ВЕСТНИК УДМУРТСКОГО УНИВЕРСИТЕТА. МАТЕМАТИКА. МЕХАНИКА. КОМПЬЮТЕРНЫЕ НАУКИ
МАТЕМАТИКА 2020. Т. 30. Вып. 3. С. 410–428.

УДК 519.853, 517.98

© *М. И. Сумин*

О РЕГУЛЯРИЗАЦИИ ПРИНЦИПА ЛАГРАНЖА И ПОСТРОЕНИИ ОБОБЩЕННЫХ МИНИМИЗИРУЮЩИХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ В ВЫПУКЛЫХ ЗАДАЧАХ УСЛОВНОЙ ОПТИМИЗАЦИИ

Рассматривается регуляризация принципа Лагранжа (ПЛ) в выпуклой задаче условной оптимизации с операторным ограничением-равенством в гильбертовом пространстве и конечным числом функциональных ограничений-неравенств. Целевой функционал задачи не является, вообще говоря, сильно выпуклым, а на множество ее допустимых элементов, которое также принадлежит гильбертову пространству, не накладывается условие ограниченности. Получение регуляризованного ПЛ основано на методе двойственной регуляризации и предполагает использование двух параметров регуляризации и двух соответствующих условий согласования одновременно. Один из регуляризирующих параметров «отвечает» за регуляризацию двойственной задачи, другой же содержится в сильно выпуклом регуляризирующем добавке к целевому функционалу исходной задачи. Основное предназначение регуляризованного ПЛ — устойчивое генерирование обобщенных минимизирующих последовательностей, аппроксимирующих точное решение задачи по функции и по ограничениям, для целей ее непосредственного практического устойчивого решения.

Ключевые слова: условная оптимизация, неустойчивость, двойственная регуляризация, регуляризованный принцип Лагранжа, обобщенная минимизирующая последовательность.

DOI: 10.35634/vm200305

Введение

Основным изучаемым объектом в статье является каноническая задача выпуклого программирования (см., например, [1, п. 3.3.1]) для пары гильбертовых пространств Z, H

$$(P) \qquad f(z) \to \min, \quad Az = h, \quad g_1(z) \leqslant 0, \dots, g_m(z) \leqslant 0, \quad z \in \mathcal{D} \subset Z,$$

где f, $g_i\colon \mathcal{D} \to \mathbb{R}^1$, $i = 1, \dots, m$, — выпуклые непрерывные функционалы, $A\colon Z \to H$ — линейный ограниченный оператор, $h \in H$ — заданный элемент, $\mathcal{D}$ — выпуклое замкнутое множество. Главным теоретическим результатом для задач класса (P), а также и для всей теории условной оптимизации, служит, как известно, принцип Лагранжа (ПЛ) (см., например, [1, 2]). Общепринятым является тот факт (подробности можно найти в [1, глава 1]), что ПЛ появился вследствие потребностей решения разного рода практических экстремальных задач. Вместе с тем следует признать, что применение этого фундаментального результата для непосредственного практического решения большого числа задач условной оптимизации и многих сводящихся к ним задач современного естествознания встречается с трудностями принципиального характера, обусловленными неразрывно связанными с ПЛ свойствами некорректности. Здесь имеются в виду такие ее проявления как неустойчивость и невыполнимость ПЛ. Мы говорим о неустойчивости ПЛ в задаче класса (P), если выделяемые им в задачах, «сколь угодно близких» к исходной (невозмущенной) задаче, «приближенные» оптимальные элементы могут сколь угодно сильно отличаться от своего невозмущенного аналога как по аргументу, так и по функции [3–6]. В свою очередь, невыполнимость ПЛ, в той или иной конкретной задаче класса (P), понимается как принципиальная невозможность записать его для этой конкретной задачи в той «привычной» форме, в которой он записывается в «большинстве» других аналогичных задач этого класса [1, с. 260], [3, 4, 6]. Приведем характерный пример невыполнимости ПЛ. Он показывает, что существуют важные с точки зрения многих приложений задачи условной оптимизации, в которых его формальное применение, совершенно «бессмысленно» по причине, что его в этих задачах просто «невозможно записать». Другие содержательные примеры некорректности ПЛ в задачах класса (P) см. в [3–6].

Пример 1. Рассмотрим задачу класса (P)

$$\|z\|^2 \to \min, \quad Az = h \tag{0.1}$$

с инъективным и самосопряженным оператором $A\colon Z \to Z$, Z — гильбертово пространство, таким, что $R(A) \neq Z$ (например, A может быть интегральным оператором Фредгольма с замкнутым симметрическим ядром). Покажем, что в задаче (0.1) при выбранных определенным образом h ПЛ [1] не выполняется. Пусть $z^0 \in Z$, но $z^0 \notin R(A)$. Тогда рассматриваем задачу (0.1) с $h = Az^0$. В ней классический ПЛ в дифференциальной форме [1, п. 3.2.2] (см. также параметрические ПЛ в [3,4,6]), а как следствие, и в недифференциальной форме, не выполняется. Действительно, если бы это было не так, то существовала бы невырожденная пара множителей $(\lambda_0, \lambda) \in \mathbb{R}^1 \times Z$ такая, что $2\lambda_0 z^0 + A\lambda = 0$. В этом случае при $\lambda_0 = 0$ получаем $\lambda = 0$ в силу инъективности A, а при $\lambda_0 = 1$, соответственно, противоречивое равенство $z^0 = -1/2A\lambda$, что и доказывает невыполнимость классического ПЛ в задаче (0.1) с выбранным h[1]. Наконец, можно также утверждать, что хорошо известным классическим фактом является то, что задача (0.1) для всех h, для которых она разрешима, а вместе с этим и классические ПЛ, теорема Куна–Таккера для нее, в случае их применимости, неустойчивы по отношению к ошибкам исходных данных [7]. В качестве конкретной задачи условной оптимизации вида (0.1) можно взять, например, задачу поиска нормального решения интегрального уравнений Фредгольма первого рода (подобные задачи относятся к числу классических в теории некорректных задач [7])

$$\|z\|^2_{2,(0,1)} \to \min, \quad \int_0^1 K(x,s)z(s)ds = h(x), \ \ 0 \leqslant x \leqslant 1, \ \ z \in L_2(0,1), \tag{0.2}$$

с симметричным ядром $K(x,s) = \{(1-x)s,\ 0 \leqslant s \leqslant x;\ x(1-s),\ x \leqslant s \leqslant 1\}$ и с $Z = L_2(0,1)$. В этом случае в соответствии с теоремой Пикара (см., например, [8, с. 148]) уравнение (0.2) из-за замкнутости ядра может быть только однозначно разрешимо (см. анализ примера 1 на с. 149 в [8]). Поэтому, в соответствии со сказанным выше, можно утверждать, что для любого $h(\cdot) = \int_0^1 K(\cdot,s)z(s)ds$ с такими $z \in L_2(0,1)$, которые не являются непрерывными функциями (соответствующий класс эквивалентности не содержит непрерывной функции), в задаче условной оптимизации (0.2) ПЛ не выполняется.

В статье показывается как основанный на двойственности подход к регуляризации [9] в задачах класса (P) (см. также [3, 4]) порождает при общих условиях на их исходные данные и соответствующую регуляризацию ПЛ в этих задачах, превращающую его в универсальный инструмент устойчивого решения всех задач этого класса, в том числе, и классической некорректной задачи (0.1). Эта регуляризация отличается от регуляризации «непосредственно самих» задач класса (P) (см., например, [2, гл. 9]). Она естественным образом трансформирует ПЛ в различного вида теоремы существования ограниченных обобщенных

[1]Можно заметить, что указанное обстоятельство имеет место на плотном подмножестве множества всех тех h, для которых задача (0.1) разрешима. При этом в каждой точке из этого плотного множества неразрешимой является соответствующая двойственная задача.

минимизирующих последовательностей (ОМП) в задачах указанного класса, являющиеся одновременно новыми регуляризирующими алгоритмами для их решения, то есть устойчивого построения ОМП, и преодоления тем самым свойств некорректности ПЛ [3–6]. В теории математического программирования применяемые в работе ОМП известны под названием обобщенных планов [10]. В свою очередь, в теории оптимального управления они получили название минимизирующих приближенных решений [11, гл. III]. Подчеркнем, что широко используемое в оптимизации понятие ОМП органично сочетает в себе учет как запросов строгой математической оптимизационной теории [10], [11, гл. IV–VIII], так и потребностей инженерной практики, предполагающей неизбежное наличие у приближенных решений ненулевых зазоров и при выполнении ограничений задачи и при приближении значений функционала цели к ее (обобщенной) нижней грани [11, гл. III].

Основное предназначение регуляризованных ПЛ, как уже отмечено выше, — устойчивое генерирование ОМП в задачах класса (P) для целей их непосредственного практического устойчивого решения. По сложившейся традиции в теории некорректных задач [2, 7] регуляризация ПЛ согласована с соответствующим ей понятием регуляризирующего оператора (алгоритма). Это понятие в задаче условной оптимизации (P), а также производное от него понятие ОМП-образующего оператора, введенные ранее в [12], существенным образом «привязаны» к задаче условной оптимизации и согласованы именно с понятием ОМП. Их отличительной особенностью является то, что генерируемые в соответствии с ними приближенные решения задачи «должны аппроксимировать» ее точное решение, одновременно, как по функции, так и «по ограничениям» без обязательного требования приближения по аргументу. Такое понятие регуляризирующего алгоритма в задачах класса (P), требующее минимальных дополнительных предположений об их исходных данных, позволяет позиционировать его как занимающее промежуточное положение между «привычными» понятиями сходимости по функции и по аргументу [2, гл. 9] (подробнее см. ниже в замечании 2).

Регуляризация ПЛ в задаче (P) проводится в настоящей работе в случае, когда ее целевой функционал не является, вообще говоря, сильно выпуклым и одновременно на множество допустимых элементов $\mathcal{D}$ не накладывается условие ограниченности. Различные варианты регуляризации ПЛ в выпуклых задачах условной оптимизации рассматривались ранее в работах [3, 4, 12–15]. Одним из основных в этих работах, как и в работах по обоснованию методов двойственной регуляризации [9], было предположение равенства в рассматриваемых задачах их обобщенной и классической нижних граней (см. определения обобщенной и классической нижних граней β и β_0 ниже в (1.2)). Такое равенство граней обеспечивалось в [3, 9, 14, 15] за счет сильной выпуклости целевого функционала задачи, а в [4, 12, 13] — благодаря ограниченности допустимого множества (множество $\mathcal{D}$ выше в задаче (P)).

