \beta_0(p,r) = \beta(p,r), \;\; i \to \infty,$$

если решение задачи $(P_{p,r})$ существует.

Пусть $\mathsf{F}$ – множество всевозможных наборов исходных данных $\mathsf{f} \equiv \{f,\, A,\, h,\, g_i\}$, каждый из которых состоит из сильно выпуклого на $\mathcal{D}$ с не зависящей от набора постоянной сильной выпуклости $\kappa > 0$ функционала $f$, линейного непрерывного оператора $A$, элемента $h$ и выпуклых на $\mathcal{D}$ функционалов $g_i$, $i = 1, \ldots, m$, причем справедливы оценки

$$|f(z_1) - f(z_2)|,\; |g_i(z_1) - g_i(z_2)| \leqslant L_M \|z_1 - z_2\| \;\forall\, z_1,\, z_2 \in \mathcal{D} \cap S_M$$

с не зависящей от набора постоянной $L_M$.

Определим наборы невозмущенных $\mathsf{f}^0$ и возмущенных $\mathsf{f}^\delta$ исходных данных соответственно: $\mathsf{f}^0 \equiv \{f^0,\, A^0,\, h^0,\, g_i^0\}$ и $\mathsf{f}^\delta \equiv \{f^\delta,\, A^\delta,\, h^\delta,\, g_i^\delta\}$, $\delta \in (0, \delta_0]$, $\delta_0 > 0$ - некоторое число. Будем считать, что

$$|f^\delta(z) - f^0(z)|,\, |g^\delta(z) - g^0(z)| \leqslant C\delta(1 + \|z\|),\; \|A^\delta - A^0\|,\;\; \|h^\delta - h^0\| \leqslant C\delta, \tag{5}$$

где $C > 0$ не зависит от $\delta$, $g^\delta \equiv (g_1^\delta, \ldots, g_m^\delta)^*$.

К линейно выпуклым задачам общего вида $(P_{p,r})$ сводятся самые разнообразные задачи. К ним относятся, например, линейно выпуклые задачи оптимального управления, линейные, а также некоторые нелинейные (см. пример 4) обратные задачи.

Одним из характерных примеров задач оптимального управления, которая может быть записана в форме $(P_{p,r})$ в гильбертовом пространстве $Z = L_2(0,T)$ является, в частности, параметрическая задача оптимального управления с фиксированным временем, а также с операторным ограничением типа равенства и конечным числом функциональных ограничений типа неравенства (см. также [11])

$$g_0(u) \equiv \int_0^T (\langle F(t)x[u](t), x[u](t)\rangle + \langle G(t)u(t), u(t)\rangle)dt \to \min,$$

$$\langle \varphi_1(t), x[u](t)\rangle = h(t) + p(t) \text{ при п.в. } t \in X,\;\; \varphi_{2,i}(x[u](T)) \leqslant r_i \;\; i = 1, \ldots, m,$$

$$\dot{x} = A(t)x + B(t)u(t)),\;\; x(0) = x_0 \in R^n,\quad t \in [0,T],$$

$$u \in \mathcal{D} \subset L_2(0,T),\quad p \in H,\; r = (r_1, \ldots, r_m)^* - \text{ параметры},$$

где $g_0 : L_2(0,T) \to R^1$ – сильно выпуклый функционал с постоянной $\kappa$, $F$, $A : [0,T] \to R^{n\times n}$, $B : [0,T] \to R^{n\times m}$, $G : [0,T] \to R^{m\times m}$ – непрерывные матрицы, $\varphi_1$, $h \in C[0,T]$, – заданные функции, $\varphi_{2,i} : R^n \to R^1$ – выпуклая, непрерывная вместе с градиентом $\nabla_x \varphi_{2,i}$ функция, $\mathcal{D} \equiv \{u \in \in L_2(0,T) : u(t) \in U \text{ при п.в. } t \in (0,T)\}$, $U \subset R^m$ – выпуклый компакт, $X \subset [0,T]$, $X = X$, $\mathcal{H} \equiv L_2(X)$. Считаем при этом одновременно, что для каждого управления $u \in \mathcal{D}$ существует единственное решение задачи Коши $x[u](t)$, $t \in [0,T]$, причем все эти решения равномерно ограничены.

Если же вести речь о линейных обратных задачах, которые могут быть записаны в форме задачи $(P_{p,r})$, то такой задачей может служить, в частности, обратная задача финального наблюдения из рассмотренного выше примера 3 (см. также [9]). Естественно, к виду $(P_{p,r})$ сводятся и многие другие линейные обратные задачи. Конкретные прикладные линейные обратные задачи рассматривались на основе идеологии двойственной регуляризации, например, в [19, 20].

Введем далее функционал Лагранжа

$$L^{\delta}_{p,r}(z,\lambda,\mu) \equiv f^{\delta}(z) + \langle \lambda, A^{\delta}z - h^{\delta} - p\rangle + \langle \mu, g^{\delta}(z) - r\rangle,\ \ z \in \mathcal{D}$$

и вогнутый двойственный функционал – функционал значений

$$V^{\delta}_{p,r}(\lambda,\mu) \equiv \inf_{z\in\mathcal{D}} L^{\delta}_{p,r}(z,\lambda,\mu),\ \ \lambda \in H,\ \mu \in R^m.$$

Ввиду сильной выпуклости функционала Лагранжа для любой пары $(\lambda,\mu) \in H \times R^m_+$, где $R^m_+ \equiv \{x = (x_1,\dots,x_m)^* \in R^m : x_i \geqslant 0,\ i = 1,\dots,m\}$, значение $V^{\delta}_{p,r}(\lambda,\mu)$ достигается на единственном элементе $z^{\delta}[\lambda,\mu] \equiv argmin\,\{L^{\delta}_{p,r}(z,\lambda,\mu),\ z \in \mathcal{D}\}$. В случае ограниченного множества $\mathcal{D}$ двойственный функционал $V^{\delta}_{p,r}$, очевидно, определен и конечен для любого элемента $(\lambda,\mu) \in H \times \times R^m$.

Приведем формулу для градиента функционала $V^{\delta}_{p,r}$ на $H \times R^m_+$ при $\delta \geqslant 0$, доказательство которой можно найти в [4].

Л е м м а 3. *Супердифференциал (в смысле выпуклого анализа) вогнутого функционала $V^{\delta}_{p,r}$ в каждой внутренней точке $(\lambda,\mu)$ множества $H \times R^m_+$ равен*

$$\partial V^{\delta}_{p,r}(\lambda,\mu) = (A^{\delta}z^{\delta}[\lambda,\mu] - h^{\delta}, g^{\delta}(z^{\delta}[\lambda,\mu]))$$

*и совпадает с производной в смысле Фреше этого функционала. В каждой точке $(\lambda,\mu)$ множества $H \times R^m_+$ имеет место включение*

$$(A^{\delta}z^{\delta}[\lambda,\mu] - h^{\delta}, g^{\delta}(z^{\delta}[\lambda,\mu])) \in \partial V^{\delta}_{p,r}(\lambda,\mu).$$

*В случае ограниченного множества $\mathcal{D}$, а также в случае глобальной липшицевости функций $g^{\delta}_i$, $i = 1,\dots,m$ при неограниченном $\mathcal{D}$ элемент $(A^{\delta}z^{\delta}[\lambda,\mu] - h^{\delta}, g^{\delta}(z^{\delta}[\lambda,\mu]))$ удовлетворяет в $H \times R^m_+$ условию Липшица с постоянной $C/\kappa$, где $C > 0$ не зависит от $(\lambda^1,\mu^1)$, $(\lambda^2,\mu^2) \in H \times \times R^m_+$.*

## 1.2. Параметрический принцип Лагранжа

Сформулируем принцип Лагранжа в параметрической задаче $(P^0_{p,r})$, доказательство которого можно найти в [13, 14]. Предположим, что функционалы $f^0 : Z \to R^1$, $g^0_i : Z \to R^1$ являются непрерывно дифференцируемыми по Фреше.

Т е о р е м а 1. *[Параметрический принцип Лагранжа] Пусть $\beta^0(p,r) < +\infty$ и $z^0_{p,r} \in \mathcal{D}^0_{p,r} \equiv \{z \in \mathcal{D} : A^0z - h^0 - p = 0,\ g^0_i(z) \leqslant r_i,\ i = 1,\dots,m\}$ – оптимальный элемент в задаче $(P^0_{p,r})$, то есть $f^0(z^0_{p,r}) = \beta^0(p,r)$. Тогда, если $\zeta \in \partial\beta^0(p,r)$, где $\partial\beta^0(p,r)$ – субдифференциал в смысле выпуклого анализа, то для множителей Лагранжа $y \in H$, $\xi \in R^m_+$, $(y,\xi) = -\zeta$, при $\mu_0 = 1$ выполняются соотношения*

$$\langle \mu_0\nabla f^0(z^0_{p,r}) + A^{0*}y + \sum_{i=1}^{m}\xi_i\nabla g^0_i(z^0_{p,r}), z - z^0_{p,r}\rangle \geqslant 0\ \ \forall\, z \in \mathcal{D}, \tag{6}$$

$$\xi_i(g^0_i(z^0_{p,r}) - r_i) = 0,\ \ i = 1,\dots,m$$

*и при этом $-\zeta = (y,\xi)$ – вектор Куна-Таккера задачи $(P^0_{p,r})$.*

*Если же $\zeta \in \partial^{\infty}\beta^0(p,r)$, где $\partial^{\infty}\beta^0(p,r)$ – сингулярный (асимптотический) субдифференциал, определяемый формулой*

$$\partial^{\infty}\beta^0(p,r) \equiv \{(y,\xi) \in H \times R^m : ((y,\xi),0) \in N_{epi\,\beta^0}((p,r),\beta^0(p,r))\},$$

*то для множителей Лагранжа* $y \in H,\ \xi \in R_+^m,\ (y, \xi) = -\zeta$, *соотношения* (6) *выполняются при* $\mu_0 = 0$.

*И, наоборот, если* $z_{p,r}^0 \in \mathcal{D}_{p,r}^0$ *такой элемент, что при некоторых* $\mu_0 > 0,\ y \in H,\ \xi \in R_+^m$ *выполняются соотношения* (6), *то этот элемент оптимален в задаче* $(P_{p,r}^0)$, *пара* $(y/\mu_0, \xi/\mu_0)$ *является вектором Куна-Таккера для нее и одновременно* $(-y/\mu_0, -\xi/\mu_0) \in \partial\beta^0(p,r)$.

*Если же* $z_{p,r}^0 \in \mathcal{D}_{p,r}^0$ *такой элемент, что при* $\mu_0 = 0$ *и некоторых* $y \in H,\ \xi \in R_+^m,\ (y, \xi) \neq 0$, *выполняются соотношения* (6), *то* $(p, r) \in \partial\,\beta^0$ *и одновременно* $(-y, -\xi) \in \partial^\infty \beta^0(p, r)$.

З а м е ч а н и е 1. Важным является то, что этим классическим принципом Лагранжа «не охватываются» задачи $(P_{p,r})$, для которых одновременно $\partial\beta(p, r) = \emptyset$ и $\partial^\infty\beta(p, r) = \{0\}$, что вполне возможно для задач с ограничениями, задаваемыми операторами с бесконечномерными образами. Одной из таких задач является задача минимизации

$$\|z\|^2 \to \min,\ A(z)(x) \equiv \int_a^b K(x, s)z(s)\,ds = p(x),\ a \leqslant x \leqslant b \tag{7}$$

с $z \in L_2(a, b)$, с параметром $p \in L_2(a, b)$ и с замкнутым симметрическим непрерывным на квадрате $\Pi \equiv [a, b] \times [a, b]$ ядром $K$ интегрального оператора. Задача (7) эквивалентна задаче поиска нормального решения интегрального уравнения Фредгольма первого рода

$$A(z)(x) \equiv \int_a^b K(x, s)z(s)\,ds = p(x),\ \ a \leqslant x \leqslant b.$$

Как показано выше при анализе примера 2 (см. также [4]), в параметрической задаче (7) плотно во множестве всех $p$, для которых это уравнение разрешимо, лежат такие $p$, для которых в задаче (7) принцип Лагранжа не выполняется, что равносильно в рассматриваемой ситуации совокупности соотношений $\partial\beta(p, r) = \emptyset$, $\partial^\infty\beta(p, r) = \{0\}$.

## 1.3. Параметрическая двойственная регуляризация

Опишем регуляризованный двойственный алгоритм [4,8–11, 13, 14] для задачи $(P_{p,r}^0)$ в случае ограниченного множества $\mathcal{D}$ и изучим его поведение в зависимости от параметров $p, r$ и дифференциальных свойств функции значений $\beta^0$.