Одновременно отметим и существенное различие результатов по регуляризации ПЛ в случае выпуклого (и, вообще говоря, не сильно выпуклого) целевого функционала [4, 13] от результатов [3], где рассматривалась задача с сильно выпуклым целевым функционалом (см. также [14, 15]). Кратко это различие можно описать следующим образом. Во-первых, как в том, так и в другом случае, ОМП конструируется из точек минимума функции Лагранжа задачи, соответствующих значениям двойственных переменных из некоторой последовательности, определяемой регуляризованным ПЛ. В случае сильно выпуклого целевого функционала сильно выпуклой по исходной (прямой) переменной является и функция Лагранжа и, как следствие, однозначно и корректно определяются и элементы ОМП. В отсутствие же сильной выпуклости, при ограниченном множестве допустимых элементов, гарантируется лишь существование элемента ОМП в соответствующем множестве минималей функции Лагранжа (см. теорему 1.5 в [4] и теорему 2 в [13]). Таким образом, генерирование ОМП в силу регуляризованного ПЛ в ситуации [4, 13] в существенной степени теряет свою конструктивность.

Постановка задачи настоящей работы не предполагает, что в задаче априори имеет место равенство $\beta = \beta_0$. Для рассмотрения задачи в такой ситуации мы вместо одного используем два параметра регуляризации. Один из них, как и в [3, 4, 13–15], «отвечает» за регуляризацию двойственной задачи, другой же содержится в сильно выпуклом регуляризирующем добавке к целевому функционалу исходной задачи. Несколько ранее подобный подход к регуляризации ПЛ в задаче класса (P), с применением двух параметров регуляризации, был предложен в работе [12], однако, в [12] существенным было предположение об ограниченности множества $\mathcal{D}$. При используемом подходе к регуляризации и при дополнительном условии ограниченности снизу целевого функционала (см. условие (2.3)) в работе на основе результатов [3] прежде всего конструируется ОМП-образующий оператор (см. теорему 2), который можно рассматривать как метод двойственной регуляризации [9] с дополнительным параметром регуляризации в целевом функционале в случае, когда одновременно на множество $\mathcal{D}$ не накладывается условие ограниченности, функционал цели может быть не сильно выпуклым и возможно строгое неравенство $\beta < \beta_0$. Для элементов конструируемой таким образом ОМП приводятся оценки точности приближения к решению по функции и по ограничениям. Этот ОМП-образующий оператор затем используется для доказательства регуляризованного ПЛ в рассматриваемой задаче в ситуации равенства двух граней задачи (см. теорему 3). Аналогичный регуляризованный ПЛ в случае ограниченности $\mathcal{D}$ получен в [12]. В обоих случаях при таком подходе преодолевается отмеченная выше неконструктивность регуляризованных ПЛ в [4, 13].

§ 1. Постановка задачи условной оптимизации

Рассмотрим задачу (P), исходные данные которой снабжены верхним индексом δ, говорящим о том, что эти данные либо являются точными ($\delta = 0$), либо возмущенными ($\delta > 0$), то есть заданными с ошибкой, величину которой и характеризует число $\delta \in [0, \delta_0]$ ($\delta_0 > 0$ — некоторое фиксированное число),

$$(P^\delta) \qquad f^\delta(z) \to \min, \ \ A^\delta z = h^\delta, \ \ g_i^\delta(z) \leqslant 0, \ \ i = 1, \dots, m, \quad z \in \mathcal{D} \subset Z,$$

где f^δ, $g_i^\delta\colon \mathcal{D} \to \mathbb{R}^1$, $i = 1, \dots, m$, — непрерывные выпуклые функционалы, $A^\delta\colon Z \to H$ — линейный ограниченный оператор, $h^\delta \in H$ — заданный элемент, $\mathcal{D}$ — выпуклое замкнутое множество, Z, H — гильбертовы пространства. Ниже будем использовать обозначение $g^\delta(z) \equiv (g_1^\delta(z), \dots, g_m^\delta(z))$. Если невозмущенная задача (P^0) разрешима, то обозначаем через z^0 ее решения, всю же совокупность этих решений, в свою очередь, обозначаем через Z^0.

Будем считать, что

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

где $L_M > 0$ — независящая от δ постоянная, $S_M \equiv \{z \in Z\colon \|z\| \leqslant M\}$. Будем также считать, что выполняются следующие оценки для отклонений возмущенных исходных данных от точных

$$\begin{gathered} |f^\delta(z) - f^0(z)| \leqslant C\delta(1 + \|z\|^2), \ \ \|A^\delta z - A^0 z\| \leqslant C\delta(1 + \|z\|), \\ |g^\delta(z) - g^0(z)| \leqslant C\delta(1 + \|z\|^2) \ \ \forall z \in \mathcal{D}, \ \ \|h^\delta - h^0\| \leqslant C\delta, \end{gathered} \tag{1.1}$$

где $C > 0$ не зависит от δ. Введем обозначения:

$$L^\delta(z, \lambda, \mu) \equiv f^\delta(z) + \langle \lambda, A^\delta z - h^\delta \rangle + \langle \mu, g^\delta(z) \rangle,$$

$$\mathcal{D}^{\delta,\epsilon} \equiv \{z \in \mathcal{D}\colon \|A^\delta z - h^\delta\| \leqslant \epsilon, \ g_i^\delta(z) \leqslant \epsilon, \ i = 1, \dots, m\}, \ \ \epsilon \geqslant 0, \ \ \mathcal{D}^{0,0} \equiv \mathcal{D}^0.$$

Определим обобщенное (β) и классическое (β_0) значения (нижние грани) задачи (P^0)

$$\beta \equiv \beta_{+0} \equiv \lim_{\epsilon \to +0} \beta_\epsilon, \quad \beta_\epsilon \equiv \inf_{z \in \mathcal{D}^{0,\epsilon}} f^0(z), \quad \beta_\epsilon \equiv +\infty, \text{ если } \mathcal{D}^{0,\epsilon} = \emptyset, \quad \beta_0 \equiv \inf_{z \in \mathcal{D}^0} f^0(z) \tag{1.2}$$

Ниже центральным понятием для нас будет понятие обобщенной минимизирующей последовательности (ОМП). Под ОМП понимается последовательность $z^k \in \mathcal{D}$, $k = 1, 2, \ldots$, если

$$f^0(z^k) \to \beta, \quad z^k \in \mathcal{D}^{0,\epsilon^k}, \quad \epsilon^k \to 0, \quad k \to \infty. \tag{1.3}$$

Очевидно, в общей ситуации $\beta \leqslant \beta_0$. В то же время, в задаче (P^0) с неограниченным $\mathcal{D}$ с формальной точки зрения возможно и строгое неравенство $\beta_0 < \beta$. В качестве иллюстрации этого обстоятельства можно указать на пример с разрешимой задачей условной минимизации в [10, с. 173, 174].

Однако, в задаче (P^0) заведомо имеет место равенство $\beta = \beta_0$ по крайней мере в двух случаях: 1) функционал f^0 — сильно выпуклый; 2) множество $\mathcal{D}$ ограничено. В указанных двух случаях для любой ОМП z^i, $i = 1, 2, \ldots$, задачи (P^0) заведомо справедливо предельное соотношение $f^0(z^i) \to f^0(z^0)$, $i \to \infty$.

Замечание 1. В случае сильно выпуклого функционала f^0, а также в случае ограниченности множества $\mathcal{D}$ обобщенная нижняя грань β может быть записана в виде

$$\beta \equiv \{f^0(z^0), \text{ если } z^0 \text{ существует; } +\infty \text{ в противном случае}\}.$$

Итак, в общей ситуации задачи (P^0) при условии ее разрешимости ($Z^0 \neq \emptyset$) мы формально не можем исключать строгого неравенства $\beta < \beta_0$. Различные достаточные условия выполнимости равенства $\beta = \beta_0$ в задачах выпуклого программирования общего вида в банаховых пространствах можно найти в [10, гл. 3]. Однако, далее мы будем, в первую очередь, конструировать для выпуклой задачи общего вида (P^0), строго говоря, не ОМП в указанном выше смысле (1.3), а минимизирующую последовательность $z^k \in \mathcal{D}$, $k = 1, 2, \ldots$, в задаче (P^0), удовлетворяющую соотношениям

$$f^0(z^k) \leqslant \beta_0 + \gamma^k = f^0(z^0) + \gamma^k, \quad z^k \in \mathcal{D}^{0,\epsilon^k}, \quad \gamma^k, \quad \epsilon^k \to 0, \quad k \to \infty, \tag{1.4}$$

которую будем называть ОМП в смысле (1.4) и которая, очевидно, заведомо является ОМП (в смысле (1.3)) в случае $\beta = \beta_0$.

Далее, следуя традиции теории некорректных задач [2, 7], введем, как и в [12], понятие регуляризирующего алгоритма в задаче условной оптимизации (P^0). Предположим, что обобщенная нижняя грань β задачи (P^0) удовлетворяет строгому неравенству $\beta < +\infty$.

Определение 1. Зависящий от параметра $\delta \in (0, \delta_0)$ оператор $R(\cdot, \cdot, \cdot, \cdot, \delta)$, ставящий в соответствие каждой четверке исходных данных f^δ, A^δ, h^δ, g^δ, удовлетворяющих оценкам (1.1), элемент $R(f^\delta, A^\delta, h^\delta, g^\delta, \delta) \equiv z^\delta \in \mathcal{D}$ такой, что

$$f^0(z^\delta) \to \beta, \;\; \|A^0 z^\delta - h^0\| \to 0, \;\; \min_{x \in \mathbb{R}^m_-} |g^0(z^\delta) - x| \to 0, \;\; \delta \to 0,$$

называется регуляризирующим в задаче (P^0).

Замечание 2. Определения регуляризирующих алгоритмов для задач математического программирования с конечным числом функциональных ограничений типа равенства и неравенства можно найти, например, в [2, гл. 9]. Эти определения даны в случае задач первого типа (то есть задач, в которых ищется только нижняя грань, см. [2, гл. 9, с. 802]) и второго типа (то есть задач, в которых ищется и нижняя грань и оптимальный элемент, см. [2, гл. 9, с. 837]). С формальной точки зрения данное выше определение 1 занимает промежуточное положение между двумя указанными выше определениями [2, гл. 9]. В отличие от определения [2, гл. 9, с. 802] в определении 1 речь идет не только о приближении к нижней грани задачи, но и, параллельно, о выполнении «в пределе» ее ограничений с одновременным представлением «сходящихся» при $\delta \to 0$ как по функции, так и по «ограничениям» элементов $R(f^\delta, A^\delta, h^\delta, g^\delta, \delta) \equiv z^\delta \in \mathcal{D}$. В то же время, в отличие от определения [2, гл. 9, с. 837] в определении 1 не идет речь о какой либо сходимости (сильной, слабой) при $\delta \to 0$ самих элементов семейства $R(f^\delta, A^\delta, h^\delta, g^\delta, \delta) \equiv z^\delta \in \mathcal{D}$ к какому либо конкретному элементу, например, к точному решению задачи (P^0) в случае существования последнего. Такая сходимость (сильная, слабая) является уже следствием как того факта, что элементы $R(f^\delta, A^\delta, h^\delta, g^\delta, \delta) \equiv z^\delta \in \mathcal{D}$ при $\delta \to 0$ сходятся одновременно и по функции и по «ограничениям», так и дополнительных свойств исходных данных задачи.

Так как основной целью работы является построение ОМП в задаче (P^0), но семейство $\{z^\delta \in \mathcal{D}\colon \delta \in (0, \delta_0]\}$ из определения 1 не является последовательностью, то помимо введенного выше определения регуляризирующего оператора в задаче (P^0) введем его «след» — определение ОМП-образующего оператора в этой задаче.

Определение 2. Пусть $\delta^k \in (0, \delta_0)$, $k = 1, 2, \ldots,$ — сходящаяся к нулю последовательность положительных чисел. Зависящий от δ^k, $k = 1, 2, \ldots,$ оператор $R(\cdot, \cdot, \cdot, \cdot, \delta^k)$, ставящий в соответствие каждому набору исходных данных $(f^{\delta^k}, A^{\delta^k}, h^{\delta^k}, g^{\delta^k})$, удовлетворяющих оценкам (1.1) при $\delta = \delta^k$, элемент $z^{\delta^k} \in \mathcal{D}$, называется ОМП-образующим в задаче (P^0), если последовательность z^{δ^k}, $k = 1, 2, \ldots,$ есть ОМП в этой задаче. Если ОМП в задаче (P^0) понимается в смысле (1.4), то будем говорить, соответственно, и об ОМП-образующем операторе в задаче (P^0) в смысле (1.4).

Определим далее двойственную задачу

$$V^\delta(\lambda, \mu) \equiv \inf_{z \in \mathcal{D}} L^\delta(z, \lambda, \mu) \to \sup, \quad (\lambda, \mu) \in H \times \mathbb{R}^m_+.$$

В случае сильной выпуклости f^δ значение $V^\delta(\lambda, \mu)$ достигается при любых $(\lambda, \mu) \in H \times \mathbb{R}^m_+$, где $\mathbb{R}^m_+ \equiv \{x = (x_1, \ldots, x_m) \in \mathbb{R}^m\colon x_i \geqslant 0,\, i = 1, \ldots, m\}$, на единственном элементе $z^\delta[\lambda, \mu] \equiv \operatorname{argmin}\{L^\delta(z, \lambda, \mu),\, z \in \mathcal{D}\}$ (в этом случае функционал Лагранжа $L^\delta(z, \lambda, \mu)$, $z \in \mathcal{D}$, является непрерывным и сильно выпуклым). В случае ограниченного множества $\mathcal{D}$ двойственный функционал V^δ, очевидно, определен и конечен для любого элемента $(\lambda, \mu) \in H \times \mathbb{R}^m$, а его значение при $(\lambda, \mu) \in H \times \mathbb{R}^m_+$ достигается на элементах множества $Z^\delta[\lambda, \mu] \equiv \operatorname{Argmin}\{L^\delta(z, \lambda, \mu),\, z \in \mathcal{D}\}$.

§ 2. Регуляризованные ПЛ в выпуклых задачах условной оптимизации

Итак, основной нашей целью является доказательство регуляризованного ПЛ (другими словами, регуляризованной теоремы Куна–Таккера) в задаче условной оптимизации (P^0) с выпуклым (и, вообще говоря, не сильно выпуклым) функционалом качества в случае, когда на множество $\mathcal{D}$ не накладывается условие ограниченности. Будем опираться для достижения указанной цели на результаты статьи [3], в которой аналогичное утверждение было доказано в предположении сильной выпуклости f^0 (и f^δ). При этом нам потребуется, прежде всего, уточнение теоремы 3.1 сходимости метода двойственной регуляризации в [3].

2.1. Двойственная регуляризация в выпуклой задаче условной оптимизации с сильно выпуклым функционалом цели

Предполагаем в данном разделе, что задача (P^0) разрешима, ее решение обозначаем через z^0, а функции $f^\delta,\ \ \delta \in [0,\delta_0]$, считаем сильно выпуклыми на $\mathcal{D}$ с постоянной сильной выпуклости κ, которая не зависит от $\delta \in [0,\delta_0]$. Обозначим, как и в [3], через $(\lambda^{\delta,\alpha},\mu^{\delta,\alpha})$ единственную в $H\times\mathbb{R}^m_+$ точку, дающую на этом множестве максимум сильно вогнутому функционалу

$$R^{\delta,\alpha}(\lambda,\mu)\equiv V^\delta(\lambda,\mu)-\alpha\|\lambda\|^2-\alpha|\mu|^2,\ \ (\lambda,\mu)\in H\times\mathbb{R}^m_+.$$

Подчеркнем, что в данной работе, в отличие от [3], мы рассматриваем непараметрическую задачу (P^0) (то есть задачу, не содержащую параметры в ограничениях). По этой причине мы используем здесь все те же обозначения, что и в [3], но пара (p,r) при этом из всех обозначений [3] удаляется.

Базовым результатом [3] являлось доказательство того факта, что регуляризованные элементы $z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}]$ в случае сильно выпуклого целевого функционала сильно сходятся при $\delta\to 0$ к единственному в этой ситуации решению z^0 задачи (P^0). При этом существенным было именно то, что точки $z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}]$, как точки минимума функции Лагранжа $L^\delta(\cdot,\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)})$, благодаря сильной выпуклости целевого функционала определялись единственным образом. Благодаря этому и в основной теореме 4.2 в [3] точки $z^{\delta^k}[\lambda^k,\mu^k]$, аппроксимирующие решение z^0 (также единственное благодаря сильно выпуклости f^0) определялись корректно, то есть единственным образом.

Пусть выполняется условие согласования

$$\frac{\delta}{\alpha(\delta)}\to 0,\ \ \alpha(\delta)\to 0,\quad \delta\to 0. \tag{2.1}$$

Напомним, что в [3] в предположении независимости постоянной сильной выпуклости κ от δ было показано (см. неравенство (3.25) в [3]), что

$$\|z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}]\|\leqslant L, \tag{2.2}$$

где постоянная $L>0$ не зависит от δ, но зависит, вообще говоря, от κ. Указанная постоянная L была использована затем в формулировке теоремы сходимости метода двойственной регуляризации в [3] (теорема 3.1 в [3]). При получении этой оценки использовалась, в частности, доказанная в [3] оценка (см. оценку (3.17) в [3])

$$f^\delta(z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}])\geqslant K,$$

где постоянная K также не зависит от δ, но зависит, вообще говоря, от κ.

2.1.1. Оценки отклонения приближенных решений по функции и по ограничениям

Итак, получим оценку для величины $\|z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}]\|$ с новой мажорирующаей постоянной (вместо L), которая явным образом зависит от постоянной сильной выпуклости κ и величины F, равномерно по $\delta\in[0,\delta_0]$ ограничивающей снизу на $\mathcal{D}$ функцию f^δ:

$$f^\delta(z)\geqslant F\quad \forall\, z\in\mathcal{D},\ \ \delta\in[0,\delta_0]. \tag{2.3}$$

Как результат, мы уточним теорему 3.1 в [3], заменив в ней постоянную L новой мажорирующей постоянной, а затем применим полученный результат для получения регуляризованного ПЛ в выпуклой задаче (P^0), функционал качества которой, вообще говоря, не является сильно выпуклым. На этом пути главной целью всех построений данного раздела является получение оценок точности приближения элементов $z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}]$ к решению z^0 задачи (P^0) по функционалу и по ограничениям через величины δ, $\alpha(\delta)$, κ, $\|z^0\|$, $f^0(z^0)$ и F в предположении (2.3).

Для достижения указанной цели нам потребуются следующие оценки из [3]. Во-первых, это оценка (3.23) из [3] (постоянная C в ней это постоянная из (1.1)))

$$\begin{gathered} f^\delta(z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}]) \leqslant f^0(z^0) + C\delta(1+\|z^0\|^2) + C\delta\|\lambda^{\delta,\alpha(\delta)}\|(2+\|z^0\|) + \\ + C\delta|\mu^{\delta,\alpha(\delta)}|(1+\|z^0\|^2) \equiv f^0(z^0)+\psi(\delta) \qquad \forall\,\delta\in(0,\delta_0). \end{gathered} \tag{2.4}$$

Она является следствием неравенств, во втором из которых используются оценки (1.1) (напомним, что $z^\delta[\lambda,\mu]\equiv \operatorname{argmin}\{L^\delta(z,\lambda,\mu),\ z\in\mathcal{D}\}$)

$$L^\delta(z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}],\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}) \leqslant L^\delta(z^0,\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}) \leqslant \tag{2.5}$$

$$f^0(z^0) + [f^\delta(z^0)-f^0(z^0)] + \|\lambda^{\delta,\alpha(\delta)}\|\|A^\delta z^0 - h^\delta\| + |\mu^{\delta,\alpha(\delta)}||g^\delta(z^0)-g^0(z^0)| \leqslant$$

$$f^0(z^0) + C\delta(1+\|z^0\|^2) + C\delta\|\lambda^{\delta,\alpha(\delta)}\|(2+\|z^0\|) + C\delta|\mu^{\delta,\alpha(\delta)}|(1+\|z^0\|^2)$$

с учетом неравенства (см. неравенство (3.14) в [3])

$$\langle (A^\delta z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}] - h^\delta, g^\delta(z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}])), (\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)})\rangle \geqslant 0. \tag{2.6}$$

Во вторых, это оценка (3.18) из [3]

$$\alpha(\delta)\sqrt{\|\lambda^{\delta,\alpha(\delta)}\|^2+|\mu^{\delta,\alpha(\delta)}|^2} \leqslant C_2\delta + \sqrt{C_2^2\delta^2 - 8\alpha(\delta)K(\delta)} \equiv \phi(\delta,\alpha(\delta)) \quad \forall\,\delta\in(0,\delta_0), \tag{2.7}$$

с $C_2=\sqrt{2}C_1(1+\|z^0\|^2)$ (см. (3.16) в [3]), $C_1 = 3/2C$ (см. (3.15) в [3]), а также с $K(\delta)\equiv$ $\equiv f^\delta(z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}]) - f^0(z^0) - C\delta(1+\|z^0\|^2)$.

Наконец, в третьих, это оценки (3.20) из [3]

$$\begin{gathered} \|A^\delta z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}] - h^\delta\| \leqslant \phi(\delta,\alpha(\delta)) \quad \forall\,\delta\in(0,\delta_0), \\ g_i^\delta(z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}]) \leqslant \phi(\delta,\alpha(\delta))\ \forall\,\delta\in(0,\delta_0), \quad \phi(\delta,\alpha(\delta))\to 0,\ \delta\to 0,\ i=1,\dots,m. \end{gathered} \tag{2.8}$$

С целью получения анонсированной оценки заметим, прежде всего, что, в силу известного неравенства [2, теорема 8.2.10] для сильно выпуклых функционалов можем записать

$$\kappa\|z^0 - z^*_\delta\|^2 \leqslant f^\delta(z^0) - f^\delta(z^*_\delta) \leqslant f^\delta(z^0) - F, \tag{2.9}$$

$$\kappa\|z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}] - z^*_\delta\|^2 \leqslant f^\delta(z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}]) - f^\delta(z^*_\delta) \leqslant f^\delta(z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}]) - F.$$

где $z^*_\delta\equiv\operatorname{argmin}\{f^\delta(z),\ z\in\mathcal{D}\}$.