Обозначим через $(\lambda_{p,r}^{\delta,\alpha}, \mu_{p,r}^{\delta,\alpha})$ единственную в $H \times R_+^m$ точку, дающую на этом множестве максимум функционалу

$$R_{p,r}^{\delta,\alpha}(\lambda, \mu) \equiv V_{p,r}^\delta(\lambda, \mu) - \alpha\|\lambda\|^2 - \alpha|\mu|^2,\ \ (\lambda, \mu) \in H \times R_+^m.$$

Пусть выполняется условие согласования

$$\frac{\delta}{\alpha(\delta)} \to 0,\ \ \alpha(\delta) \to 0,\quad \delta \to 0. \tag{8}$$

Аппроксимация решения $z_{p,r}^0$ задачи $(P_{p,r}^0)$ при $\delta \to 0$ происходит с помощью регуляризованных элементов $z^\delta[\lambda^{\delta,\alpha(\delta)}, \mu^{\delta,\alpha(\delta)}]$.

Справедлива следующая теорема, доказательство которой можно найти в [4,9–11, 13, 14].

Т е о р е м а 2. [*Регуляризирующий двойственный алгоритм*] *Вне зависимости от того, разрешима или нет, двойственная к* $(P_{p,r}^0)$ *задача, при условии согласования* (8) *выполняются соотношения*

$$\alpha(\delta)\|(\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)})\| \to 0,\quad f^0(z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]) \to f^0(z_{p,r}^0),\ \ \delta \to 0,$$

$$A^0 z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}] - h^0 \to 0,\ \ g_i(z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}]) \leqslant \kappa(\delta),\ \kappa(\delta) \to 0,\ \delta \to 0.$$

*Если же сильно выпуклый функционал $f^0$ является и субдифференцируемым (в смысле выпуклого анализа) в точках $\mathcal{D}$, то справедливо и предельное соотношение*

$$\|z^\delta[\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)}] - z_{p,r}^0\| \to 0, \ \delta \to 0.$$

*Другими словами, вне зависимости от того, разрешима или нет двойственная задача, регуляризованный двойственный алгоритм представляет собой регуляризирующий в смысле [12] алгоритм.*

Итак, в зависимости от того, имеет или нет решение двойственная к $(P_{p,r}^0)$ задача, сконструированный выше двойственный алгоритм ведет себя двояко. В случае существования решения двойственной задачи генерируемое в соответствии с алгоритмом семейство двойственных переменных $(\lambda_{p,r}^{\delta,\alpha(\delta)}, \mu_{p,r}^{\delta,\alpha(\delta)})$, $\delta \to 0$ ограничено в $H \times R_+^m$. Если же такого решения нет, то нормы элементов этого семейства неограниченно возрастают при $\delta \to 0$.

В случае неограниченного множества $\mathcal{D}$ сформулированный алгоритм заведомо сохраняет силу при $\delta = 0$ (подробности можно найти в [4, 13, 14]).

### 1.4. Применение двойственной регуляризации в теории двойственности

Выясним, как устроено множество всех значений параметров $(p, r) \in \beta^0$, для которых соответствующая двойственная задача разрешима.

Оказывается, для функции значений $\beta^0(p, r) : H \times R^m \to R^1 \bigcup \{+\infty\}$, которую можно определить посредством равенства

$$\beta^0(p, r) \equiv \{f^0(z_{p,r}^0), \text{ если } z_{p,r}^0 \text{существует; } +\infty \text{ в противном случае}\} =$$

$$= \inf_{z \in \mathcal{D}} \sup_{(\lambda,\mu) \in H \times R_+^m} L_{p,r}^0(z, \lambda, \mu),$$

справедливо представление

$$\beta^0(p, r) = \sup_{(\lambda,\mu) \in H \times R_+^m} V_{p,r}^0(\lambda, \mu) = \sup_{(\lambda,\mu) \in H \times R_+^m} \min_{z \in \mathcal{D}} L_{p,r}^0(z, \lambda, \mu) \ \forall \ (p, r) \in H \times R^m,$$

которое может быть переписано в форме классического равенства

$$\sup_{(\lambda,\mu) \in H \times R_+^m} \min_{z \in \mathcal{D}} L_{p,r}^0(z, \lambda, \mu) = \inf_{z \in \mathcal{D}} \sup_{(\lambda,\mu) \in H \times R_+^m} L_{p,r}^0(z, \lambda, \mu) \ \forall \ (p, r) \in H \times R^m. \tag{9}$$

Равенство (9) уточняет классический результат – теорему о минимаксе из книги [21], представляющую собой формулировку достаточного условия для его выполнимости, согласно которой в рассматриваемой задаче равенство (9) выполняется заведомо лишь в точках непрерывности функции значений (S-функции) $\beta^0$. Однако можно утверждать, что в случае сильно выпуклой целевой функции в гильбертовом пространстве для выполнимости равенства (9) никаких дополнительных условий на функцию значений (S-функцию) накладывать не требуется, и оно выполняется для всех $(p, r) \in \beta^0$. При этом функция значений $\beta^0$ может не иметь вообще ни одной точки непрерывности в $\beta^0$ (см. пример 2). В то же время можно заметить, что равенство (9) является следствием несимметричной теоремы о минимаксе [22] (см. гл.6, теорему 2.7).

Обозначим далее через $Q$ множество всех значений параметра $(p, r) \in H \times R^m$, для которых задача двойственная к невозмущенной задаче $(P_{p,r}^0)$ разрешима, то есть имеет место равенство

$$\max_{(\lambda,\mu) \in H \times R_+^m} \min_{z \in \mathcal{D}} L_{p,r}^0(z, \lambda, \mu) = \inf_{z \in \mathcal{D}} \sup_{(\lambda,\mu) \in H \times R_+^m} L_{p,r}^0(z, \lambda, \mu). \tag{10}$$

Можно утверждать (подробности и доказательство см. в [13], [14]), что множество $Q$ совпадает со множеством всех тех $(p,r) \in \beta^0 \subset H \times R^m$, для которых не пуст субдифференциал (в смысле выпуклого анализа) $\partial\beta^0(p,r)$.

Таким образом, для выполнимости более «сильного» равенства (10) или, другими словами, существования седловой точки функции Лагранжа $L^0_{p,r}(z,\lambda,\mu)$, $(z,\lambda,\mu) \in \mathcal{D}\times H\times R^m_+$, необходима и достаточна лишь непустота субдифференциала $\partial\beta^0(p,r)$. Так как функционал $\beta^0 : H \times R^m \to$ $\to R^1 \cup \{+\infty\}$ выпуклый полунепрерывный снизу (см. лемму 1 ), множество $Q$ плотно в $\beta^0$ [23] (см. теорему 4.3).

Для иллюстрации рассмотрим два важных частных случая задачи $(P^0_{p,r})$.

**Параметрическая задача с операторным ограничением типа равенства:**

$$(P^0_p) \qquad f^0(z) \to \min,\ \ A^0 z = h^0 + p,\ \ z \in \mathcal{D} \subset Z.$$

В этом случае в соответствии с полученными выше результатами имеем

$$V^0_p(\lambda) \equiv \min_{z\in\mathcal{D}} L^0_p(z,\lambda) = V^0_0(\lambda) - \langle\lambda,p\rangle,\quad \sup_{\lambda\in H} V^0_p(\lambda) = \beta^0(p)$$

и, значит,

$$\beta^0(p) = \sup_{\lambda\in H}(\langle\lambda,-p\rangle - (-V^0_0)(\lambda)) = (-V^0_0)^*(-p),$$

где $(-V^0_0)^*$ функция, сопряженная (по Фенхелю) к функции $-V^0_0$.

Нас интересует разрешимость двойственной задачи, которая эквивалентна принадлежности точки $p$ множеству $Q$ всех тех точек $p \in \beta^0$, для которых выполняется соответствующее этому случаю равенство (10). Это множество, как доказано выше, совпадает со множеством всех тех $p$, для которых не пуст субдифференциал (в смысле выпуклого анализа) $\partial\beta^0(p)$, что, в свою очередь, равносильно выполнимости в задаче $(P^0_p)$ регулярного принципа Лагранжа. Именно в этом и только в этом случае, то есть при $-\lambda \in \partial\beta^0(p)$, имеет место равенство $\partial V^0_p(\lambda) = 0$ или $-p = \partial(-V^0_0)(\lambda)$, которое, в свою очередь, выполняется в силу классического результата из выпуклого анализа (см., например, [23, стр.65, следствие 4.1]) тогда и только тогда, когда имеет место включение $\lambda \in \partial(-V^0_0)^*(-p)$. Другими словами, в рассматриваемом частном случае мы получили дополнительный критерий разрешимости двойственной задачи, выраженный в терминах функции, сопряженной к функции значений, взятой с противоположным знаком.

Заметим, что результат, согласно которому включение $-\lambda \in \partial\beta^0(p)$ равносильно тому, что $\lambda$ является решением двойственной задачи $V^0_0(\lambda) \to \max,\quad \lambda \in H$ применительно к задаче $(P^0_p)$ в случае сильно выпуклой целевой функции уточняет аналогичный результат из [23] (см. стр.81, следствие 5.2), в соответствии с которым для указанной выше равносильности требуется еще и выполнение условия $p \in A^0 f^0$. Отметим, что последнее условие существования внутренней точки $p$ является весьма сильным и не выполняется заведомо, если, например, оператор $A^0$ вполне непрерывен, а $H$ – бесконечномерное гильбертово пространство.

**Параметрическая задача с ограничениями типа неравенства:**

$$(P^0_r) \qquad f^0(z) \to \min,\ \ g^0_i(z) \leqslant r,\ \ i = 1,\ldots,m,\quad z \in \mathcal{D} \subset Z.$$

В этом случае, как легко видеть, функция $\beta^0 : R^m \to R^1 \cup \{+\infty\}$ является функцией монотонно убывающей по каждой переменной отдельно. Хорошо известно, что такие функции являются дифференцируемыми в обычном смысле при почти всех по мере Лебега $r \in \beta^0$ [24]. Так как к тому же функция $\beta^0$ и выпуклая, то все такие точки являются ее точками субдифференцируемости и одновременно – точками, для которых разрешима соответствующая двойственая задача. Таким образом, можно утверждать, что в случае задачи $(P^0_r)$ свойство разрешимости двойственной задачи является свойством общего положения.

### 1.5. Двойственная регуляризация и принцип Лагранжа

Покажем далее, как описанный выше алгоритм двойственной регуляризации одновременно с доказательством его сходимости позволяет вывести и необходимые условия оптимальности элемента $z^0_{p,r}$ в случае, когда субдифференциал $\partial\beta^0(p,r)$ не пуст. Положим $\delta = 0$, $(\lambda^{0,\alpha}_{p,r}, \mu^{0,\alpha}_{p,r}) \equiv (\lambda^\alpha_{p,r}, \mu^\alpha_{p,r})$.

Предположим для упрощения изложения, что функционалы $f^0$, $g^0_j$ являются еще и непрерывно дифференцируемыми по Фреше. С одной стороны, так как функционал $V^0_{p,r}(\lambda,\mu) - \alpha\|\lambda\|^2 - \alpha|\mu|^2$ достигает максимума в точке $(\lambda^\alpha_{p,r}, \mu^\alpha_{p,r})$, то выполняется неравенство

$$\langle (A^0 z^0[\lambda^\alpha_{p,r}, \mu^\alpha_{p,r}] - h^0, g^0(z^0[\lambda^\alpha_{p,r}, \mu^\alpha_{p,r}])) - 2\alpha(\lambda^\alpha_{p,r}, \mu^\alpha_{p,r}), (\lambda,\mu) - (\lambda^\alpha_{p,r}, \mu^\alpha_{p,r})\rangle \leqslant 0 \,\forall\, (\lambda,\mu) \in H \times R^m_+.$$

Из этого неравенства непосредственно вытекают соотношения

$$A^0 z^0[\lambda^\alpha_{p,r}, \mu^\alpha_{p,r}] - h^0 = 2\alpha(\delta)\lambda^\alpha_{p,r}, \quad \langle g^0(z^0[\lambda^\alpha_{p,r}, \mu^\alpha_{p,r}]) - 2\alpha(\delta)\mu^\alpha_{p,r}, \mu - \mu^\alpha_{p,r}\rangle \leqslant 0 \;\forall \mu \in R^m_+. \tag{11}$$

Из (11) следует, что если $\mu^\alpha_{p,r,j} > 0$ для некоторого $j \in \{1,\ldots,m\}$, то

$$g^0_j(z^0[\lambda^\alpha_{p,r}, \mu^\alpha_{p,r}]) - 2\alpha\mu^\alpha_{p,r,j} = 0, \quad g^0_j(z^0[\lambda^\alpha_{p,r}, \mu^\alpha_{p,r}])\mu^\alpha_{p,r,j} > 0. \tag{12}$$