Тогда из первой оценки (2.9) выводим оценку

$$\|z^0 - z^*_\delta\| \leqslant \frac{\sqrt{f^\delta(z^0)-F}}{\sqrt{\kappa}} \leqslant \frac{\sqrt{f^0(z^0)+C\delta(1+\|z^0\|^2)-F}}{\sqrt{\kappa}},$$

из которой получаем

$$\|z^*_\delta\| \leqslant \|z^0\| + \frac{\sqrt{f^0(z^0)+C\delta(1+\|z^0\|^2)-F}}{\sqrt{\kappa}}.$$

В свою очередь, из второй оценки (2.9) аналогичным образом и с учетом последней оценки получаем неравенство

$$\|z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}]\| \leqslant \|z^0\| + \frac{\sqrt{f^0(z^0)+C\delta(1+\|z^0\|^2)-F}}{\sqrt{\kappa}} + \frac{\sqrt{f^\delta(z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}])-F}}{\sqrt{\kappa}},$$

из которого в силу оценки (2.4) следует

$$\begin{gathered}\|z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}]\| \leqslant \|z^0\| + \frac{\sqrt{f^0(z^0)+\delta(1+\|z^0\|^2)-F}}{\sqrt{\kappa}} + \\ + \frac{\sqrt{f^0(z^0)+\psi(\delta)-F}}{\sqrt{\kappa}} \equiv \|z^0\| + \frac{\xi(\delta,\alpha(\delta))}{\sqrt{\kappa}} \quad \forall\,\delta\in(0,\delta_0).\end{gathered} \tag{2.10}$$

Последняя оценка, правая часть которой явным образом зависит от постоянной сильной выпуклости κ, с определенной в (2.4) величиной $\psi(\delta)$, и является искомой оценкой в предположении (2.3) для величины $\|z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}]\|$. Величину же $\|z^0\| + \frac{\xi(\delta,\alpha(\delta))}{\sqrt{\kappa}}$ в правой части неравенства (2.10) можно взять вместо величины L в теореме 3.1 в [3] при указанном предположении (см. ниже теорему 1).

Далее, оценивая величину $\phi(\delta,\alpha(\delta))$, можем записать (напомним, что $C_2 = C(3\sqrt{2}/2)\cdot$ $\cdot(1+\|z^0\|^2)$)

$$\begin{gathered}\phi(\delta,\alpha(\delta)) \equiv C_2\delta + \sqrt{C_2^2\delta^2 - 8\alpha(\delta)K(\delta)} \equiv \\ \equiv C_2\delta + \sqrt{C_2^2\delta^2 - 8\alpha(\delta)(f^\delta(z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}]) - f^0(z^0) - C\delta(1+\|z^0\|^2))} \leqslant \\ \leqslant C_2\delta + \sqrt{C_2^2\delta^2 - 8\alpha(\delta)(F - f^0(z^0) - C\delta(1+\|z^0\|^2))} \quad \forall\,\delta\in(0,\delta_0).\end{gathered} \tag{2.11}$$

Таким образом, учитывая оценку (2.7), получаем оценку

$$\alpha(\delta)\sqrt{\|\lambda^{\delta,\alpha(\delta)}\|^2 + |\mu^{\delta,\alpha(\delta)}|^2} \leqslant C_2\delta + \sqrt{C_2^2\delta^2 - 8\alpha(\delta)(F - f^0(z^0) - C\delta(1+\|z^0\|^2))}\,\forall\,\delta\in(0,\delta_0). \tag{2.12}$$

Наконец, оценка (2.4) в совокупности с оценками (2.7), (2.12) и условием согласования (2.1) обеспечивает и оценку величины $\psi(\delta)$ через величины δ, $\alpha(\delta)$, κ, $\|z^0\|$, $f^0(z^0)$ и F в предположении (2.3).

Подытоживая проведенные рассуждения, получаем с учетом оценок (1.1), (2.8), (2.10) оценки для любого $\delta\in(0,\delta_0)$

$$\begin{gathered}\|A^0 z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}] - h^0\| \leqslant \\ \leqslant \|A^\delta z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}] - h^\delta\| + \|A^0 z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}] - \\ - A^\delta z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}]\| + \|h^0 - h^\delta\| \leqslant \\ \leqslant \|A^\delta z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}] - h^\delta\| + C\delta(1 + \|z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}]\|) + C\delta \leqslant \\ \leqslant \phi(\delta,\alpha(\delta)) + C\delta(2 + \|z^0\| + \frac{\xi(\delta,\alpha(\delta))}{\sqrt{\kappa}}),\end{gathered} \tag{2.13}$$

$$\begin{gathered}g_i^0(z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}]) \leqslant \\ \leqslant g_i^\delta(z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}]) + g_i^0(z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}]) - g_i^\delta(z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}]) \leqslant \\ \leqslant g_i^\delta(z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}]) + C\delta(1 + \|z^\delta[\lambda^{\delta,\alpha(\delta)},\mu^{\delta,\alpha(\delta)}]\|^2) \leqslant \\ \leqslant \phi(\delta,\alpha(\delta)) + C\delta(1 + (\|z^0\| + \frac{\xi(\delta,\alpha(\delta))}{\sqrt{\kappa}})^2).\end{gathered} \tag{2.14}$$

Одновременно, оценки (1.1), (2.4), (2.10) дают возможность записать оценку для любого $\delta \in (0, \delta_0)$

$$\begin{aligned} & f^0(z^\delta[\lambda^{\delta,\alpha(\delta)}, \mu^{\delta,\alpha(\delta)}]) \leqslant \\ \leqslant f^0(z^0) + \psi(\delta) + f^0(z^\delta[\lambda^{\delta,\alpha(\delta)}, \mu^{\delta,\alpha(\delta)}]) - f^\delta(z^\delta[\lambda^{\delta,\alpha(\delta)}, \mu^{\delta,\alpha(\delta)}]) \leqslant \\ \leqslant f^0(z^0) + \psi(\delta) + C\delta(1 + \|z^\delta[\lambda^{\delta,\alpha(\delta)}, \mu^{\delta,\alpha(\delta)}]\|^2) \leqslant \\ \leqslant f^0(z^0) + \psi(\delta) + C\delta(1 + (\|z^0\| + \frac{\xi(\delta,\alpha(\delta))}{\sqrt{\kappa}})^2). \end{aligned} \tag{2.15}$$

Полученные оценки (2.13), (2.14), (2.15), в которых величина $\psi(\delta)$ задается в (2.4), в совокупности с оценками (2.7), (2.11), (2.12) и условием согласования (2.1), представляют собой явные оценки точности приближения элементов $z^\delta[\lambda^{\delta,\alpha(\delta)}, \mu^{\delta,\alpha(\delta)}]$ к решению z^0 задачи (P^0) по ограничениям и по функционалу качества через величины δ, $\alpha(\delta)$, κ, $\|z^0\|$, $f^0(z^0)$ и F в предположении (2.3).

Замечание 3. Заметим, наконец, что если бы мы вместо оценки (2.10) имели бы, как и в [3], оценку (2.2) (см. оценку (3.25) в [3]), то очевидно, вместо оценок (2.13), (2.14), (2.15) мы бы получили, как и в [3], оценки

$$\begin{aligned} \|A^0 z^\delta[\lambda^{\delta,\alpha(\delta)}, \mu^{\delta,\alpha(\delta)}] - h^0\| \leqslant \phi(\delta,\alpha(\delta)) + C\delta(2+L), \\ g_i^0(z^\delta[\lambda^{\delta,\alpha(\delta)}, \mu^{\delta,\alpha(\delta)}]) \leqslant \phi(\delta,\alpha(\delta)) + C\delta(1+L^2), \\ f^0(z^\delta[\lambda^{\delta,\alpha(\delta)}, \mu^{\delta,\alpha(\delta)}]) \leqslant f^0(z^0) + \psi(\delta) + C\delta(1+L^2). \end{aligned}$$

2.1.2. Уточненная теорема сходимости метода двойственной регуляризации в выпуклом программировании с сильно выпуклым функционалом цели

Полученные выше оценки (2.13), (2.14), (2.15) позволяют нам уточнить теорему 3.1 сходимости метода двойственной регуляризации в [3] (здесь нам будет нужна только часть ее утверждений).

Теорема 1 (Регуляризирующий двойственный алгоритм). *Пусть задача (P^0) разрешима. Тогда, вне зависимости от того, разрешима или нет двойственная к (P^0) задача, в предположении (2.3) и при выполнении условия согласования (2.1) выполняются соотношения*

$$\begin{aligned} \alpha(\delta)\|(\lambda^{\delta,\alpha(\delta)}, \mu^{\delta,\alpha(\delta)})\|^2 \to 0, \quad \delta \to 0, \\ f^0(z^\delta[\lambda^{\delta,\alpha(\delta)}, \mu^{\delta,\alpha(\delta)}]) \leqslant f^0(z^0) + \psi_1(\delta), \\ \psi_1(\delta) \equiv \psi(\delta) + C\delta(1 + (\|z^0\| + \frac{\xi(\delta,\alpha(\delta))}{\sqrt{\kappa}})^2) \to 0, \quad \delta \to 0, \end{aligned} \tag{2.16}$$

$$\|A^0 z^\delta[\lambda^{\delta,\alpha(\delta)}, \mu^{\delta,\alpha(\delta)}] - h^0\| \leqslant \phi(\delta,\alpha(\delta)) + C\delta(2 + \|z^0\| + \frac{\xi(\delta,\alpha(\delta))}{\sqrt{\kappa}}) \to 0, \tag{2.17}$$

$$g_i^0(z^\delta[\lambda^{\delta,\alpha(\delta)}, \mu^{\delta,\alpha(\delta)}]) \leqslant \phi(\delta,\alpha(\delta)) + C\delta(1 + (\|z^0\| + \frac{\xi(\delta,\alpha(\delta))}{\sqrt{\kappa}})^2) \to 0, \tag{2.18}$$

$$\langle(\lambda^{\delta,\alpha(\delta)}, \mu^{\delta,\alpha(\delta)}), (A^\delta z^\delta[\lambda^{\delta,\alpha(\delta)}, \mu^{\delta,\alpha(\delta)}] - h^\delta, g^\delta(z^\delta[\lambda^{\delta,\alpha(\delta)}, \mu^{\delta,\alpha(\delta)}]))\rangle \to 0, \ \delta \to 0,$$

где C, $C_2 = C(3\sqrt{2}/2)(1+\|z^0\|^2)$ — независящие от δ постоянные (см. формулы (1.1), (2.7)), а величины $\psi(\delta)$, $\phi(\delta,\alpha(\delta))$, $\xi(\delta,\alpha(\delta))$ определены в (2.4), (2.7), (2.10) соответственно.