С другой стороны, так как элемент $z^0[\lambda^\alpha_{p,r}, \mu^\alpha_{p,r}]$ доставляет минимальное значение функционалу $L^0(z, \lambda^\alpha_{p,r}, \mu^\alpha_{p,r})$, $z \in \mathcal{D}$, то

$$\langle \nabla^* f^0(z^0[\lambda^\alpha_{p,r}, \mu^\alpha_{p,r}]) + A^{0*}\lambda^\alpha_{p,r} + \nabla^* g^0(z^0[\lambda^\alpha_{p,r}, \mu^\alpha_{p,r}])\mu^\alpha_{p,r}, z - z^0[\lambda^\alpha_{p,r}, \mu^\alpha_{p,r}]\rangle \geqslant 0 \;\forall\, z \in \mathcal{D}. \tag{13}$$

Так как в этом случае функционал $V^0_{p,r}$ достигает максимума, имеем предельные соотношения

$$(\lambda^\alpha_{p,r}, \mu^\alpha_{p,r}) \to (\lambda_{p,r}, \mu_{p,r}), \quad z^0[(\lambda^\alpha_{p,r}, \mu^\alpha_{p,r})] \to z^0_{p,r}, \quad \alpha \to 0,$$

где $(\lambda_{p,r}, \mu_{p,r})$ – минимальный по норме элемент, максимизирующий функционал $V^0_{p,r}$ на $H \times R^m_+$. Поэтому, переходя к пределу в (13) при $\alpha \to 0$, получаем

$$\langle \nabla^* f^0(z^0_{p,r}) + A^{0*}\lambda_{p,r} + \nabla^* g^0(z^0_{p,r})\mu_{p,r}, z - z^0_{p,r}\rangle \geqslant 0 \;\forall\, z \in \mathcal{D}. \tag{14}$$

Одновременно, так как в силу (12) в случае $\mu^\alpha_{p.r.j} > 0$ с необходимостью выполняется неравенство $g^0_j(z^0[\lambda^\alpha_{p,r}, \mu^\alpha_{p,r}]) > 0$, то, если $g^0_j(z^0_{p,r}) < 0$, тогда при достаточно малых $\alpha$ имеем $g^0_j(z^0[\lambda^\alpha_{p,r}, \mu^\alpha_{p,r}]) < 0$ и, значит, с необходимостью при таких $j$ и при тех же $\alpha$ имеем $\mu^\alpha_{p,r,j} = 0$ и, как следствие, $\mu_j = 0$. Последними рассуждениями показано, что $\mu_{p,r,j} g^0_j(z^0_{p,r}) = = 0$, $j = 1,\ldots,m$. Таким образом, алгоритм двойственной регуляризации естественным путем привел и к классическим необходимым условиям – принципу Лагранжа в задаче $(P^0_{p,r})$ в случае, когда субдифференциал $\partial\beta^0(p,q)$ не пуст или, другими словами, когда функционал $V^0_{p,r}$ достигает максимума. При этом можно показать, что в качестве множителей Лагранжа $(\lambda_{p,r}, \mu_{p,r})$ в этом принципе Лагранжа может быть взята любая пара множителей $(\lambda,\mu)$, на которой достигает максимума функционал значений $V^0_{p,r}$ или, другими словами, любой элемент $(\lambda,\mu) \in \partial\beta^0(p,r)$.

Рассмотрим далее случай, когда субдифференциал $\partial\beta^0(p,r)$ пуст, а, в то же время, сингулярный субдифференциал $\partial^\infty\beta^0(p,r)$ не пуст. В этом случае воспользуемся известным представлением (подробности в [13, 14])

$$\partial^\infty\beta^0(p,r) = \limsup_{(p',r') \xrightarrow{\beta^0} (p,r),\, t\downarrow 0} t\partial\beta^0(p',r') \equiv$$

$$\{w - \lim_{k\to\infty} t_k\zeta_k : t_k \downarrow 0,\ \zeta_k \in \partial\beta^0(p^k, r^k),\ (p^k, r^k) \xrightarrow{\beta^0} (p, r)\},$$

где $(p', r') \xrightarrow{\beta^0} (p, r)$ означает, что $((p', r'), \beta^0(p', r')) \to ((p, r), \beta^0(p, r))$, а $t \downarrow 0$ означает сходимость к нулю справа, и умножим неравенство (14) на $t > 0$

$$\langle t\nabla^* f^0(z^0_{p,r}) + A^{0*} t\lambda_{p,r} + \nabla^* g^0(z^0_{p,r}) t\mu_{p,r}, z - z^0_{p,r}\rangle \geqslant 0\ \forall z \in \mathcal{D}. \tag{15}$$

Тогда для любой слабой предельной точки вида

$$(\tilde\lambda_{p,r}, \tilde\mu_{p,r}) = w - \lim_{k\to\infty,\,(p^k, r^k)\xrightarrow{\beta^0}(p,r),\,t_k\downarrow 0} t_k(\lambda^k_{p^k,r^k}, \mu^k_{p^k,r^k})$$

с $(\lambda^k_{p^k,r^k}, \mu^k_{p^k,r^k}) \in \partial\beta^0(p^k, r^k)$ можем записать после очевидного предельного перехода в (15) при $(p, r) = (p^k, r^k)$, $(\lambda_{p,r}, \mu_{p,r}) = (\lambda^k_{p^k,r^k}, \mu^k_{p^k,r^k})$, $t = t_k$

$$\langle A^{0*}\tilde\lambda_{p,r} + \nabla^* g^0(z^0_{p,r})\tilde\mu_{p,r}, z - z^0_{p,r}\rangle \geqslant 0\ \forall z \in \mathcal{D}. \tag{16}$$

При этом в силу условия дополняющей нежесткости $\mu_{p^k,r^k,j} g^0_j(z^0_{p^k,r^k}) = 0,\ j = 1, \ldots, m$ в результате предельного перехода при $k \to \infty$ получаем $\tilde\mu_{p,r,j} g^0_j(z^0_{p,r}) = 0,\ j = 1, \ldots, m$, что в совокупности с (16) и означает выполнимость нерегулярного принципа Лагранжа.

Суммируя полученные выше в этом разделе результаты, можем сформулировать следующую теорему.

Т е о р е м а 3. [*Регуляризованный принцип Лагранжа*] *Если субдифференциал $\partial\beta^0(p, r)$ не пуст, то найдется пара двойственных переменных $(\lambda, \mu) \in H \times R^m_+$ такая, что*

$$(\lambda, \mu) \in \{V^0_{p,r}(\lambda', \mu') : (\lambda', \mu') \in H \times R^m_+\},$$

*а решение $z^0_{p,r}$ задачи $(P^0_{p,r})$ доставляет минимальное значение функционалу Лагранжа $L^0_{p,r}(z, \lambda, \mu)$, $z \in \mathcal{D}$ и удовлетворяет, естественно, регулярному принципу Лагранжа в дифференциальной форме теоремы 1.*

*Если субдифференциал $\partial\beta^0(p, r)$ пуст, мы можем в общей ситуации лишь утверждать, что найдется последовательность $(\lambda^k, \mu^k)$, $k = 1, 2, \ldots$ двойственных переменных такая, что*

$$V^0_{p,r}(\lambda^k, \mu^k) \to \beta^0(p, r) = \sup_{(\lambda', \mu') \in H \times R^m_+} V^0_{p,r}(\lambda', \mu'),\ k \to \infty,$$

*а точки $z^k$, $k = 1, 2, \ldots$, доставляющие минимум функционалу Лагранжа $L^0_{p,r}(z, \lambda^k, \mu^k)$, $z \in \mathcal{D}$, сходятся при $k \to \infty$ к решению $z^0_{p,r}$ задачи $(P^0_{p,r})$. При этом элемент $z^0_{p,r}$ в совокупности с любой слабой предельной точкой, возможно вырожденной, $(0, \lambda, \mu)$ нормированной последовательности*

$$\xi^k \equiv (1/||(\lambda^k, \mu^k)||, \lambda^k/\|(\lambda^k, \mu^k)\|, \mu^k/\|(\lambda^k, \mu^k)\|)$$

*удовлетворяет соотношениям нерегулярного принципа Лагранжа в дифференциальной форме*

$$\langle A^{0*}\lambda + \nabla^* g^0(z^0_{p,r})\mu, z - z^0_{p,r}\rangle \geqslant 0\ \forall z \in \mathcal{D},\quad \mu_j g^0_j(z^0_{p,r}) = 0,\ j = 1, \ldots, m.$$

*В последней ситуации результат может быть уточнен если мы имеем информацию о компактности единичной сферы пространства двойственных переменных.*

*Если субдифференциал $\partial\beta^0(p, r)$ пуст, то в случае компактности единичной сферы двойственного пространства элемент $z^0_{p,r}$ в совокупности с любой предельной точкой $(0, \lambda, \mu)$ (она* *в данном случае является невырожденной) нормированной последовательности $\xi_k$ удовлетворяет соотношениям невырожденного принципа Лагранжа в дифференциальной форме*

$$\langle A^{0*}\lambda + \nabla^* g^0(z^0_{p,r})\mu, z - z^0_{p,r}\rangle \geqslant 0 \ \forall\, z \in \mathcal{D}, \quad \mu_j g^0_j(z^0_{p,r}) = 0, \ j = 1, \ldots, m.$$

*Если же в случае пустоты субдифференциала $\partial\beta^0(p,r)$ мы не имеем информации о компактности единичной сферы пространства двойственных переменных, но сингулярный (асимптотический) субдифференциал $\partial^\infty\beta^0(p,r)$ состоит не из одного нуля, то найдутся последовательность $(p^k, r^k)$, $k = 1, 2, \ldots$, $(p^k, r^k) \to (p, r)$, $k \to \infty$ и последовательность двойственных переменных $(\lambda^k, \mu^k)$, $k = 1, 2, \ldots$, такие, что $V^0_{p^k,r^k}(\lambda^k, \mu^k) \to \beta^0(p, r)$, $k \to \infty$, а минимизирующие функционал Лагранжа $L^0_{p^k,r^k}(z, \lambda^k, \mu^k)$, $z \in \mathcal{D}$ точки $z^k$, $k = 1, 2, \ldots$ сходятся при $k \to \infty$ к решению $z^0_{p,r}$ задачи $(P^0_{p,r})$. При этом элемент $z^0_{p,r}$ в совокупности с любой существующей в этом случае ненулевой слабой предельной точкой $(\lambda, \mu)$ последовательности $t_k(\lambda^k, \mu^k)$, $t_k \downarrow 0$, $k \to \infty$ удовлетворяет соотношениям невырожденного принципа Лагранжа в дифференциальной форме*

$$\langle A^{0*}\lambda + \nabla^* g^0(z^0_{p,r})\mu, z - z^0_{p,r}\rangle \geqslant 0 \ \forall\, z \in \mathcal{D}, \quad \mu_j g^0_j(z^0_{p,r}) = 0, \ j = 1, \ldots, m.$$

Рассмотрим следующий простой иллюстративный пример.

П р и м е р 5. Пусть в задаче $(P_p)$ $\mathcal{D} = Z$, $h \in R(A)$, оператор $A$ самосопряжен и $R(A) \neq \overline{R(A)}$ (например, $A$ может быть интегральным оператором Фредгольма с симметрическим ядром). Тогда классический принцип Лагранжа для гладких задач с равенствами (см., например, [21], стр.253) в этом случае не применим ни при одном $p \in R(A)$. В то же время, в соответствии со сказанным выше, в этом случае при любом $p \in R(A)$ найдется такая последовательность двойственной переменной $\lambda^k$, $k = 1, 2, \ldots$, что $V_p(\lambda^k) \to \beta(p)$, а точки $z^k$, $k = 1, 2, \ldots$, минимизирующие функционал Лагранжа $L_p(z, \lambda^k)$, $z \in Z$, сходятся при $k \to \infty$ к решению $z^0_p$ задачи $(P_p)$.

Так как выпуклый функционал $\beta$ полунепрерывен снизу, то для плотного в $\beta$ множества точек $p$ субдифференциал $\partial\beta(p)$ не пуст [23] и, как следствие, для таких $p$ выполняется регулярный принцип Лагранжа. Таким образом, в случае приведенного в предыдущем абзаце примера мы сталкиваемся с ситуацией, когда классический принцип Лагранжа для гладких задач с равенствами не применим ни при одном $p \in R(A)$, тогда как непустота субдифференциала $\partial\beta(p)$ для плотного в $\beta$ множества обеспечивает выполнимость для всех таких $p \in \beta$ регулярного принципа Лагранжа в задаче $(P_p)$.