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

$$\|z^\delta[\lambda^{\delta,\alpha(\delta)}, \mu^{\delta,\alpha(\delta)}] - z^0\| \to 0, \quad \delta \to 0. \tag{2.19}$$

Другими словами, вне зависимости от того, разрешима или нет двойственная задача, зависящий от параметра $\delta \in (0, \delta_0)$ оператор (алгоритм) $R(\cdot,\cdot,\cdot,\cdot,\delta)$, ставящий в соответствие каждой четверке исходных данных f^δ, A^δ, h^δ, g^δ, удовлетворяющих оценкам (1.1), элемент $R(f^\delta, A^\delta, h^\delta, g^\delta, \delta) = z^\delta[\lambda^{\delta,\alpha(\delta)}, \mu^{\delta,\alpha(\delta)}]$, является регуляризирующим в смысле определения 1, причем в случае субдифференцируемости f^0 в точках $\mathcal{D}$ имеет место и сильная сходимость (2.19). Если же такой субдифференцируемости нет, то можно говорить лишь о слабой сходимости $z^\delta[\lambda^{\delta,\alpha(\delta)}, \mu^{\delta,\alpha(\delta)}]$ к z^0 при $\delta \to 0$.

В случае ограниченного $\mathcal{D}$ и, как следствие, оценки $\|z^\delta[\lambda^{\delta,\alpha(\delta)}, \mu^{\delta,\alpha(\delta)}]\| \leqslant L$, $\delta \in (0, \delta_0)$ (например, с $L = \sup\limits_{z\in\mathcal{D}} \|z\|$), величину $\|z^0\| + \frac{\xi(\delta,\alpha(\delta))}{\sqrt{\kappa}}$ в (2.16), (2.17), (2.18) следует заменить на L (см. замечание 3).

Замечание 4. В формулировке теоремы 3.1 в [3] были пропущены полученные при ее доказательстве слагаемые: $\psi(\delta)$ (см. неравенства (3.23), (3.26) в [3]) — в выражении для $\psi_1(\delta)$ в формуле, соответствующей (2.16); $C\delta(2 + L)$ — в правой части неравенства, соответствующего (2.17); $C\delta(1 + L^2)$ — в правой части неравенства, соответствующего (2.18). В приводимой здесь формулировке все указанные слагаемые восстановлены с заменой L на $\|z^0\| + \frac{\xi(\delta,\alpha(\delta))}{\sqrt{\kappa}}$.

2.2. Двойственная регуляризация в задаче выпуклого программирования с выпуклым функционалом цели

В предыдущем разделе была описана процедура двойственной регуляризации в задаче выпуклого программирования с сильно выпуклым функционалом качества и получены оценки сходимости этой процедуры по функции и по ограничениям. В настоящем разделе мы покажем как эти оценки могут быть применены к задаче выпуклого программирования с выпуклым функционалом качества, который не обязательно является сильно выпуклым. А именно, мы организуем процесс двойственной регуляризации с помощью двух регуляризирующих добавков, один из которых находится, как и ранее, в двойственной задаче, а другой — в функционале качества. При этом существенную роль будут играть полученные в предыдущем разделе оценки (2.13), (2.14), (2.15). Они в совокупности с оценками (2.7), (2.11), (2.12) и условием согласования (2.1) представляют собою явные оценки для отклонений приближенных решений $z^\delta[\lambda^{\delta,\alpha(\delta)}, \mu^{\delta,\alpha(\delta)}]$, аппроксимирующих точное решение z^0 задачи (P^0) с сильно выпуклым функционалом f^0, по функции и по ограничениям через величины δ, $\alpha(\delta)$, κ, $\|z^0\|$, $f^0(z^0)$ и F в предположении (2.3).

Итак, рассматриваем сформулированную в § 1 задачу (P^0), функционал $f^0\colon \mathcal{D} \to \mathbb{R}^1$ которой является выпуклым на выпуклом замкнутом (возможно неограниченном) множестве $\mathcal{D}$, точнее, выпуклыми являются как точный функционал f^0, так и его возмущение f^δ. Пусть решение задачи (P^0) (единственное, в случае строгой (сильной) выпуклости f^0) существует, то есть $Z^0 \neq \emptyset$. Заметим, что из существования ограниченной ОМП в задаче (P^0), очевидно, следует и разрешимость последней.

Далее будем обозначать через z^0 одно конкретное из решений задачи (P^0), а именно, ее нормальное, то есть минимальное по норме, решение. Предположим, что выполняется условие (2.3), в котором величина F не зависит от $\delta \in [0, \delta_0]$.

Рассмотрим семейство регуляризованных задач

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

с их единственными и одновременно нормальными решениями z^δ_ε, $\varepsilon > 0$, $\delta \geqslant 0$, $z^0_0 \equiv z^0$. Очевидно, в каждой из задач (P^δ_ε) с $\varepsilon > 0$ функционал качества $f^\delta(\cdot) + \varepsilon\|\cdot\|^2$ является непрерывным и сильно выпуклым на $\mathcal{D}$ с постоянной сильной выпуклости $\varepsilon > 0$, причем в силу условия (2.3) справедлива оценка

$$f^\delta(z) + \varepsilon\|z\|^2 \geqslant F \quad \forall\, z \in \mathcal{D}, \quad \delta \in [0, \delta_0], \quad \varepsilon \geqslant 0. \tag{2.20}$$

Как известно из результатов обычной теории тихоновской стабилизации [2, § 4, теорема 4], в этом случае для любой зависимости $\varepsilon(\delta) > 0$, $\varepsilon(\delta) \to 0$, $\delta \to 0$, имеет место предельное соотношение $z^0_{\varepsilon(\delta)} \to z^0$, $\delta \to 0$, где, как уже определено выше, z^0 — нормальное решение исходной задачи (P^0).

Введем регулярный функционал Лагранжа

$$L^{\delta,\varepsilon}(z, \lambda, \mu) \equiv f^\delta(z) + \varepsilon\|z\|^2 + \langle \lambda, A^\delta z - h^\delta\rangle + \langle \mu, g^\delta(z)\rangle, \quad z \in \mathcal{D}, \quad \lambda \in H, \quad \mu \in \mathbb{R}^m_+,$$

и двойственную задачу

$$V^{\delta,\varepsilon}(\lambda, \mu) \equiv \inf_{z\in\mathcal{D}} L^{\delta,\varepsilon}(z, \lambda, \mu) \to \sup, \quad (\lambda, \mu) \in H \times \mathbb{R}^m_+, \quad V^{\delta,0}(\lambda, \mu) \equiv V^\delta(\lambda, \mu).$$

Пусть $z^{\delta,\varepsilon}[\lambda, \mu] \equiv \operatorname{argmin}\{L^{\delta,\varepsilon}(z, \lambda, \mu),\ z \in \mathcal{D}\}$ при условии, что $(\lambda, \mu) \in H \times \mathbb{R}^m_+$, $\varepsilon > 0$, $\delta \geqslant 0$. Обозначим через $(\lambda^{\delta,\alpha(\delta),\varepsilon}, \mu^{\delta,\alpha(\delta),\varepsilon}) \in H \times \mathbb{R}^m_+$ решение регуляризованной двойственной задачи

$$V^{\delta,\varepsilon}(\lambda, \mu) - \alpha(\delta)\|(\lambda, \mu)\|^2 \to \max, \quad (\lambda, \mu) \in H \times \mathbb{R}^m_+,$$

с условием согласования (2.1). Пусть помимо условия согласования (2.1) выполняется и еще одно подобное условие

$$\frac{\delta}{\varepsilon(\delta)} \to 0, \quad \varepsilon(\delta) \to 0, \quad \delta \to 0. \tag{2.21}$$

Сформулированные условия в силу оценок (2.13), (2.14), (2.15), с учетом оценки (2.20), позволяют записать

$$\begin{gathered}\|A^0 z^{\delta,\varepsilon(\delta)}[\lambda^{\delta,\alpha(\delta),\varepsilon(\delta)}, \mu^{\delta,\alpha(\delta),\varepsilon(\delta)}] - h^0 - p\| \leqslant \\ \leqslant \phi(\delta, \alpha(\delta), \varepsilon(\delta)) + C\delta(2 + \|z^0_{\varepsilon(\delta)}\| + \frac{\xi(\delta, \alpha(\delta), \varepsilon(\delta))}{\sqrt{\varepsilon(\delta)}}),\end{gathered} \tag{2.22}$$

$$\begin{gathered}g^0_i(z^{\delta,\varepsilon(\delta)}[\lambda^{\delta,\alpha(\delta),\varepsilon(\delta)}, \mu^{\delta,\alpha(\delta),\varepsilon(\delta)}]) \leqslant \\ \leqslant \phi(\delta, \alpha(\delta), \varepsilon(\delta)) + C\delta(1 + (\|z^0_{\varepsilon(\delta)}\| + \frac{\xi(\delta, \alpha(\delta), \varepsilon(\delta))}{\sqrt{\varepsilon(\delta)}})^2), \quad i = 1, \ldots, m,\end{gathered} \tag{2.23}$$

$$\begin{gathered}f^0(z^{\delta,\varepsilon(\delta)}[\lambda^{\delta,\alpha(\delta),\varepsilon(\delta)}, \mu^{\delta,\alpha(\delta),\varepsilon(\delta)}]) + \varepsilon(\delta)\|z^{\delta,\varepsilon(\delta)}[\lambda^{\delta,\alpha(\delta),\varepsilon(\delta)}, \mu^{\delta,\alpha(\delta),\varepsilon(\delta)}\|^2 \leqslant \\ \leqslant f^0(z^0_{\varepsilon(\delta)}) + \varepsilon(\delta)\|z^0_{\varepsilon(\delta)}\|^2 + \psi(\delta, \varepsilon(\delta)) + C\delta(1 + (\|z^0_{\varepsilon(\delta)}\| + \frac{\xi(\delta, \alpha(\delta), \varepsilon(\delta))}{\sqrt{\varepsilon(\delta)}})^2),\end{gathered} \tag{2.24}$$

$$\|z^{\delta,\varepsilon(\delta)}[\lambda^{\delta,\alpha(\delta),\varepsilon(\delta)}, \mu^{\delta,\alpha(\delta),\varepsilon(\delta)}]\| \leqslant \|z^0_{\varepsilon(\delta)}\| + \frac{\xi(\delta, \alpha(\delta), \varepsilon(\delta))}{\sqrt{\varepsilon(\delta)}}, \tag{2.25}$$