Можно рассуждать и несколько иначе. Если, например, функционал $f(z) = \|z\|^2$, а оператор $A$, к тому же, и взаимнооднозначный (например, интегральный оператор Фредгольма с симметрическим замкнутым ядром), что и имеет место в случае задачи примера в замечании 1 (см. также пример 2 в [4], то согласно классическим свойствам разрешимости операторных уравнений, сопряженное уравнение $A\lambda = g$ плотно разрешимо в $H$. Значит, элементы вида $A\lambda$, $\lambda \in H$ лежат всюду плотно в $Z$, то есть для всюду плотного множества точек $p$ из $R(A)$ выполняются одновременно равенства $Az = h + p$, $z = A\lambda$ или равенства $Az = h + p$, $z = -1/2A\lambda$. Так как целевая двойственная функция $V_p(\lambda)$, $\lambda \in H$ в этом случае имеет обычный градиент $\partial V_p(\lambda) = Az[\lambda] - h - p$, где $z[\lambda] = -1/2A\lambda$, то в силу сказанного выше этот градиент зануляется для плотного в $\beta$ множества точек $p$, то есть для всех таких $p$ двойственная задача $V_p(\lambda) \to \sup$, $\lambda \in H$ разрешима, что означает непустоту соответствующего субдифференциала $\partial\beta(p)$ и, как следствие, выполнимость соотношений регулярного принципа Лагранжа. Подчеркнем, что в последнем конкретном примере мы сталкиваемся с ситуацией, когда классический принцип Лагранжа для гладких задач с равенствами не применим ни при одном $p \in R(A)$, тогда как непустота субдифференциала $\partial\beta(p)$ для плотного в $\beta$ множества обеспечивает выполнимость для всех таких $p \in \beta$ регулярного принципа Лагранжа в задаче $(P_p)$. Одновременно, плотно в $\beta$ (см. пример 2, а также анализ примера 2 в [4]) лежат точки $p$ такие, для которых субдифференциал $\partial\beta(p)$ пуст, а асимптотический субдифференциал $\partial^\infty\beta(p)$ состоит из одного нуля (те точки, для которых не выполняется принцип Лагранжа). Они «не охватываются» классическим принципом Лагранжа, но, в соответствии со сказанным выше, для таких точек найдется такая последовательность двойственной переменной $\lambda^k,\ k = 1, 2, \dots$, что $V_p(\lambda^k) \to \beta(p)$, а точки $z^k,\ k = 1, 2, \dots$, минимизирующие функционал Лагранжа $L_p(z, \lambda^k),\ z \in Z$, сходятся при $k \to \infty$ к решению $z_p$ задачи $(P_p)$.

Если в задаче $R(A) = \overline{R(A)}$, но $R(A) \neq Z$, то это означает, что в любой точке $p \in \beta = R(A)$ асимптотический субдифференциал $\partial^\infty\beta(p)$ состоит не из одного нуля (в качестве его ненулевого элемента можно взять любой элемент $r \neq 0$ из ортогонального дополнения пространства $R(A)$). Условие $R(A) = \overline{R(A)}$ является в этом случае достаточным условием выполнимости невырожденного принципа Лагранжа как в соответствии с классическим принципом Лагранжа для гладких задач с равенствами, так и в соответствии со сказанным выше.

### 1.6. Параметрическая итеративная двойственная регуляризация, правило останова итерационного процесса

Конечно, точное решение регуляризованной двойственной задачи при каждом фиксированном значении параметра регуляризации $\alpha$ на множестве $H \times R^m_+ \equiv \Lambda$, которое предполагается в выше приведенных рассуждениях, практически не реализуемо.

Оказывается, однако, что для практического решения исходной задачи $(P^0_{p,r})$ как в случае сильно выпуклого, так и выпуклого функционала цели может быть непосредственно использована процедура итеративной регуляризации двойственного алгоритма. Такие процедуры в теории некорректных задач и в задачах оптимизации можно найти в книгах [1, 25]. Приведем здесь лишь теорему сходимости процесса итеративной двойственной регуляризации в случае ограниченного множества $\mathcal{D}$.

Отметим здесь, что описанный выше двойственный алгоритм естественно в случае $\delta = 0$ рассматривать как базовый метод. Последовательность, вырабатываемая методом итеративной регуляризации, которому посвящен данный раздел, будет стабилизироваться к последовательности, вырабатываемой этим базовым методом.

Можно показать (см., например, [4, 14]), что в случае ограниченного множества $\mathcal{D}$ справедлива оценка

$$|\partial V^\delta_{p,r}(\lambda, \mu) - \partial V^0_{p,r}(\lambda, \mu)| \leqslant \tau(\delta)(1 + \|(\lambda, \mu)\|),$$

где $\tau(\delta) \geqslant 0,\ \tau(\delta) \to 0,\ \delta \to 0$, причем $\tau(\delta) \equiv C\sqrt{\delta}$ в случае линейно-выпуклой задачи $(P^0_{p,r})$ с сильно выпуклым функционалом $f^0$.

Пусть последовательность $(\bar\lambda^k, \bar\mu^k),\ k = 1, 2, \dots$, конструируется по правилу

$$\begin{gathered}(\bar\lambda^{k+1}, \bar\mu^{k+1}) = Pr_\Lambda((\bar\lambda^k, \bar\mu^k) + \beta^k \partial V^{\delta^k}_{p,r}(\bar\lambda^k, \bar\mu^k) - 2\beta^k\alpha^k(\bar\lambda^k, \bar\mu^k)),\\ k = 1, 2, \dots;\ (\bar\lambda^1, \bar\mu^1) \in H \times R^m_+,\end{gathered} \tag{17}$$

где $\Lambda \equiv H \times R^m_+$, а последовательности $\tau^k \equiv \tau(\delta^k),\ \alpha^k,\ \beta^k,\ k = 1, 2, \dots$, удовлетворяют условиям

$$\tau^k \geqslant 0,\ \alpha^k > 0,\ \beta^k > 0,\ \lim_{k\to\infty}(\tau^k + \alpha^k + \beta^k) = 0,\ \frac{\alpha^k}{\alpha^{k+1}} \leqslant C_0, \tag{18}$$

$$\frac{|\alpha^{k+1} - \alpha^k|}{(\alpha^k)^3\beta^k} \leqslant C,\ \frac{\beta^k}{(\alpha^k)^3} \leqslant C,\ \frac{\tau^k}{(\alpha^k)^3} \leqslant C,\ \sum_{k=1}^\infty \alpha^k\beta^k = +\infty.$$

Можно заметить, что условие согласования $\lim\limits_{k\to\infty} \tau^k/(\alpha^k)^3 = 0$ в случае сильно выпуклого целевого функционала имеет вид $\lim\limits_{k\to\infty} \delta^k/(\alpha^k)^6 = 0$.

Легко проверить, что последовательности $\alpha^k$, $\beta^k$, $k=1,2,\ldots$, удовлетворяющие соотношениям (18), существуют. В качестве одного из возможных примеров таких последовательностей можно взять, в частности, $\alpha^k = k^{-1/6}$, $\beta^k = k^{-1/(5/3)}$.

Введем обозначение: $(\lambda^{0,\alpha^k}, \mu^{0,\alpha^k}) \equiv (\lambda^{\alpha^k}, \mu^{\alpha^k}) \equiv (\lambda^k, \mu^k)$. Справедлива следующая теорема, доказательство которой можно найти в [4,9–11, 14].

Т е о р е м а 4. [*Итеративная двойственная регуляризация*] *Пусть $z^0_{p,r}$ – решение исходной оптимизационной задачи $(P^0_{p,r})$ и выполняются условия согласования (18). Тогда для любых реализаций наборов исходных данных $\mathsf{f}^{\delta^k}$, удовлетворяющих соотношениям (5) выполняются соотношения*

$$(\bar\lambda^k, \bar\mu^k) - (\lambda^k, \mu^k) \to 0, \quad f^0(z^{\delta^k}[\bar\lambda^k, \bar\mu^k]) \to f^0(z^0),\ k \to \infty,$$

$$A^0 z^{\delta^k}[\bar\lambda^k, \bar\mu^k] - h^0 \to 0,\ \ g_i(z^{\delta^k}[\bar\lambda^k, \bar\mu^k]) \leqslant \varkappa^k,\ \ \varkappa^k \to 0,\ k \to \infty,$$

*а при дополнительном условии субдифференцируемости $f^0$ (в смысле выпуклого анализа) в точках $\mathcal{D}$ и соотношение*

$$\|z^{\delta^k}[\bar\lambda^k, \bar\mu^k] - z^0_{p,r}\| \to 0,\ k \to \infty.$$

*Другими словами, вне зависимости от того, разрешима или нет двойственная задача, итерационная процедура (17), (18) представляет собой регуляризирующий алгоритм.*

В случае разрешимости двойственной задачи, что равносильно непустоте субдифференциала $\partial\beta^0(p,r)$, двойственные переменные $(\bar\lambda^k, \bar\mu^k)$, $(\lambda^k, \mu^k)$ равномерно по $k$ ограничены. Это естественно приводит к менее жестким условиям согласованного стремления к нулю величин $\tau^k$, $\alpha^k$, $\beta^k$, $k=1,2,\ldots$. А, именно, вторая группа условий в (18) может быть заменена в этом случае на условия

$$\lim_{k\to\infty} \frac{|\alpha^{k+1} - \alpha^k|}{(\alpha^k)^2 \beta^k} \leqslant C,\ \ \lim_{k\to\infty} \frac{\beta^k}{\alpha^k} \leqslant C,\ \ \lim_{k\to\infty} \frac{\tau^k}{\alpha^k} \leqslant C,\ \ \sum_{k=1}^{\infty} \alpha^k \beta^k = +\infty.$$

В качестве одного из возможных примеров таких последовательностей можно взять, в частности, $\alpha_k = k^{-1/3}$, $\beta_k = k^{-1/2}$. В этом случае сформулированная выше теорема остается в силе при указанных менее жестких условиях согласования.

Обсудим далее важное с практической точки зрения правило останова итерационного процесса (17) в случае, когда исходные данные оптимизационной задачи задаются с определенной фиксированной (конечной) погрешностью $\delta > 0$. Будем следовать при этом подходу [1] (см. гл.9, §8, теорема 2). Пусть последовательности $\tau^k$, $\alpha^k$, $\beta^k$, $k=1,2,\ldots$, удовлетворяют условиям (18) или, в случае существования решения двойственной задачи, сформулированным выше менее жестким условиям согласования. Зафиксируем следующее правило останова процесса (17)

$$(\bar\lambda^{k+1}, \bar\mu^{k+1}) = (\bar\lambda^k, \bar\mu^k) + \beta^k \partial V^\delta_{p,r}(\bar\lambda^k, \bar\mu^k) - 2\beta^k \alpha^k (\bar\lambda^k, \bar\mu^k),\ \ k=1,2,\ldots;\ (\bar\lambda^1, \bar\mu^1) \in H \times R^m_+, \quad (19)$$

при фиксированном конечном уровне погрешности $\delta > 0$: при каждом $\delta > 0, \delta \leqslant \delta^1$ итерации продолжаются до такого наибольшего номера $k = k(\delta)$, при котором выполняются неравенства

$$\delta^k \geqslant \delta,\ \ k = 1,2,\ldots,k(\delta).$$

Справедлива следующая (подробности и доказательство см. в [4, 14])

Т е о р е м а 5. [*Правило останова итерационного процесса*] *Вне зависимости от того, разрешима или нет двойственная к $(P^0_{p,r})$ задача, справедливы предельные соотношения*

$$f^0(z^\delta[\bar\lambda^{k(\delta)}, \bar\mu^{k(\delta)}]) \to f^0(z^0),\ \ A^0 z^\delta[\bar\lambda^{k(\delta)}, \bar\mu^{k(\delta)}] - h^0 \to 0,$$

$$g_i^{0+}(z^\delta[\bar\lambda^{k(\delta)}, \bar\mu^{k(\delta)}]) \to 0,\ \ i=1,\ldots,m,\ \delta \to 0,$$

*а в случае субдифференцируемости $f^0$ и предельное соотношение*

$$\|z^\delta[\bar\lambda^{k(\delta)},\bar\mu^{k(\delta)}]-z^0_{p,r}\|\to 0,\ \ \delta\to 0,$$

*где $z^\delta[\bar\lambda^{k(\delta)},\bar\mu^{k(\delta)}]$ – результат $k(\delta)$ итераций итерационного процесса (19). Другими словами, указанное правило останова порождает регуляризирующий алгоритм в задаче $(P^0_{p,r})$.*

## 1.7. Параметрическая двойственная регуляризация в случае выпуклого целевого функционала

Рассмотрим далее регуляризованный двойственный алгоритм в случае, когда функционал $f^0$ задачи $(P^0_{p,r})$, а также его возмущение $f^\delta$ являются только выпуклыми. Считаем выполненными все сделанные ранее при постановке задачи предположения за исключением лишь предположения сильной выпуклости функций $f^\delta$, $\delta\in[0,\delta_0]$. Пусть множество $\mathcal{D}$ ограничено, и будем считать, что задача $(P^0_{p,r})$ разрешима. Пусть опять $z^0_{p,r}$ – нормальное решение этой задачи. Двойственный алгоритм решения задачи $(P^0_{p,r})$ можно организовать на основе рассмотренного выше двойственного алгоритма в случае сильно выпуклого функционала $f^0$, например, следующим образом (подробности и доказательства см. в [4, 14]).

Рассмотрим семейство регуляризованных задач

$$(P^0_{p,r,\varepsilon})\qquad f^0(z)+\varepsilon\|z\|^2\to\min,\ \ A^0z=h^0,\ \ g^0_i(z)\leqslant 0,\ \ i=1,\dots,m,\quad z\in\mathcal{D}\subset Z,$$

с их единственными нормальными решениями $z^0_{p,r,\varepsilon}$, $\varepsilon>0$, $z^0_{p,r,0}\equiv z^0_{p,r}$.

Известно, что в этом случае для любой последовательности положительных сходящихся к нулю чисел $\varepsilon^k$, $k=1,2,\dots$, предел последовательности $z^0_{p,r,\varepsilon^k}$, $k=1,2,\dots$, является нормальным решением исходной задачи $(P^0_{p,r})$.

Определим функцию значений

$$V^{\varepsilon,\delta}_{p,r}(\lambda,\mu)\equiv\inf_{z\in\mathcal{D}}L^{\varepsilon,\delta}_{p,r}(z,\lambda,\mu),\ \ \lambda\in H,\ \ \mu\in R^m,\ \ \delta\geqslant 0,$$

где

$$L^{\varepsilon,\delta}_{p,r}(z,\lambda,\mu)\equiv f^\delta(z)+\varepsilon\|z\|^2+\langle\lambda,A^\delta z-h^\delta-p\rangle+\langle\mu,g^\delta(z)-r\rangle,\ \ z\in\mathcal{D},\ \lambda\in H,\ \ \mu\in R^m.$$

Пусть $z^\delta_\varepsilon[\lambda,\mu]\equiv argmin\,\{L^{\varepsilon,\delta}_{p,r}(z,\lambda,\mu),\ z\in\mathcal{D}\}$ при $(\lambda,\mu)\in H\times R^m_+$, $\varepsilon>0$, $\delta\geqslant 0$.

Зададимся произвольной последовательностью положительных сходящихся к нулю чисел $\varepsilon^s$, $s=1,2,\dots$ и будем при каждом $s$ решать регуляризованную задачу $(P^0_{p,r,\varepsilon^s})$ с помощью алгоритма итеративной регуляризации в случае сильно выпуклого целевого функционала.

Тогда в силу теоремы 4 при выполнении условия согласования (18) для любых реализаций наборов исходных данных $\mathsf{f}^{\delta^k}$, удовлетворяющих соотношениям (5), выполняются соотношения

$$f^0(z^{\delta^k}_{\varepsilon^s}[\bar\lambda^k_s,\bar\mu^k_s])+\varepsilon^s\|z^{\delta^k}_{\varepsilon^s}[\bar\lambda^k_s,\bar\mu^k_s]\|^2\to f^0(z^0_{p,r,\varepsilon^s})+\varepsilon^s\|z^0_{p,r,\varepsilon^s}\|^2,\ \ k\to\infty,$$

$$A^0z^{\delta^k}_{\varepsilon^s}[\bar\lambda^k_s,\bar\mu^k_s]-h^0\to 0,\ \ g^0_i(z^{\delta^k}_{\varepsilon^s}[\bar\lambda^k_s,\bar\mu^k_s])\leqslant\varkappa^k_s,\ \ \varkappa^k_s\to 0,\ k\to\infty,$$

а при дополнительном условии субдифференцируемости $f^0$ (в смысле выпуклого анализа) в точках $\mathcal{D}$ и соотношение

$$\|z^{\delta^k}_{\varepsilon^s}[\bar\lambda^k_s,\bar\mu^k_s]-z^0_{p,r,\varepsilon^s}\|\to 0,\ \ k\to\infty,$$

где через $(\bar\lambda^k_s,\bar\mu^k_s)$, $k=1,2,\dots$ обозначена последовательность двойственных переменных, вырабатываемая при каждом $s=1,2,\dots$ с помощью алгоритма итеративной регуляризации (17).

Предположим, что при каждом $s$ итеративная регуляризация проводится так, что

$$|f^0(z_{\varepsilon^s}^{\delta^{\bar{k}(s)}}[\bar{\lambda}_s^{\bar{k}(s)},\bar{\mu}_s^{\bar{k}(s)}])+\varepsilon^s\|z_{\varepsilon^s}^{\delta^{\bar{k}(s)}}[\bar{\lambda}_s^{\bar{k}(s)},\bar{\mu}_s^{\bar{k}(s)}]\|^2-f^0(z_{p,r,\varepsilon^s}^0)-\varepsilon^s\|z_{p,r,\varepsilon^s}^0\|^2|\leqslant\gamma^s,$$

$$\|A^0z_{\varepsilon^s}^{\delta^{\bar{k}(s)}}[\bar{\lambda}_s^{\bar{k}(s)},\bar{\mu}_s^{\bar{k}(s)}]-h^0\|\leqslant\gamma^s,\ \ g_i^0(z_{\varepsilon^s}^{\delta^{\bar{k}(s)}}[\bar{\lambda}_s^{\bar{k}(s)},\bar{\mu}_s^{\bar{k}(s)}])\leqslant\gamma^s,\ i=1,\ldots,m,$$

а при дополнительном условии субдифференцируемости $f^0$ (в смысле выпуклого анализа) в точках $\mathcal{D}$ и соотношение

$$\|z_{\varepsilon^s}^{\delta^{\bar{k}(s)}}[\bar{\lambda}_s^{\bar{k}(s)},\bar{\mu}_s^{\bar{k}(s)}]-z_{p,r,\varepsilon^s}^0\|\leqslant\gamma^s,$$

где $\gamma^s$, $s=1,2,\ldots,$ – произвольная фиксированная последовательность сходящихся к нулю положительных чисел, $\bar{k}(s)$ – тот наименьший номер, при котором выполняются эти соотношения, такой, что $\bar{k}(s+1)>\bar{k}(s)$.

Т е о р е м а 6. *Пусть $z_{p,r}^0$ – решение исходной оптимизационной задачи $(P_{p,r}^0)$ с выпуклым функционалом цели, и выполняются условия согласования (18). Тогда для любых реализаций наборов исходных данных $\mathsf{f}^{\delta^k}$, удовлетворяющих соотношениям (5), выполняются соотношения*

$$f^0(z_{\varepsilon^s}^{\delta^{\bar{k}(s)}}[\bar{\lambda}_s^{\bar{k}(s)},\bar{\mu}_s^{\bar{k}(s)}])\to f^0(z_{p,r}^0),\ \ s\to\infty,$$

$$A^0z_{\varepsilon^s}^{\delta^{\bar{k}(s)}}[\bar{\lambda}_s^{\bar{k}(s)},\bar{\mu}_s^{\bar{k}(s)}]-h^0\to 0,\ \ g_i(z_{\varepsilon^s}^{\delta^{\bar{k}(s)}}[\bar{\lambda}_s^{\bar{k}(s)},\bar{\mu}_s^{\bar{k}(s)}])\leqslant\gamma^s,\ \ \gamma^s\to 0,\ s\to\infty,$$

*а при дополнительном условии субдифференцируемости $f^0$ (в смысле выпуклого анализа) при $z\in\in\mathcal{D}$ точки $z^{\delta^{\bar{k}(s)}}[\bar{\lambda}_s^{\bar{k}(s)},\bar{\mu}_s^{\bar{k}(s)}],\ s=1,2,\ldots$ сильно в $Z$ сходятся к нормальному решению исходной задачи $(P_{p,r}^0)$.*

## 1.8. Параметрическая двойственная регуляризация в линейно выпуклой задаче оптимального управления с поточечными фазовыми ограничениями типа равенства и неравенства

Рассмотрим параметрическую линейно выпуклую задачу оптимального управления с поточечными фазовыми ограничениями типа равенства и неравенства, понимаемыми как ограничения в гильбертовом пространстве $\mathcal{H}\equiv L_2(X)$ [26], [27]

$$(P_{p,r}^{OC})\qquad g_0(u)\equiv\int_0^T(\langle F(t)x[u](t),x[u](t)\rangle+\langle G(t)u(t),u(t)\rangle)dt\to\min,\quad u\in\mathcal{D}\subset L_2(0,T),$$

$$g_1(u)(t)\equiv\langle\varphi_1(t),x[u](t)\rangle=h(t)+p(t),\ \ g_2(u)(t)\equiv\varphi_2(t,x[u](t))\leqslant r(t)\ \ \text{при п.в. } t\in X.$$

Здесь: $p\in\mathcal{H}$, $r\in\mathcal{H}$ – параметры, $g_0:\ L_2(0,T)\to R^1$ – сильно выпуклый функционал с постоянной $\kappa$, $F$, $A:\ [0,T]\to R^{n\times n}$, $B:\ [0,T]\to R^{n\times m}$, $G:\ [0,T]\to R^{m\times m}$ – измеримые по Лебегу ограниченные матрицы, $\varphi_1$, $h\in L_\infty(0,T)$ – заданные функции, $\varphi_2:\ [0,T]\times R^n\to R^1$ – выпуклая по $x$, непрерывная вместе с градиентом $\nabla_x\varphi_2$ функция, $\mathcal{D}\equiv\{u\in L_2(0,T):\ u(t)\in U$ при п.в. $t\in(0,T)\}$, $U\subset R^m$ – выпуклый компакт, $X\subset[0,T]$, $X=X$, $x[u](t),\ t\in[0,T]$ – решение задачи Коши

$$\dot{x}=A(t)x+B(t)u(t),\ \ x(0)=x_0\in R^n,\quad t\in[0,T].$$

Обозначим единственное решение задачи $(P_{p,r}^{OC})$, если оно существует, через $u_{p,r}^0$.

С формальной точки зрения задача $(P_{p,r}^{OC})$ [26], [27] принципиально отличается от рассмотренной в предыдущих разделах задачи математического программирования $(P_{p,r})$, так как множества $\mathcal{H}_-\equiv\{z\in L_2(X):\ z(t)\leqslant 0$ при п.в. $t\in X\}$, $\mathcal{H}_+\equiv\{z\in L_2(X):\ z(t)\geqslant 0$ при п.в. $t\in X\}$ не имеют внутренних точек в топологии несущего пространства $\mathcal{H}$.

Тем не менее, описанный выше процесс двойственной регуляризации для задачи $(P_{p,r})$ полностью сохраняет свою силу и в случае задачи $(P_{p,r}^{OC})$. Он также заключается в непосредственном решении двойственной к $(P^{OC}_{p,r})$ и регуляризованной по Тихонову задачи (считаем для простоты ошибку исходных данных равной нулю)

$$R^{\alpha}_{p,r}(\lambda,\mu) \equiv V_{p,r}(\lambda,\mu) - \alpha\|(\lambda,\mu)\|^2 \to \sup, \quad (\lambda,\mu) \in \mathcal{H}\times\mathcal{H}_+, \tag{20}$$

$$V_{p,r}(\lambda,\mu) \equiv \min_{u\in\mathcal{D}} L_{p,r}(u,\lambda,\mu), \quad L_{p,r}(u,\lambda,\mu) \equiv g_0(u) + \langle\lambda, g_1(u) - h - p\rangle + \langle\mu, g_2(u) - r\rangle,$$

$$u[\lambda,\mu] \equiv \arg\min\{L_{p,r}(u,\lambda,\mu) : u \in \mathcal{D}\},\ (\lambda^{\alpha}_{p,r}, \mu^{\alpha}_{p,r}) \equiv \arg\max\{R^{\alpha}_{p,r}(\lambda,\mu) : (\lambda,\mu) \in \mathcal{H}\times\mathcal{H}_+\},$$

в результате чего происходит аппроксимация в метрике $L_2(0,T)$ при $\alpha \to 0$ решения $u^0_{p,r}$ элементами $u[\lambda^{\alpha}_{p,r}, \mu^{\alpha}_{p,r}]$. Наличие параметра $(p,r) \in \mathcal{H}\times\mathcal{H}$ в исходной задаче также позволяет исследовать зависимость от него свойств сходимости двойственного алгоритма (20): связь с дифференциальными свойствами функции значений ($S$-функции)

$$\beta(p,r) \equiv \{g_0(u^0_{p,r}), \text{если задача } (P_{p,r}) \text{ разрешима}; +\infty \text{ в противоположном случае}\},$$

разрешимостью двойственной задачи, принципом Лагранжа и принципом максимума Понтрягина.

Подчеркнем, что важнейшее преимущество рассмотрения ограничений задачи $(P^{OC}_{p,r})$, как ограничений в $L_2(X)$, заключается, прежде всего, в том, что это приводит к устойчивому к ошибкам исходных данных алгоритму ее решения. В то же время, при определенных условиях на исходные данные эти ограничения можно, естественно, трактовать и как ограничения в $L_\infty(X)$ ($p,\ r \in L_\infty(X)$ ) и $C(X)$ ($\varphi_1,\ h,\ p,\ r \in C(X)$). При этом понятия оптимальности управления в указанных частных случаях эквивалентны понятию оптимальности для случая, когда «те же» ограничения рассматриваются в $L_2(X)$.