где (см. формулу (2.11))

$$\begin{gathered}\phi(\delta, \alpha(\delta), \varepsilon(\delta)) \leqslant \\ \leqslant C_2(\varepsilon(\delta))\delta + \sqrt{(C_2(\varepsilon(\delta)))^2\delta^2 - 8\alpha(\delta)(F - f^0(z^0_{\varepsilon(\delta)}) - \varepsilon(\delta)\|z^0_{\varepsilon(\delta)}\|^2 - C\delta(1 + \|z^0_{\varepsilon(\delta)}\|^2))},\end{gathered} \tag{2.26}$$

$$\psi(\delta,\varepsilon(\delta))\equiv \tag{2.27}$$
$$\equiv C\delta(1+\|z^0_{\varepsilon(\delta)}\|^2)+C\delta(\|z^0_{\varepsilon(\delta)}\|+2)\|\lambda^{\delta,\alpha(\delta),\varepsilon(\delta)}\|+C\delta(1+\|z^0_{\varepsilon(\delta)}\|^2)|\mu^{\delta,\alpha(\delta),\varepsilon(\delta)}|\to 0,\quad \delta\to 0,$$

$$\xi(\delta,\alpha(\delta),\varepsilon(\delta))\equiv$$
$$\equiv\sqrt{f^0(z^0_{\varepsilon(\delta)})+\varepsilon(\delta)\|z^0_{\varepsilon(\delta)}\|^2+\delta(1+\|z^0_{\varepsilon(\delta)}\|^2)-F}+ \tag{2.28}$$
$$+\sqrt{f^0(z^0_{\varepsilon(\delta)})+\varepsilon(\delta)\|z^0_{\varepsilon(\delta)}\|^2+\psi(\delta,\varepsilon(\delta))-F},$$

$$\alpha(\delta)\sqrt{\|\lambda^{\delta,\alpha(\delta),\varepsilon(\delta)}\|^2+|\mu^{\delta,\alpha(\delta),\varepsilon(\delta)}|^2}\leqslant \tag{2.29}$$
$$\leqslant C_2(\varepsilon(\delta))\delta+\sqrt{(C_2(\varepsilon(\delta)))^2\delta^2-8\alpha(\delta)(F-f^0(z^0_{\varepsilon(\delta)})-\varepsilon(\delta)\|z^0_{\varepsilon(\delta)}\|^2-C\delta(1+\|z^0_{\varepsilon(\delta)}\|^2))}.$$

Здесь постоянная $C>0$ берется, как и выше, из (1.1) и не зависит от δ, $C_2(\varepsilon(\delta))=$ $=C(3\sqrt{2}/2)(1+\|z^0_{\varepsilon(\delta)}\|^2)$.

Полученные оценки (2.22), (2.23), (2.24) в совокупности с оценкой (2.26), равенствами (2.27), (2.28) и оценкой (2.29), а также с учетом условий согласования (2.1), (2.21) и предельного соотношения $z^0_{\varepsilon(\delta)}\to z^0$, $\delta\to 0$, где z^0 — нормальное решение исходной задачи (P^0), дают

$$A^0z^{\delta,\varepsilon(\delta)}[\lambda^{\delta,\alpha(\delta),\varepsilon(\delta)},\mu^{\delta,\alpha(\delta),\varepsilon(\delta)}]-h^0\to 0, \tag{2.30}$$
$$g^0_i(z^{\delta,\varepsilon(\delta)}[\lambda^{\delta,\alpha(\delta),\varepsilon(\delta)},\mu^{\delta,\alpha(\delta),\varepsilon(\delta)}])\leqslant\varphi(\delta)\to 0,\quad i=1,\dots,m,$$
$$f^0(z^{\delta,\varepsilon(\delta)}[\lambda^{\delta,\alpha(\delta),\varepsilon(\delta)},\mu^{\delta,\alpha(\delta),\varepsilon(\delta)}])\leqslant f^0(z^0)+\varphi_1(\delta),\quad \varphi_1(\delta)\to 0,$$
$$\alpha(\delta)\|\lambda^{\delta,\alpha(\delta),\varepsilon(\delta)}\|\to 0,\quad \alpha(\delta)|\mu^{\delta,\alpha(\delta),\varepsilon(\delta)}|\to 0,\quad \delta\to 0, \tag{2.31}$$

откуда следует, что любая последовательность $z^{\delta^k,\varepsilon(\delta^k)}[\lambda^{\delta^k,\alpha(\delta^k),\varepsilon(\delta^k)},\mu^{\delta^k,\alpha(\delta^k),\varepsilon(\delta^k)}]$ с $\delta^k\geqslant 0$, $\delta^k\to 0$, $k\to\infty$, представляет собою ОМП в смысле (1.4) в рассматриваемой задаче (P^0), причем $\alpha(\delta^k)\|(\lambda^{\delta^k,\alpha(\delta^k),\varepsilon(\delta^k)},\mu^{\delta^k,\alpha(\delta^k),\varepsilon(\delta^k)})\|\to 0$, $k\to\infty$. При этом все слабые предельные точки построенной ОМП в случае существования таковых (они заведомо существуют, если она ограничена) являются решениями задачи (P^0).

Замечание 5. В случае ограниченного $\mathcal{D}$ и оценки $\|z^{\delta,\varepsilon(\delta)}[\lambda^{\delta,\alpha(\delta),\varepsilon(\delta)},\mu^{\delta,\alpha(\delta),\varepsilon(\delta)}]\|\leqslant L$, $\delta\in(0,\delta_0)$ (см. замечание 3), величину $\|z^0_{\varepsilon(\delta)}\|+\frac{\xi(\delta,\alpha(\delta),\varepsilon(\delta))}{\sqrt{\varepsilon(\delta)}}$ в (2.22), (2.23), (2.24), (2.25) следует заменить на L. Как следствие, в этом случае условие согласования (2.21) становится излишним, можно формально положить $\varepsilon(\delta)=\varepsilon$ и предельные соотношения (2.30), (2.31) можно переписать в виде

$$A^0z^{\delta,\varepsilon}[\lambda^{\delta,\alpha(\delta),\varepsilon},\mu^{\delta,\alpha(\delta),\varepsilon}]-h^0\to 0,\qquad g^0_i(z^{\delta,\varepsilon}[\lambda^{\delta,\alpha(\delta),\varepsilon},\mu^{\delta,\alpha(\delta),\varepsilon}])\leqslant\varphi(\delta,\varepsilon)\to 0,\quad i=1,\dots,m,$$
$$f^0(z^{\delta,\varepsilon}[\lambda^{\delta,\alpha(\delta),\varepsilon},\mu^{\delta,\alpha(\delta),\varepsilon}])\leqslant f^0(z^0)+\varphi_1(\delta,\varepsilon),\qquad \varphi_1(\delta,\varepsilon)\to 0,$$
$$\alpha(\delta)\|(\lambda^{\delta,\alpha(\delta),\varepsilon},\mu^{\delta,\alpha(\delta),\varepsilon})\|\to 0,\quad \delta\to 0,\quad \varepsilon\to 0.$$

Итак, в настоящем разделе построена ОМП в смысле (1.4) в задаче (P^0) или, другими словами, доказана следующая теорема.

Теорема 2. *Пусть $Z^0\neq\emptyset$ и выполняются условия согласования* (2.1) *и* (2.21), $\delta^k\in(0,\delta_0)$, $k=1,2,\dots,$ *— последовательность сходящихся к нулю положительных чисел. Пусть также выполняется условие* (2.3). *Тогда оператор $R(\cdot,\cdot,\cdot,\cdot,\delta^k)$, ставящий в соответствие любому набору исходных данных $\{f^{\delta^k},A^{\delta^k},h^{\delta^k},g^{\delta^k}\}$, удовлетворяющих оценкам* (1.1) *при*

$\delta = \delta^k$, элемент $R(f^{\delta^k}, A^{\delta^k}, h^{\delta^k}, g^{\delta^k}, \delta^k) \equiv z^{\delta^k,\varepsilon(\delta^k)}[\lambda^{\delta^k,\alpha(\delta^k),\varepsilon(\delta^k)}, \mu^{\delta^k,\alpha(\delta^k),\varepsilon(\delta^k)}]$, является ОМП-образующим в смысле (1.4), причем $\alpha(\delta^k)\|(\lambda^{\delta^k,\alpha(\delta^k),\varepsilon(\delta^k)}, \mu^{\delta^k,\alpha(\delta^k),\varepsilon(\delta^k)})\| \to 0,\ k \to \infty$. Одновременно полученные выше оценки (2.22), (2.23), (2.24) (в совокупности с оценкой (2.26), равенствами (2.27), (2.28) и оценкой (2.29), а также с учетом условий согласования (2.1), (2.21) и предельного соотношения $z^0_{\varepsilon(\delta)} \to z^0,\ \delta \to 0$) естественно трактовать как оценки точности построенного ОМП по ограничениям и функционалу.

В случае ограниченного $\mathcal{D}$ условие согласования (2.21), а также условие ограниченности снизу (2.3) становятся излишними. В этом случае ОМП-образующий оператор $R(\cdot,\cdot,\cdot,\cdot,\delta^k)$, ставящий в соответствие любому набору исходных данных $\{f^{\delta^k}, A^{\delta^k}, h^{\delta^k}, g^{\delta^k}\}$, удовлетворяющих оценкам (1.1) при $\delta = \delta^k$, элемент из $\mathcal{D}$, задается равенством

$$R(f^{\delta^k}, A^{\delta^k}, h^{\delta^k}, g^{\delta^k}, \delta^k) \equiv z^{\delta^k,\varepsilon^k}[\lambda^{\delta^k,\alpha(\delta^k),\varepsilon^k}, \mu^{\delta^k,\alpha(\delta^k),\varepsilon^k}],$$

где $\varepsilon^k,\ k = 1, 2, \ldots,$ — произвольная последовательность сходящихся к нулю положительных чисел.

Очевидно, при дополнительном предположении $\beta = \beta_0$ построенная ОМП в смысле (1.4) о задаче (P^0) является одновременно и ОМП в смысле (1.3). Это позволит нам далее доказать в терминах ОМП регуляризованный ПЛ для задачи (P^0), являющийся одновременно регуляризирующим алгоритмом в задаче в смысле определения 1.