Как и в предыдущих разделах, алгоритм (20) для решения задачи $(P^{OC}_{p,r})$ ведет себя двояко в зависимости от того, разрешима или нет двойственная к $(P^{OC}_{p,r})$ задача. В первом случае семейство двойственных переменных $(\lambda^{\alpha}_{p,r}, \mu^{\alpha}_{p,r})$, $\alpha \to 0$ ограничено в $\mathcal{H}\times\mathcal{H}$, во втором – нет. Однако в обоих случаях «параллельно» с построением минимизирующего приближенного решения в смысле Дж. Варги алгоритм (20) ведет и к соответствующим необходимым условиям – так называемому регуляризованному принципу Лагранжа в недифференциальной форме в задаче $(P^{OC}_{p,r})$ и, соответственно, регуляризованному принципу максимума Понтрягина. Главное отличие последних от их классических аналогов заключается в том, что, во-первых, они записываются в терминах минимизирующих последовательностей (а не оптимальных управлений), а, во-вторых, выполняются в любой задаче $(P^{OC}_{p,r})$, имеющей решение (как известно, эти классические аналоги в задачах $(P^{OC}_{p,r})$ с ограничениями в $L_2(X)$ могут не быть справедливыми).

Указанные выше регуляризованные принцип Лагранжа и принцип максимума Понтрягина в задаче $(P^{OC}_{p,r})$ оказываются полезными для решения задач оптимизации, оптимального управления, а также некорректных обратных задач. Подчеркнем также, что если в задаче отсутствует ограничение-равенство ($(P^{OC}_{p,r}){=}(P^{OC}_{r})$), а ограничение-неравенство понимается как ограничение в $C(X)$, $r \in C(X)$, то регуляризованный принцип максимума после некоторой естественной перенормировки множителя $\mu^{\alpha}_{r}$, в независимости от того, разрешима или нет двойственная задача, приводит к невырожденному классическому принципу максимума Понтрягина для управления $u^0_r$, в записи которого участвует традиционная в таком важном частном случае мера Радона. Все необходимые подробности, доказательства, иллюстративные примеры см. в [27].

## 2. Параметрическая двойственная регуляризация в нелинейной задаче математического программирования

Идеология двойственной регуляризации естественным образом распространяется и на нелинейные задачи [28]. При этом, в отличие от линейно выпуклого случая, где центральной конструкцией является функция Лагранжа, в нелинейном случае центральную роль играет модифицированная функция Лагранжа. Естественным образом модифицированные функции Лагранжа возникают именно в нелинейных параметрических задачах математического программирования. В этом случае их конструкции являются следствиями дифференциальных свойств функций значений ($S$-функций) параметрических задач. Приведем краткое изложение результатов работы [29], сосредоточив главное внимание именно на связи дифференциальных свойств функций значений с модифицированными функциями Лагранжа.

### 2.1. Постановка параметрической нелинейной задачи математического программирования

Рассмотрим параметрическую задачу минимизации

$$(P_p) \qquad f(z) \to \inf, \;\; g(z) = p, \;\; z \in \mathcal{D} \subset Z,$$

где $f : \mathcal{D} \to R^1$ – непрерывный функционал, $g : \mathcal{D} \to H$ – вполненепрерывный оператор, $\mathcal{D} \subset \subset Z$ – замкнутое ограниченное множество, $Z$, $H$ – гильбертовы пространства, $p \in H$ – параметр. Будем считать, что для любых двух последовательностей $\bar{z}^i$, $\bar{\bar{z}}^i \in \bar{\mathcal{D}}$, $i = 1, 2, \ldots$, таких, что $\|\bar{z}^i - \bar{\bar{z}}^i\| \to 0,\; i \to \infty$, имеет место предельное соотношение $\|g(\bar{z}^i) - g(\bar{\bar{z}}^i)\| \to 0,\; i \to \infty$.

Обозначим: $\mathcal{D}_p^\varepsilon \equiv \{z \in \mathcal{D} : \|g(z) - p\| \leqslant \varepsilon\},\; \varepsilon \geqslant 0$. Определим функцию значений ($S$-функцию) $\beta : H \to R^1 \cup \{+\infty\}$ задачи $(P_p)$

$$\beta(p) \equiv \lim_{\varepsilon \to +0} \beta_\varepsilon(p), \;\; \beta_\varepsilon(p) \equiv \inf_{z \in \mathcal{D}_p^\varepsilon} f(z), \qquad \beta_\varepsilon(p) \equiv +\infty, \text{ если } \mathcal{D}_p^\varepsilon = \emptyset.$$

Очевидно, в общей ситуации $\beta(p) \leqslant \beta_0(p)$, где $\beta_0(p)$ – классическое значение задачи $(P_p)$.

Справедлива следующая важная для дальнейших построений

Л е м м а 4. *Функция значений $\beta : H \to R^1 \cup \{+\infty\}$ является полунепрерывной снизу.*

Как и в линейно выпуклом случае, определим минимизирующую последовательность – минимизирующее приближенное решение в смысле Дж. Варги [18] в задаче $(P_p)$, то есть последовательность элементов $z^i \in \mathcal{D},\; i = 1, 2, \ldots$, такую, что $f(z^i) \leqslant \beta(p) + \delta^i$, $z^i \in \mathcal{D}_p^{\varepsilon^i}$ для некоторых последовательностей сходящихся к нулю неотрицательных чисел $\delta^i$, $\varepsilon^i$, $i = 1, 2, \ldots$.

Пусть $\mathsf{F}$ – множество всевозможных наборов исходных данных $\mathsf{f} \equiv \{f, g\}$, каждый из которых состоит из непрерывного на $\mathcal{D}$ функционала $f$ и вполненепрерывного оператора $g$ с указанным выше дополнительным свойством непрерывности. Определим наборы невозмущенных $\mathsf{f}^0$ и возмущенных $\mathsf{f}^\delta$ исходных данных соответственно: $\mathsf{f}^0 \equiv \{f^0, g^0\}$ и $\mathsf{f}^\delta \equiv \{f^\delta, g^\delta\}$, $\delta \in (0, \delta_0]$, $\delta_0 > 0$ - некоторое число. Будем считать, что выполняются следующие оценки

$$|f^\delta(z) - f^0(z)|, \;\; \|g^\delta(z) - g^0(z)\| \leqslant K\delta \;\; \forall\, z \in \mathcal{D},$$

где $K > 0 > 0$ – некоторая не зависящая от $\delta$ постоянная.

Обозначим задачу $(P_p)$, функционал $f$, оператор $g$, функцию значений $\beta$ и т.п., соответствующие набору исходных данных $\mathsf{f}^\delta$, $\delta \in [0, \delta_0]$, через $(P_p^\delta)$, $f^\delta$, $g^\delta$, $\beta^\delta$, соответственно.

К нелинейным задачам математического программирования общего вида $(P_p)$ сводятся самые разнообразные задачи. К ним относятся, например, нелинейные задачи оптимального управления, различные нелинейные обратные задачи.

В качестве одного из возможных примеров нелинейной задачи оптимального управления рассмотрим, например, в гильбертовом пространстве $Z = L_2(0, T)$ параметрическую задачу оптимального управления с фиксированным временем

$$f(u) \equiv \int_0^T F(t, x[u](t), u(t))dt \to \min, \;\; g(u)(t) \equiv \varphi(t, x[u](t)) = p(t) \text{ при п.в. } t \in X,$$

$$\dot{x} = h(t, x, u(t)), \;\; x(0) = x_0 \in R^n, \;\; t \in [0,T], \;\; u \in \mathcal{D} \subset L_2(0,T), \;\; p \in H \; - \text{ параметр},$$

где $F : [0,T] \times R^n \times R^m \to R^1$, $h : [0,T] \times R^n \times R^m \to R^n$, $\varphi : [0,T] \times R^n \to R^1$, – непрерывные вместе с градиентами $\nabla_x F$, $\nabla_x h$, $\nabla_x \varphi$ функции, $\mathcal{D} \equiv \{u \in L_2(0,T) : u(t) \in U$ при п.в. $t \in (0,T)\}$, $U \subset R^m$ – компакт, $X \subset [0,T]$, $X = X$, $H \equiv L_2(X)$. Считаем при этом одновременно для простоты, что для каждого управления $u \in \mathcal{D}$ существует единственное решение задачи Коши $x[u](t)$, $t \in [0,T]$, причем все эти решения равномерно ограничены.

В качестве одного из возможных важных примеров нелинейной обратной задачи, сводящейся к форме $(P_p)$, укажем на обратную задачу в следующей постановке.

Пусть задана линейная система дифференциальных уравнений

$$\dot{x} = A(t, s, u(t))x, \;\; x(0) = h \in R^n, \;\; t \in [0,T], \;\; s \in [a,b] \; - \;\; \text{числовой параметр},$$

где $a$, $b$ – заданные числа, $u \in \mathcal{D} \equiv \{u \in L_2(0,T) : u(t) \in U$ при п.в. $t \in (0,T)\}$, $U \subset R^m$ – компакт, $A : [0,T] \times [a,b] \times R^m \to R^{m \times n}$ – некоторая заданная функция. Требуется по наблюдениям траектории указанной системы $x[u, h; s](T)$ для значений параметра $s \in [a,b]$ (параметр $s$ может играть, например, роль частоты электромагнитного излучения) в финальный момент времени $T$ определить начальное условие $h$ и возмущающее воздействие $u \in \mathcal{D}$. Подобные постановки обратных задач встречаются в различных физических приложениях (см., например, [30]).

### 2.2. Субдифференциалы полунепрерывных снизу функций и модифицированные функции Лагранжа

В рассмотренном выше линейно выпуклом случе центральную роль играет классическая конструкция функции Лагранжа и качественные свойства алгоритма двойственной регуляризации полностью определяются дифференциальными свойствами полунепрерывной снизу выпуклой функции значений (теоремы 1, 3). В нелинейном случае роль субдифференциалов в смысле выпуклого анализа линейно выпуклого случая берут на себя субдифференциалы [31–35] полунепрерывных снизу нелинейных функций значений. Именно они порождают конструкции модифицированных функций Лагранжа и определяют качественные свойства алгоритма двойственной регуляризации в нелинейной задаче математического программирования.

Введем прежде всего понятие проксимального субградиента полунепрерывной снизу функции (см., например, [31–33]). Для этого сначала напомним понятие проксимальной нормали.

О п р е д е л е н и е 1. (а) Пусть $H$ - гильбертово пространство, $S \subset H$ - замкнутое множество, $\bar{s} \in S$. Вектор $\zeta \in H$ называется проксимальной нормалью к множеству $S$ в точке $\bar{s} \in S$, если существует постоянная $M > 0$ такая, что

$$\langle \zeta, s - \bar{s} \rangle \leqslant M \|s - \bar{s}\|^2 \;\; \forall\, s \in S. \tag{21}$$

Множество всех таких векторов $\zeta$, представляющее собой конус, обозначим через $\hat{N}_S(\bar{s})$ и назовем проксимальным нормальным конусом.

(б) Пусть $f : H \to R^1 \cup \{+\infty\}$ полунепрерывная снизу функция и $\bar{x} \in dom\, f$. Вектор $\zeta \in H$ называется проксимальным субградиентом функции $f$ в точке $\bar{x}$, если

$$(\zeta, -1) \in \hat{N}_{epi\, f}(\bar{x}, f(\bar{x})).$$

Множество всех таких векторов $\zeta$ обозначим через $\partial_{PN} f(\bar{x})$ и назовем проксимальным субградиентом $f$ в точке $\bar{x}$.

З а м е ч а н и е 2. Неравенство (21) может быть записано в виде

$$\langle \frac{1}{2M}\zeta, s - \bar{s} \rangle \leqslant \frac{1}{2}\|s - \bar{s}\|^2 \;\; \forall\, s \in S.$$

которое, согласно лемме 3C.2 в [32] (см. также [33]), эквивалентно включению

$$\bar{s} \in Pr_S(\bar{s} + \frac{1}{2M}\zeta).$$

Другими словами, $\zeta$ есть проксимальная нормаль к $S$ в $\bar{s}$ тогда и только тогда, когда $\bar{s}$ есть ближайшая в $S$ точка к некоторой точке вида $\bar{s} + t\zeta,\ t > 0$.

Л е м м а 5. *Пусть $f : H \to R^1 \cup \{+\infty\}$ полунепрерывная снизу функция и $\bar{x} \in dom\, f$. Вектор $\zeta \in H$ является проксимальным субградиентом функции $f$ в точке $\bar{x}$ тогда и только тогда, когда существуют постоянные $R > 0$ и $\delta > 0$ такие, что*

$$\langle \zeta, x - \bar{x} \rangle \leqslant f(x) - f(\bar{x}) + R\|x - \bar{x}\|^2 \ \ \forall\, x \in S_\delta(\bar{x}) \equiv \{x' \in H :\ \|x' - \bar{x}\| < \delta\}$$

*или*

$$f(\bar{x}) - \langle \zeta, \bar{x} \rangle \leqslant f(x) - \langle \zeta, x \rangle + R\|x - \bar{x}\|^2 \ \ \forall\, x \in S_\delta(\bar{x}).$$

Определим далее понятие нормали Фреше [31, 34, 35] к замкнутому множеству в банаховом пространстве, а также соответствующее понятие субдифференциала полунепрерывной снизу функции. Следующие два определения могут быть найдены в [34, 35].

О п р е д е л е н и е 2. Пусть $\Omega$ - непустое множество банахова пространства $X$, $\varepsilon \geqslant 0$. Пусть $x \in cl\,\Omega$. Тогда непустое множество

$$\hat{N}_\varepsilon(x;\Omega) \equiv \{x^* \in X^* :\ \limsup_{u \xrightarrow{\Omega} x} \frac{\langle x^*, u - x\rangle}{\|u - x\|} \leqslant \varepsilon\}$$

называется множеством $\varepsilon$-нормалей Фреше ко множеству $\Omega$ в точке $x$. При $\varepsilon = 0$ это множество является конусом, называется нормальным конусом Фреше к $\Omega$ в $x$ и обозначается $\hat{N}(x;\Omega)$. При $x \notin cl\,\Omega$ полагается $\hat{N}_\varepsilon(x;\Omega) = \emptyset$ для всех $\varepsilon \geqslant 0$.

О п р е д е л е н и е 3. Пусть $f : X \to R^1 \cup \{+\infty\}$ - полунепрерывная снизу функция, определенная на банаховом пространстве $X$, $\bar{x} \in dom\, f$. Множество

$$\hat{\partial} f(\bar{x}) \equiv \{x^* \in X^* :\ (x^*, -1) \in \hat{N}((\bar{x}, f(\bar{x})); epi\, f)\},$$

называется субдифференциалом Фреше функции $f$ в точке $\bar{x}$. При этом полагается $\hat{\partial} f(\bar{x}) = \emptyset$ в случае $x \notin dom\, f$.

Справедливо следующее [34, 35].

З а м е ч а н и е 3. Субдифференциал $\hat{\partial} f(\bar{x})$ может быть записан в виде

$$\hat{\partial} f(\bar{x}) = \{x^* \in X^* :\ \liminf_{u \to x} \frac{f(u) - f(x) - \langle x^*, u - x\rangle}{\|u - x\|} \geqslant 0\}.$$

Справедлива также следующая [34]

Л е м м а 6. *Пусть $f : X \to R^1 \cup \{+\infty\}$ - полунепрерывная снизу функция, определенная на банаховом пространстве $X$, $x \in dom\, f$, $\varepsilon > 0$. Тогда $x^* \in \hat{\partial} f(x)$ в том и только в том случае, если существует окрестность $X_\varepsilon$ точки $x$ такая, что*

$$f(x') - f(x) - \langle x^*, x' - x\rangle + \varepsilon\|x' - x\| \geqslant 0 \ \ \forall\, x' \in X_\varepsilon$$

*или*

$$f(x) - \langle x^*, x\rangle \leqslant f(x') - \langle x^*, x'\rangle + \varepsilon\|x' - x\| \ \ \forall\, x' \in X_\varepsilon.$$

Важнейшим свойством полунепрерывных снизу функций $f : X \to R^1 \cup \{+\infty\}$ является то, что как множество $\partial_{PN} f(x)$ в случае гильбертова пространства $X$, так и множество $\hat{\partial} f(x)$ в случае пространства $X$ из достаточно обширного класса банаховых пространств (подробности см., например, в [31–35] не пусто для плотного в $f$ множества. В настоящей работе в качестве пространства $X$ выступает гильбертово пространство $H$, для которого указанные выше свойства заведомо справедливы.

Если точка $p \in \beta$ такова, что $\partial_{PN}\beta(p) \neq \emptyset$ и $\zeta \in \partial_{PN}\beta(p)$, то из леммы 5 следует, что существуют постоянные $R > 0$ и $\delta > 0$ такие, что

$$\beta(p) - \langle \zeta, p \rangle \leqslant \beta(p') - \langle \zeta, p' \rangle + R\|p' - p\|^2 \ \ \forall\, p' \in S_\delta(p). \tag{22}$$

Так как в силу ограниченности множества $\mathcal{D}$ эффективное множество $\beta$ ограничено и функция $\beta$ ограничена на $\beta$, то в силу неравенства (22) можем записать для некоторой постоянной $c = c(p, \zeta) > 0$

$$\beta(p) - \langle \zeta, p \rangle \leqslant \beta(p') - \langle \zeta, p' \rangle + c\|p' - p\|^2 \ \ \forall\, p' \in \mathcal{H} \tag{23}$$

или

$$\beta(p) - \langle \zeta, p \rangle < \beta(p') - \langle \zeta, p' \rangle + \hat{c}\|p' - p\|^2 \ \ \forall\, p' \in \mathcal{H},\ \ p' \neq p, \quad \hat{c} > c,$$

откуда в силу полунепрерывности снизу функции значений $\beta$ и ограниченности $\beta$ следует, что минимизирующей последовательностью в задаче минимизации

$$\beta(p') - \langle \zeta, p' \rangle + \hat{c}\|p' - p\|^2 \to \inf,\ \ p' \in \mathcal{H} \tag{24}$$

является лишь любая последовательность $p^k$, $k = 1, 2, \ldots$, сходящаяся к точке $p$ такая, что $\beta(p^k) \to \beta(p)$, $k \to \infty$, и никакая другая последовательность. Отсюда следует, что в задаче минимизации модифицированной функции Лагранжа

$$f(z) - \langle \zeta, g(z) - p \rangle + \hat{c}\|g(z) - p\|^2 \to \inf,\ \ u \in \mathcal{D} \tag{25}$$

минимизирующей является лишь последовательность $z^k$, $k = 1, 2, \ldots$ такая, что $f(z^k) \to \beta(p)$, $g(z^k) \to p$, $k \to \infty$, и никакая другая последовательность. При этом справедливо равенство

$$\inf_{z \in \mathcal{D}} (f(z) - \langle \zeta, g(z) - p \rangle + \hat{c}\|g(z) - p\|^2) = \beta(p). \tag{26}$$

Одновременно, если для некоторых $\zeta \in \mathcal{H}$, $\hat{c} > 0$ выполняется равенство (26), то выполняется и неравенство (23) при $c = \hat{c}$, а следовательно и неравенство (22) при $R = \hat{c}$ и любом $\delta > 0$, то есть $\zeta \in \partial_{PN}\beta(p)$. Отсюда можно вывести следующее важное для нас следствие.

Если $p \in \beta$ такая точка, что $\partial_{PN}\beta(p) \neq \emptyset$ и для всех $\zeta \in \partial_{PN}\beta(p)$ определенный выше в неравенстве (23) штрафной коэффициент $c = c(p, \zeta)$ можно взять независимым от $\zeta \in \partial_{PN}\beta(p)$, то есть $c = c(p)$, то тогда, как легко заметить, этот коэффициент можно считать столь большим, что неравенство (23) имеет место лишь для $\zeta \in \partial_{PN}\beta(p)$, причем для всех таких $\zeta$ минимизирующей последовательностью в задаче минимизации (24) является лишь любая последовательность $p^k$, $k = 1, 2, \ldots$, сходящаяся к точке $p$ такая, что $\beta(p^k) \to \beta(p)$, $k \to \infty$, и никакая другая последовательность.

При этом одновременно, в задаче минимизации модифицированной функции Лагранжа (25) равенство (26) выполняется лишь при $\zeta \in \partial_{PN}\beta(p)$ и минимизирующей в ней является лишь последовательность $z^k$, $k = 1, 2, \ldots$ такая, что $f(z^k) \to \beta(p)$, $g(z^k) \to p$, $k \to \infty$ и никакая другая последовательность.

Если точка $p \in \beta$ такова, что $\hat{\partial}\beta(p) \neq \emptyset$ и $\zeta \in \hat{\partial}\beta(p)$, то из леммы 6 следует, что существуют постоянные $R > 0$ и $\delta > 0$ такие, что

$$\beta(p) - \langle \zeta, p \rangle \leqslant \beta(p') - \langle \zeta, p' \rangle + R\|p' - p\| \ \ \forall\, p' \in S_\delta(p). \tag{27}$$

Так как в силу ограниченности множества $\mathcal{D}$ эффективное множество $\beta$ ограничено и функция $\beta$ ограничена на множестве $\beta$, то в силу неравенства (27) можем записать для некоторой постоянной $c = c(p, \zeta) > 0$

$$\beta(p) - \langle \zeta, p \rangle \leqslant \beta(p') - \langle \zeta, p' \rangle + c\|p' - p\| \ \ \forall\, p' \in H$$

или

$$\beta(p) - \langle \zeta, p \rangle \leqslant \beta(p') - \langle \zeta, p' \rangle + \hat{c}\|p' - p\| \ \ \forall\, p' \in H, \ \ p' \neq p, \qquad \hat{c} > c,$$

откуда в силу полунепрерывности снизу функции значений $\beta$ и ограниченности множества $\beta$ следует, что минимизирующей последовательностью в задаче минимизации

$$\beta(p') - \langle \zeta, p' \rangle + \hat{c}\|p' - p\| \to \inf, \ \ p' \in H$$

является лишь любая последовательность $p^k$, $k = 1, 2, \ldots$, сходящаяся к точке $p$ такая, что $\beta(p^k) \to \beta(p)$, $k \to \infty$, и никакая другая последовательность. Отсюда следует, что в задаче минимизации модифицированной функции Лагранжа

$$f(z) - \langle \zeta, g(z) - p \rangle + \hat{c}\|g(z) - p\| \to \inf, \ \ z \in \mathcal{D} \tag{28}$$

минимизирующей является лишь последовательность $z^k$, $k = 1, 2, \ldots$ такая, что $f(z^k) \to \beta(p)$, $g(z^k) \to p$, $k \to \infty$, и никакая другая последовательность. При этом справедливо равенство

$$\inf_{z \in \mathcal{D}} (f(z) - \langle \zeta, g(z) - p \rangle + \hat{c}\|g(z) - p\|) = \beta(p).$$

Определим, далее с учетом конструкций модифицированных функций Лагранжа задач (25), (28) (в связи с указанными конструкциями см. также [2, 36, 37]) модифицированную функцию Лагранжа задачи $(P_p^\delta)$

$$L_{p,c}^\delta(z, \lambda) \equiv f^\delta(z) + \langle \lambda, g^\delta(z) - p \rangle + c\psi(\|g^\delta(z) - p\|), z \in \mathcal{D}, \ \lambda \in H, \ c \geqslant 0,$$

где штрафная функция $\psi : R_+^1 \to R_+^1$, определяется формулой $\psi(t) \equiv l_1 t + l_2 t^2$, $t \in R_+^1$, в которой весовые множители $l_1$, $l_2 \in \{0, 1\}$ – фиксированные числа. В этой конструкции штрафной функции $\psi$ мы учитываем сразу обе возможности для функции значений $\beta^0$ иметь в точке $p$ либо проксимальный субградиент, либо субдифференциал Фреше, либо и то и другое одновременно.

Определим, в свою очередь, и модифицированную двойственную задачу

$$V_{p,c}^\delta(\lambda) \to \sup, \ \ \lambda \in H, \quad V_{p,c}^\delta(\lambda) \equiv \inf_{z \in \mathcal{D}} L_{p,c}^\delta(z, \lambda).$$

Условия на исходные данные задачи $(P_p^0)$ таковы, что функция $V_{p,c}^\delta$ при $\delta \in [0, \delta_0]$ является определенной (конечной) при любом $c \in R^1$ для любой точки $\lambda \in H$ и выполняется оценка

$$|V_{p,c}^\delta(\lambda) - V_{p,c}^0(\lambda)| \leqslant C\delta(1 + \|\lambda\| + |c|) \ \ \forall\, \lambda \in H,$$

где $C > 0$ – некоторая постоянная.

Как и в линейно выпуклом случае при конструировании минимизирующей последовательности в нелинейной задаче $(P_p)$ наиважнейшее значение имеет следующая лемма [4], в которой устанавливается выражение для супердифференциала $\partial V_{p,c}^\delta$ вогнутой функции значений $V_{p,c}^\delta$ при условии полной непрерывности оператора $g : \mathcal{D} \to H$. Здесь, как и в линейно выпуклом случае, под супердифференциалом вогнутой функции $V_{p,c}^\delta$ понимается субдифференциал (в смысле выпуклого анализа) с обратным знаком выпуклой функции $-V_{p,c}^\delta$.