2.3. Регуляризованный ПЛ в задаче условной оптимизации при условии равенства двух ее значений

Покажем, что в случае $\beta = \beta_0$ при условии (2.3) выполняется и предельное соотношение

$$\langle(\lambda^\delta, \mu^\delta), (A^\delta z^{\delta,\varepsilon(\delta)}[\lambda^\delta, \mu^\delta] - h^\delta, g^\delta(z^{\delta,\varepsilon(\delta)}[\lambda^\delta, \mu^\delta]))\rangle \to 0, \quad \delta \to 0. \tag{2.32}$$

Здесь и ниже принято обозначение $(\lambda^\delta, \mu^\delta) \equiv (\lambda^{\delta,\alpha(\delta),\varepsilon(\delta)}, \mu^{\delta,\alpha(\delta),\varepsilon(\delta)})$. Для этого сначала заметим, что в данном случае из трех соотношений (2.30) следует предельное соотношение

$$f^0(z^{\delta,\varepsilon(\delta)}[\lambda^\delta, \mu^\delta]) \to f^0(z^0), \quad \delta \to 0. \tag{2.33}$$

Одновременно, в силу (2.5) можем записать в случае задачи $(P^\delta_{\varepsilon(\delta)})$

$$f^\delta(z^{\delta,\varepsilon(\delta)}[\lambda^\delta, \mu^\delta]) + \varepsilon(\delta)\|z^{\delta,\varepsilon(\delta)}[\lambda^\delta, \mu^\delta]\|^2 + \langle(\lambda^\delta, \mu^\delta), (A^\delta z^{\delta,\varepsilon(\delta)}[\lambda^\delta, \mu^\delta] - h^\delta, g^\delta(z^{\delta,\varepsilon(\delta)}[\lambda^\delta, \mu^\delta]))\rangle \leqslant$$
$$\leqslant f^0(z^0_{\varepsilon(\delta)}) + \varepsilon(\delta)\|z^0_{\varepsilon(\delta)}\|^2 + C\delta(1 + \|z^0_{\varepsilon(\delta)}\|^2) + C\delta\|\lambda^\delta\|(2 + \|z^0_{\varepsilon(\delta)}\|) + C\delta|\mu^\delta|(1 + \|z^0_{\varepsilon(\delta)}\|^2) \equiv$$
$$\equiv f^0(z^0_{\varepsilon(\delta)}) + \varepsilon(\delta)\|z^0_{\varepsilon(\delta)}\|^2 + \psi(\delta, \varepsilon(\delta)),$$

или

$$\langle(\lambda^\delta, \mu^\delta), (A^\delta z^{\delta,\varepsilon(\delta)}[\lambda^\delta, \mu^\delta] - h^\delta, g^\delta(z^{\delta,\varepsilon(\delta)}[\lambda^\delta, \mu^\delta]))\rangle \leqslant$$
$$\leqslant f^0(z^0_{\varepsilon(\delta)}) + \varepsilon(\delta)\|z^0_{\varepsilon(\delta)}\|^2 + \psi(\delta, \varepsilon(\delta)) - f^\delta(z^{\delta,\varepsilon(\delta)}[\lambda^\delta, \mu^\delta]),$$

что с учетом предельного соотношения (2.33), первой из оценок (1.1), оценки (2.25), условия согласования (2.21) и предельных соотношений $z^0_{\varepsilon(\delta)} \to z^0$, $\psi(\delta, \varepsilon(\delta)) \to 0$, $\delta \to 0$, дает оценку

$$\langle(\lambda^\delta, \mu^\delta), (A^\delta z^{\delta,\varepsilon(\delta)}[\lambda^\delta, \mu^\delta] - h^\delta, g^\delta(z^{\delta,\varepsilon(\delta)}[\lambda^\delta, \mu^\delta]))\rangle \leqslant \gamma(\delta) \to 0, \quad \delta \to 0.$$

При этом в силу (2.6) имеем $\langle(\lambda^\delta, \mu^\delta), (A^\delta z^{\delta,\varepsilon(\delta)}[\lambda^\delta, \mu^\delta] - h^\delta, g^\delta(z^{\delta,\varepsilon(\delta)}[\lambda^\delta, \mu^\delta]))\rangle \geqslant 0$. Две последние оценки дают предельное соотношение (2.32) в случае $\beta = \beta_0$.

Докажем следующий регуляризованный ПЛ, являющийся одновременно регуляризирующим алгоритмом для построения ОМП-образующего оператора в задаче (P^0) при условии $\beta = \beta_0$. Он объединяет в себе необходимые и достаточные условия существования ограниченной ОМП в этой задаче.

Теорема 3. *Пусть $\beta = \beta_0$, выполняется условие (2.3) и $\delta^k \in (0, \delta_0)$, $k = 1, 2, \ldots$, $\delta^k \to 0$, $k \to \infty$, — произвольная фиксированная последовательность. Тогда:*

1) *для существования в задаче (P^0) ограниченной ОМП (или, что эквивалентно, оптимального элемента) необходимо, чтобы существовали последовательность сходящихся к нулю положительных чисел ε^k, $k = 1, 2, \ldots$, и последовательность двойственных переменных $(\lambda^k, \mu^k) \in H \times \mathbb{R}^m_+$, $k = 1, 2, \ldots$, такие, что $\delta^k \|(\lambda^k, \mu^k)\| \to 0$, $k \to \infty$, и выполняются включения*

$$z^{\delta^k,\varepsilon^k}[\lambda^k, \mu^k] \in \mathcal{D}^{\delta^k,\gamma^k}, \quad \gamma^k \to 0, \quad k \to \infty, \tag{2.34}$$

и предельное соотношение

$$\langle(\lambda^k, \mu^k), (A^{\delta^k} z^{\delta^k,\varepsilon^k}[\lambda^k, \mu^k] - h^{\delta^k}, g^{\delta^k}(z^{\delta^k,\varepsilon^k}[\lambda^k, \mu^k]))\rangle \to 0, \quad k \to \infty, \tag{2.35}$$

а последовательность $z^{\delta^k,\varepsilon^k}[\lambda^k, \mu^k]$, $k = 1, 2, \ldots$, является (возможно неограниченной) ОМП. Другими словами, оператор, задаваемый равенством

$$R(f^{\delta^k}, A^{\delta^k}, h^{\delta^k}, g^{\delta^k}, \delta^k) = z^{\delta^k,\varepsilon^k}[\lambda^k, \mu^k], \tag{2.36}$$

является регуляризирующим в смысле определения 1, причем каждая слабая предельная точка последовательности $z^{\delta^k,\varepsilon^k}[\lambda^k, \mu^k]$, $k = 1, 2, \ldots$, в случае ее ограниченности, есть решение задачи (P^0). В качестве конкретной последовательности $(\lambda^k, \mu^k) \in H \times \mathbb{R}^m_+$, $k = 1$, $2, \ldots$, можно взять, например, последовательность $(\lambda^{\delta^k,\alpha(\delta^k),\varepsilon^k}, \mu^{\delta^k,\alpha(\delta^k),\varepsilon^k})$, $k = 1, 2, \ldots$, с $\varepsilon^k = \varepsilon(\delta^k)$ при условиях согласования (2.1) и (2.21), вырабатываемую ОМП-образующим алгоритмом теоремы 2;

2) *для существования в задаче (P^0) ограниченной ОМП достаточно, чтобы существовали последовательность сходящихся к нулю положительных чисел ε^k, $k = 1, 2, \ldots$, и последовательность двойственных переменных $(\lambda^k, \mu^k) \in H \times \mathbb{R}^m_+$, $k = 1, 2, \ldots$, такие, что $\delta^k \|(\lambda^k, \mu^k)\| \to 0$, $k \to \infty$, выполняются включения (2.34) и предельное соотношение (2.35), а последовательность $z^{\delta^k,\varepsilon^k}[\lambda^k, \mu^k]$, $k = 1, 2, \ldots$, является ограниченной. В этом случае задаваемый равенством (2.36) оператор является регуляризирующим в смысле определения 1 и каждая слабая предельная точка последовательности $z^{\delta^k,\varepsilon^k}[\lambda^k, \mu^k]$, $k = 1, 2, \ldots$, есть решение задачи (P^0).*

Замечание 6. Поясним содержательный смысл соотношений (2.34), (2.35) сформулированной теоремы 3. Последовательность включений (2.34) и предельное соотношение (2.35) естественно позиционировать как «секвенциальные обобщения» обычных для классического ПЛ (см., например, [1, п. 1.3.3]) условия допустимости оптимального элемента и условия дополняющей нежесткости соответственно. При этом на элементы последовательности $z^{\delta^k,\varepsilon^k}[\lambda^k, \mu^k]$, $k = 1, 2, \ldots$, следует смотреть как на удовлетворяющие ограничениям задачи (P^0) и обычному условию дополняющей нежесткости «лишь в пределе»[2].

Д о к а з а т е л ь с т в о. Для доказательства необходимости заметим, что в этом случае задача (P^0) разрешима, то есть $Z^0 \neq \emptyset$. Теперь включение (2.34) и предельное соотношение $\delta^k \|(\lambda^k, \mu^k)\| \to 0$, $k \to \infty$, а также предельное соотношение (2.35) теоремы вытекают из (2.30), (2.31) (см. теорему 2) и (2.32), если в качестве точек (λ^k, μ^k), $z^{\delta^k,\varepsilon^k}[\lambda^k, \mu^k]$ взять точки $(\lambda^{\delta^k,\alpha(\delta^k),\varepsilon(\delta^k)}, \mu^{\delta^k,\alpha(\delta^k),\varepsilon(\delta^k)})$, $z^{\delta^k,\varepsilon(\delta^k)}[\lambda^{\delta^k,\alpha(\delta^k),\varepsilon(\delta^k)}, \mu^{\delta^k,\alpha(\delta^k),\varepsilon(\delta^k)}]$, $k = 1, 2, \ldots$, соответственно, с $\varepsilon^k = \varepsilon(\delta^k)$ и при условиях согласования (2.1) и (2.21).

[2]Если же предположить, что последовательность $z^{\delta^k,\varepsilon^k}[\lambda^k, \mu^k]$, $k = 1, 2, \ldots$, сходится, естественно, к оптимальному элементу z^0, и $(\lambda^k, \mu^k) \to (\lambda^0, \mu^0)$, $k \to \infty$, то переход к пределу в (2.35) дает равенство $\langle \mu^0, g^0(z^0)\rangle = 0$ и, как следствие, обычное условие дополняющей нежесткости $\mu^0_i g^0_i(z^0) = 0$, $i = 1, \ldots, m$.

Для доказательства достаточности заметим, прежде всего, что множество Z^0 не пусто ввиду включения $z^{\delta^k,\varepsilon^k}[\lambda^k,\mu^k] \in \mathcal{D}^{\delta^k,\gamma^k}$, ограниченности последовательности $z^{\delta^k,\varepsilon^k}[\lambda^k,\mu^k]$, $k = 1,2,\ldots$, и условий на исходные данные задачи (P^0). Далее, можем записать

$$L^{\delta^k,\varepsilon^k}(z^{\delta^k,\varepsilon^k}[\lambda^k,\mu^k],\lambda^k,\mu^k) \leqslant L^{\delta^k,\varepsilon^k}(z,\lambda^k,\mu^k) \quad \forall\, z \in \mathcal{D}.$$

Отсюда, с учетом (2.35), ограниченности последовательности $z^{\delta^k,\varepsilon^k}[\lambda^k,\mu^k]$, $k = 1,2,\ldots$, следует

$$f^{\delta^k}(z^{\delta^k,\varepsilon^k}[\lambda^k,\mu^k]) \leqslant f^{\delta^k}(z) + \langle(\lambda^k,\mu^k),(A^{\delta^k}z - h^{\delta^k}, g^{\delta^k}(z))\rangle + \psi^k$$
$$\forall\, z \in \mathcal{D}, \quad \psi^k \to 0, \quad k \to \infty.$$

Положим здесь $z = z^0 \in Z^0$ и используем условие согласования $\delta^k\|(\lambda^k,\mu^k)\| \to 0$, $k \to \infty$, и опять ограниченность последовательности $z^{\delta^k,\varepsilon^k}[\lambda^k,\mu^k]$, $k = 1,2,\ldots$. Тогда получаем $f^0(z^{\delta^k,\varepsilon^k}[\lambda^k,\mu^k]) \leqslant f^0(z^0) + \widetilde{\psi}^k$, $\widetilde{\psi}^k \to 0$, $k \to \infty$. Так как одновременно мы имеем включение $z^{\delta^k,\varepsilon^k}[\lambda^k,\mu^k] \in \mathcal{D}^{\delta^k,\gamma^k}$, то используя классические свойства слабой компактности ограниченного выпуклого замкнутого множества и слабой полунепрерывности снизу непрерывного выпуклого функционала в гильбертовом пространстве, получаем, что $f^0(z^{\delta^k,\varepsilon^k}[\lambda^k,\mu^k]) \to f^0(z^0)$, $k \to \infty$, то есть последовательность $z^{\delta^k,\varepsilon^k}[\lambda^k,\mu^k]$, $k = 1,2,\ldots$, является ОМП в задаче (P^0) и, стало быть, задаваемый равенством (2.36) оператор является регуляризирующим в смысле определения 1, причем каждая слабая предельная точка последовательности $z^{\delta^k,\varepsilon^k}[\lambda^k,\mu^k]$, $k = 1,2,\ldots$, в силу ее ограниченности есть решение задачи (P^0). □

Финансирование. Работа выполнена при финансовой поддержке РФФИ (коды проектов 19-07-00782_а, 20-01-00199_а, 20-52-00030 Бел_а).