Л е м м а 7. *Супердифференциал (в смысле выпуклого анализа)* $\partial V^{\delta}_{p,c}(\lambda)$, $\delta \in [0, \delta_0]$, *вогнутой функции* $V^{\delta}_{p,c}$ *в точке* $\lambda \in H$ *при любом* $c \in R^1$ *выражается формулой*

$$\partial V^{\delta}_{p,c}(\lambda) = \partial_C V^{\delta}_{p,c}(\lambda) = \overline{conv}\{\lim_{i\to\infty}(g^{\delta}(z^i) - p):$$

$$z^i \in \mathcal{D},\ L^{\delta}_{p,c}(z^i, \lambda) \to \inf_{z\in\mathcal{D}} L^{\delta}_{p,c}(z, \lambda),\ i \to \infty\} \equiv \overline{conv} Q^{\delta}_{p,c}(\lambda),$$

*где* $\partial_C V^{\delta}_{p,c}(\lambda)$ *- обобщенный градиент Кларка функции* $V^{\delta}_{p,c}$ *в точке* $\lambda$.

## 2.3. Двойственная регуляризация для решения нелинейной задачи математического программирования

Сформулированные выше результаты позволяют организовать процесс двойственной регуляризации в нелинейной задаче математического программирования $(P_p)$ с целью конструирования в ней минимизирующей последовательности – минимизирующего приближенного решения. Дадим краткое описание этого процесса.

Укажем прежде всего на важные для дальнейшего обстоятельства, вытекающие из результатов предыдущего раздела (см. задачи (25), (28). Возможны две и только две ситуации для исходной задачи $(P^0_p)$:

А) в задаче имеется вектор Куна-Таккера в следующем обобщенном смысле: существует вектор $\lambda \in \mathcal{H}$, для которого $\beta^0(p) \leqslant \inf_{z\in\mathcal{D}} L^0_{p,c}(z, \lambda)$ для некоторого $c > 0$;

Б) в задаче не существует вектора Куна-Таккера в указанном смысле.

Существование вектора Куна-Таккера в указанном (обобщенном) смысле эквивалентно тому, что целевая функция $V^0_{p,c}(\lambda),\ \lambda \in \mathcal{H}$ в модифицированной двойственной задаче

$$V^0_{p,c}(\lambda) \to \sup, \quad \lambda \in \mathcal{H}$$

достигает значения $\beta^0(p)$ в некоторой точке $\lambda^0 \in \mathcal{H}$.

Лемма 7 лежит в основе организации алгоритма [29] поиска максимума в задаче максимизации при каждом $c > 0$ сильно вогнутого функционала $R^{\delta,\alpha}_{p,c}(\lambda) \equiv V^{\delta}_{p,c}(\lambda) - \alpha|\lambda|^2,\ \lambda \in \mathcal{H}$. При этом с целью конструирования минимизирующей последовательности в исходной задаче $(P^0_p)$ при $\tilde{c} > c$ рассматривается задача

$$R^{\delta,\alpha}_{p,\tilde{c}}(\lambda) \to \max, \quad \lambda \in \Lambda_{\tilde{c}} \equiv \{\lambda \in \mathcal{H}:\ \|\lambda\| \leqslant \tilde{c}\}. \tag{29}$$

Обозначим через $\lambda^{\alpha}_{p,c}$ единственную в $\Lambda_c$ точку, дающую на $\Lambda_c$ максимум функционалу $R^{\alpha}_{p,c}$ (она, очевидно, существует). Регуляризованный процесс [29] поиска точки максимума $\lambda^{\alpha}_{p,\tilde{c}}$ в модифицированной двойственной задаче (29) при выполнении условия согласования (аналогичного тому, которое имеет место в линейно выпуклом случае) $\delta/\alpha(\delta) \to 0,\ \delta \to 0,\ \alpha(\delta) \to 0$ конструктивно порождает минимизирующую последовательность $z^i \in \mathcal{D},\ i = 1, 2, \ldots$ в задаче $(P_p)$, то есть $f^0(z^i) \to \beta^0(p),\ z^i \in \mathcal{D}^{0,\varepsilon^i}_p,\ \varepsilon^i \to 0,\ i \to \infty$. При этом в случае А) величина $\tilde{c}$ может быть взята равной любому фиксированному достаточно большому положительному числу. В случае же Б) штрафной коэффициент $\tilde{c}$ необходимо стремить к $+\infty$. Основное предположение при этом заключается в том, что минимизация модифицированной функции Лагранжа может проводиться с любой наперед заданной точностью. Необходимые подробности можно найти в [29] (см. также [28]).

ЛИТЕРАТУРА

1. *Васильев Ф.П.* Методы оптимизации. М.: Факториал Пресс, 2002.

2. *Мину М.* Математическое программирование. Теория и алгоритмы. М.: Наука, 1990.

3. *Эрроу К.Дж., Гурвиц Л., Удзава Х.* Исследования по линейному и нелинейному программированию. М.: ИЛ, 1962. [Англ. оригинал: Arrow K.J., Hurwicz L., Uzawa H. Studies in Linear and Nonlinear Programming. Stanford University Press, 1958.]

4. *Сумин М.И.* Регуляризация в линейно выпуклой задаче математического программирования на основе теории двойственности // Ж. вычисл. матем. и матем. физ. 2007. Т. 47. № 4. С. 602-625.

5. *Экланд И., Темам Р.* Выпуклый анализ и вариационные проблемы. М.: Мир, 1979.

6. *Гловински Р., Лионс Ж.-Л., Тремольер Р.* Численное исследование вариационных неравенств. М.: Мир, 1979.

7. *Темам Р.* Уравнения Навье–Стокса. Теория и численный анализ. М.: Мир, 1981.

8. *Сумин М.И.* Оптимальное управление параболическими уравнениями: двойственные численные методы, регуляризация // Распределенные системы: оптимизация и приложения в экономике и науках об окружающей среде: Сб. докладов к Международной конференции (Екатеринбург, 30 мая - 2 июня 2000 г.). 2000. Екатеринбург: Изд-во Ин-та математики и механики УрО РАН, 2000. С. 66-69.

9. *Сумин М.И.* Регуляризованный градиентный двойственный метод решения обратной задачи финального наблюдения для параболического уравнения // Ж. вычисл. матем. и матем. физ. 2004. Т. 44. 11. С. 2001-2019.

10. *Сумин М.И.* Итеративная регуляризация градиентного двойственного метода для решения интегрального уравнения Фредгольма первого рода // Вестник Нижегородского университета. Серия Математика. 2004. Вып. 1(2). С. 192–208.

11. *Сумин М.И.* Регуляризованный двойственный алгоритм в задачах оптимального управления для распределенных систем // Вестник Нижегородского университета. Серия Математическое моделирование и оптимальное управление. 2006. Вып. 2(31). С. 82-101.

12. *Тихонов А.Н., Арсенин В.Я.* Методы решения некорректных задач. М.: Наука, 1986.

13. *Сумин М.И.* Метод возмущений и двойственная регуляризация в линейно выпуклой задаче математического программирования // Проблемы динамического управления. Сборник научных трудов факультета ВМиК МГУ им. М.В. Ломоносова / Под редакцией Ю.С. Осипова, А.В. Кряжимского. Вып. 3. 2008. М.: Издательский отдел факультета ВМиК МГУ, С. 200-231.

14. *Sumin M.I.* Parametric Dual Regularization in a Linear-Convex Mathematical Programming // Computational Optimization: New Research Developments. New-York: Nova Science Publishers Inc, 2010 (at press).

15. *Сумин М.И.* Некорректные задачи и методы их решения. Материалы к лекциям для студентов старших курсов: Учеб. пособие. Н. Новгород: Изд-во Нижегород. гос. ун-та, 2009.

16. *Тихонов А.Н., Гончарский А.В., Степанов В.В., Ягола А.Г.* Регуляризирующие алгоритмы и априорная информация. М.: Наука, 1983.

17. *Осипов Ю.С., Васильев Ф.П., Потапов М.М.* Основы метода динамической регуляризации. М.: Изд-во МГУ, 1999.

18. *Варга Дж.* Оптимальное управление дифференциальными и функциональными уравнениями. М.: Наука, 1977.

19. *Жидков А.А., Калинин А.В., Сумин М.И.* О некоторых обратных задачах для системы уравнений Максвелла в квазистационарном электрическом приближении // Теория и численные методы решения обратных и некорректных задач. Молодежная международная научная школа-конференция (Новосибирск, 10-20 августа 2009 г.): тезисы докладов. Новосибирск. Ин-т математики СО РАН, 2009. С. 50.

20. *Жидков А.А., Калинин А.В., Сумин М.И.* Двойственная регуляризация в обратных задачах атмосферного электричества // Супервычисления и математическое моделирование. XI Международный семинар: Тезисы. Саров: Изд-во ФГУП РФЯЦ ВНИИЭФ, 2009. С. 65-66.

21. *Алексеев В.М., Тихомиров В.М., Фомин С.В.* Оптимальное управление. М.: Наука, 1979.

22. *Обен Ж.-П., Экланд И.* Прикладной нелинейный анализ. М.: Мир, 1988.

23. *Обен Ж.-П.* Нелинейный анализ и его экономические приложения. М.: Мир, 1988.

24. *Ward A.L.* Differentiability of Vector Monotone Functions // Proc. London Math. Soc. 1935. V. 32. No. 2. P. 339-362.

25. *Бакушинский А.Б., Гончарский А.В.* Некорректные задачи. Численные методы и приложения. М.: Изд-во МГУ, 1989.

26. *Сумин М.И.* Параметрическая двойственная регуляризация и принцип максимума в задаче опти-мального управления с фазовыми ограничениями // Вестник Тамбовского ун-та. Серия.: Естественные и. технические. науки. Тамбов: Изд-во Тамбовского ун-та. 2009. Т. 14. Вып. 4. С. 807-809.

27. *Сумин М.И.* Параметрическая двойственная регуляризация для задачи оптимального управления с поточечными фазовыми ограничениями // Ж. вычисл. матем. и матем. физ. 2009. Т. 49. № 12. С. 2083-2102.

28. *Сумин М.И.* Регуляризованный двойственный метод решения нелинейной задачи математического программирования // Ж. вычисл. матем. и матем. физ. 2007. Т. 47. № 5. С. 796–816.

29. *Sumin M.I.* Parametric Dual Regularization in a Nonlinear Mathematical Programming // Advances in Mathematics Research. Volume 11. New-York: Nova Science Publishers Inc, 2010 (at press).

30. *Гайкович К.П., Кутерин Ф.А., Смирнов А.И., Сумин М.И.* Двойственная регуляризация в обратной задаче УНЧ зондирования земной коры // Вестник Нижегородского университета им. Н.И. Лобачевского. 2009. № 1. С. 47-52.

31. *Borwein J.M., Strojwas H.M.* Proximal Analysis and Boundaries of Closed Sets in Banach Space, Part I: Theory // Can. J. Math. 1986. V. 38. No. 2. P. 431-452; Part II: Applications // Can. J. Math. 1987. V. 39. № 2. P. 428-472.

32. *Loewen P.D.* Optimal Control via Nonsmooth Analysis. CRM Proceedings and Lecture Notes. V. 2. Amer. Math. Soc., Providence, RI. 1993.

33. *Clarke F.H., Ledyaev Yu.S., Stern R.J., Wolenski P.R.* Nonsmooth Analysis and Control Theory. Graduate Texts in Mathematics. V. 178. New-York: Springer-Verlag, 1998.

34. *Mordukhovich B.S.* Variational Analysis and Generalized Differentiation. I: Basic Theory. Berlin: Springer. 2006.

35. *Mordukhovich B.S.* Variational Analysis and Generalized Differentiation. II: Applications. Berlin: Springer, 2006.

36. *Бертсекас Д.* Условная оптимизация и методы множителей Лагранжа. М.: Радио и связь, 1987.

37. *Гольштейн Е.Г., Третьяков Н.В.* Модифицированные функции Лагранжа. Теория и методы оптимизации. М.: Наука, 1989.

БЛАГОДАРНОСТИ: Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (проекты 07-01-00495-а; 09-01-97019-р_поволжье_а), а также аналитической ведомственной целевой программы «Развитие научного потенциала высшей школы (2009–2010 годы)» Минобрнауки РФ (проект 2.1.1/3927) и федеральной целевой программы «Научные и научно-педагогические кадры инновационной России» на 2009-2013 годы (проект НК-13П-13).

Поступила в редакцию 5 октября 2009 г.

Sumin M. I. Parametric dual regularization in optimization, optimal control and inverse problems. The article is devoted to application of a perturbation method in the theory of dual regularization both for linear-convex, and for a nonlinear problem of mathematical programming in a Hilbert space. Primary attention is given to the qualitative properties of the dual regularization method, depending on the differential properties of the value function (S-function) in the optimization problem. It is shown that the convergence of the method is closely related to the Lagrange principle. It is shown, that the scheme of dual regularization gives a new way of the proof of the Lagrange principle and leads to its useful specifications. The so-called regularized Lagrange principle in nondifferential form is discussed. Also, possibility of application of the dual regularization method in parametric problems of optimisation, optimal control and in parametric inverse problems is discussed.

Key words: Linear-convex mathematical programming; nonlinear mathematical programming; parametric problem; minimizing sequence; Lagrange principle; duality; regularization; perturbation method; optimization; optimal control; inverse problems.