СПИСОК ЛИТЕРАТУРЫ

1. Алексеев В. М., Тихомиров В. М., Фомин С. В. Оптимальное управление. М.: Наука, 1979.
2. Васильев Ф. П. Методы оптимизации: В 2-х книгах. М.: МЦНМО, 2011.
3. Сумин М. И. Регуляризованная параметрическая теорема Куна–Таккера в гильбертовом пространстве // Журн. вычисл. матем. и матем. физ. 2011. Т. 51. № 9. С. 1594–1615.
4. Сумин М. И. Устойчивое секвенциальное выпуклое программирование в гильбертовом пространстве и его приложение к решению неустойчивых задач // Журн. вычисл. матем. и матем. физ. 2014. Т. 54. № 1. С. 25–49. https://doi.org/10.7868/S0044466914010141
5. Сумин М. И. Зачем нужна регуляризация принципа Лагранжа и принципа максимума Понтрягина и что она дает // Вестн. Тамбов. ун-та. Сер. Естественные и технические науки. 2018. Т. 23. Вып. 124. С. 757–775.
https://doi.org/10.20310/1810-0198-2018-23-124-757-775
6. Сумин М. И. Регуляризованные принцип Лагранжа и принцип максимума Понтрягина в оптимальном управлении и обратных задачах // Труды Института математики и механики УрО РАН. 2019. Т. 25. № 1. С. 279–296.
https://doi.org/10.21538/0134-4889-2019-25-1-279-296
7. Тихонов А. Н., Арсенин В. Я. Методы решения некорректных задач. М.: Наука, 1986.
8. Краснов М. Л., Киселев А. И., Макаренко Г. И. Интегральные уравнения. М.: Наука, 1976.
9. Сумин М. И. Регуляризация в линейно выпуклой задаче математического программирования на основе теории двойственности // Журн. вычисл. матем. и матем. физ. 2007. Т. 47. № 4. С. 602–625.
10. Гольштейн Е. Г. Теория двойственности в математическом программировании и ее приложения. М.: Наука, 1971.

11. Варга Дж. Оптимальное управление дифференциальными и функциональными уравнениями. М.: Наука, 1977.
12. Сумин М. И. О регуляризации классических условий оптимальности в выпуклых задачах оптимального управления // Труды Института математики и механики УрО РАН. 2020. Т. 26. № 2. С. 252–269. https://doi.org/10.21538/0134-4889-2020-26-2-252-269
13. Сумин М. И. Об устойчивом секвенциальном принципе Лагранжа в выпуклом программировании и его применении при решении неустойчивых задач // Труды Института математики и механики УрО РАН. 2013. Т. 19. № 4. С. 231-240.
14. Кутерин Ф. А., Сумин М. И. О регуляризованном принципе Лагранжа в итерационной форме и его применении для решения неустойчивых задач // Математическое моделирование. 2016. Т. 28. № 11. С. 3–18.
15. Кутерин Ф. А., Сумин М. И. Устойчивый итерационный принцип Лагранжа в выпуклом программировании как инструмент для решения неустойчивых задач // Журн. вычисл. матем. и матем. физ. 2017. Т. 57. № 1. С. 55–68. https://doi.org/10.7868/S0044466917010100

Поступила в редакцию 01.06.2020

Сумин Михаил Иосифович, д. ф.-м. н., профессор, Тамбовский государственный университет им. Г. Р. Державина, 392000, Россия, г. Тамбов, ул. Интернациональная, 33;
Нижегородский государственный университет им. Н. И. Лобаческого, 603950, Россия, г. Нижний Новгород, пр. Гагарина, 23.
E-mail: m.sumin@mail.ru

Цитирование: М. И. Сумин. О регуляризации принципа Лагранжа и построении обобщенных минимизирующих последовательностей в выпуклых задачах условной оптимизации // Вестник Удмуртского университета. Математика. Механика. Компьютерные науки. 2020. Т. 30. Вып. 3. С. 410–428.

MATHEMATICS 2020. Vol. 30. Issue 3. Pp. 410–428.

M. I. Sumin

On the regularization of the Lagrange principle and on the construction of the generalized minimizing sequences in convex constrained optimization problems

Keywords: constrained optimization, instability, dual regularization, regularized Lagrange principle, generalized minimizing sequence.

MSC2010: 90C25, 90C46, 47A52, 65F22

DOI: 10.35634/vm200305

We consider the regularization of the Lagrange principle (LP) in the convex constrained optimization problem with operator constraint-equality in a Hilbert space and with a finite number of functional inequality-constraints. The objective functional of the problem is not, generally speaking, strongly convex. The set of admissible elements of the problem is also embedded into a Hilbert space and is not assumed to be bounded. Obtaining a regularized LP is based on the dual regularization method and involves the use of two regularization parameters and two corresponding matching conditions at the same time. One of the regularization parameters is «responsible» for the regularization of the dual problem, while the other is contained in a strongly convex regularizing addition to the objective functional of the original problem. The main purpose of the regularized LP is the stable generation of generalized minimizing sequences that approximate the exact solution of the problem by function and by constraint, for the purpose of its practical stable solving.

Funding. The paper was funded by RFBR (projects 19-07-00782_a, 20-01-00199_a, 20-52-00030 Bel_a).

REFERENCES

1. Alekseev V. M., Tikhomirov V. M., Fomin S. V. *Optimal Control*, New York: Plenum Press, 1987. https://doi.org/10.1007/978-1-4615-7551-1
2. Vasil'ev F. P. *Metody optimizatsii* (Optimization methods), vols. 1, 2, Moscow: Moscow Center for Continuous Mathematical Education, 2011.
3. Sumin M. I. Regularized parametric Kuhn–Tucker theorem in a Hilbert space, *Comput. Math. Math. Phys.*, 2011, vol. 51, no. 9, pp. 1489–1509. https://doi.org/10.1134/S0965542511090156
4. Sumin M. I. Stable sequential convex programming in a Hilbert space and its application for solving unstable problems, *Comput. Math. Math. Phys.*, 2014, vol. 54, no. 1, pp. 22–44. https://doi.org/10.1134/S0965542514010138
5. Sumin M. I. Why regularization of Lagrange principle and Pontryagin maximum principle is needed and what it gives, *Vestnik Tambovskogo Universiteta. Seriya Estestvennye i Tekhnicheskie Nauki*, 2018, vol. 23, no. 124, pp. 757–775 (in Russian). https://doi.org/10.20310/1810-0198-2018-23-124-757-775
6. Sumin M. I. Regularized Lagrange principle and Pontryagin maximum principle in optimal control and in inverse problems, *Trudy Instituta Matematiki i Mekhaniki UrO RAN*, 2019, vol. 25, no. 1, pp. 279–296 (in Russian). https://doi.org/10.21538/0134-4889-2019-25-1-279-296
7. Tikhonov A. N., Arsenin V. Ya. *Solutions of Ill-Posed Problems*, New York: Halsted Press, 1977.
8. Krasnov M. L., Kiselev A. I., Makarenko G. I. *Integral'nye uravneniya* (Integral Equations), Moscow: Nauka, 1976.
9. Sumin M. I. Duality-based regularization in a linear convex mathematical programming problem, *Comput. Math. Math. Phys.*, 2007, vol. 47, no. 4, pp. 579–600. https://doi.org/10.1134/S0965542507040045
10. Gol'shtein E. G. *Teoriya dvoistvennosti v matematicheskom programmirovanii i ee prilozheniya* (Duality theory in mathematical programming and its applications), Moscow: Nauka, 1971.

11. Warga J. *Optimal control of differential and functional equations*, New York: Academic Press, 1972. Translated under the title *Optimal'noe upravlenie differentsial'nymi i funktsional'nymi uravneniyami*, Moscow: Nauka, 1977.
12. Sumin M. I. On the regularization of the classical optimality conditions in convex optimal control problems, *Trudy Instituta Matematiki i Mekhaniki UrO RAN*, 2020, vol. 26, no. 2, pp. 252–269 (in Russian). https://doi.org/10.21538/0134-4889-2020-26-2-252-269
13. Sumin M. I. On the stable sequential Lagrange principle in convex programming and its application for solving unstable problems, *Trudy Instituta Matematiki i Mekhaniki UrO RAN*, 2013, vol. 19, no. 4, pp. 231–240 (in Russian).
14. Kuterin F. A., Sumin M. I. On the regularized Lagrange principle in the iterative form and its application for solving unstable problems, *Mathematical Models and Computer Simulations*, 2017, vol. 9, no. 3, pp. 328–338. https://doi.org/10.1134/S2070048217030085
15. Kuterin F. A., Sumin M. I. The stable iterative Lagrange principle in convex programming as an instrument of solving unstable problems, *Comput. Math. Math. Phys.*, 2017, vol. 57, no. 1, pp. 71–82. https://doi.org/10.1134/S0965542517010092

Received 01.06.2020

Sumin Mikhail Iosifovich, Doctor of Physics and Mathematics, Professor, Tambov State University, ul. Internatsional'naya, 33, Tambov, 392000, Russia;
Nizhnii Novgorod State University, pr. Gagarina, 23, Nizhnii Novgorod, 603950, Russia.
E-mail: m.sumin@mail.ru

Citation: M. I. Sumin. On the regularization of the Lagrange principle and on the construction of the generalized minimizing sequences in convex constrained optimization problems, *Vestnik Udmurtskogo Universiteta. Matematika. Mekhanika. Komp'yuternye Nauki*, 2020, vol. 30, issue 3, pp. 410–428.