УДК 519.85

ОБ УСТОЙЧИВОМ КОНСТРУИРОВАНИИ МИНИМИЗИРУЮЩИХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ В ЗАДАЧАХ НЕЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ С ОГРАНИЧЕНИЯМИ ТИПА РАВЕНСТВА И НЕРАВЕНСТВА

А.В. Канатов

Нижегородский госуниверситет им. Н.И. Лобачевского

alexkanatov@yandex.ru

Поступила в редакцию 11.02.2013

Изложен метод двойственной регуляризации применительно к параметрической задаче нелинейного программирования общего вида в гильбертовом пространстве с операторным ограничением типа равенства и конечным числом функциональных ограничений типа неравенства. Данный метод обеспечивает устойчивое к ошибкам исходных данных конструирование элементов минимизирующей последовательности в исходной задаче из элементов последовательностей, являющихся минимизирующими для модифицированной функции Лагранжа, взятой при значениях двойственных переменных из соответствующей максимизирующей последовательности в модифицированной двойственной задаче. В частности, показывается, как свойства обобщенной дифференцируемости полунепрерывных снизу функций значений в бесконечномерных задачах математического программирования порождают соответствующие конструкции модифицированных функций Лагранжа. Приводится пример, иллюстрирующий неустойчивость формального построения минимизирующей последовательности без регуляризации решения модифицированной двойственной задачи.

Ключевые слова: нелинейное программирование, параметрическая задача, секвенциальная оптимизация, проксимальный субградиент, принцип Лагранжа, вектор Куна–Таккера, минимизирующее приближенное решение, двойственность, регуляризация.

Введение

Настоящая статья посвящена распространению метода двойственной регуляризации [1–4] на нелинейные бесконечномерные параметрические задачи математического программирования. В ней непосредственно продолжается изучение на основе указанного метода нелинейных задач условной оптимизации, начатое в [5, 6]; а точнее говоря, метод параметрической двойственной регуляризации, развитый в [6] для нелинейной задачи математического программирования с ограничениями типа равенства, распространяется на случай аналогичной задачи, содержащей помимо ограничений-равенств и ограничения-неравенства.

Метод параметрической двойственной регуляризации основан на идеологии секвенциальной оптимизации, т.е. теории оптимизации, в которой базовым является понятие минимизирующей последовательности, а не понятие оптимального элемента. Отличительной чертой указанного метода является применение трех классических подходов математической теории, каковыми являются метод возмущений, теория двойственности в условной оптимизации (см., например, [7, с. 263; 8–10]) и метод стабилизации (регуляризации) Тихонова (см., например, [11, 12]). Отметим две важнейшие особенности развиваемого метода параметрической двойственной регуляризации. Во-первых, объединение трех классических подходов на основе идеологии секвенциальной оптимизации позволяет параллельно с устойчивым к ошибкам исходных данных конструированием минимизирующих последовательностей формулировать и соответствующие «устойчивые» необходимые и достаточные условия на элементы минимизирующих последовательностей. В более простой ситуации задачи выпуклого программирования с ограничениями типа равенства и неравенства соответствующие результаты и их поясняющие примеры можно найти в [13]. И, во-вторых, указанное объединение трех классических подходов на базе секвенциальной оптимизации позволяет теснейшим образом увязать все основные конструкции развиваемого метода с фундаментальными в современном бесконечномерном негладком анализе понятиями нормалей к замкнутым множествам и соответствующими понятиями обобщенного дифференцирования полунепрерывных снизу функций.

Поясним сказанное выше на примере рассмотренной в [6] и являющейся частным случаем изучаемой в данной работе нелинейной параметрической задачи математического программирования общего вида с ограничением типа равенства

$$(P_p)\quad f(z)\to \inf, g(z)=p, z\in D\subset Z, \text{ где}$$

$f:D\to R^1$ – непрерывный функционал, $g:D\to H$ – вполне непрерывный оператор, $D\subset Z$ – ограниченное замкнутое множество, Z, H – гильбертовы пространства, $p\in H$ – параметр.

Напомним определение обобщенной нижней грани в задаче (P_p):

$$\beta(p)\equiv\beta_{+0}(p)\equiv\lim_{\varepsilon\to+0}\beta_\varepsilon(p),$$

$$\beta_\varepsilon(p)=\begin{cases}\inf\limits_{z\in D_p^\varepsilon} f(z), & D_p^\varepsilon\neq\varnothing,\\ +\infty, & D_p^\varepsilon=\varnothing;\end{cases}$$

где $D_p^\varepsilon\equiv\{z\in D:\|g(z)-p\|\le\varepsilon\}, \varepsilon\ge 0$.

Центральной в [6] является конструкция так называемой модифицированной функции Лагранжа (см., например, [14]). Как показано в [6], конструкция модифицированной функции Лагранжа

$$L_{p,c}(z,\lambda)\equiv f(z)+\langle\lambda,g(z)-p\rangle+$$
$$c(l_1\|g(z)-p\|+l_2\|g(z)-p\|^2),\ z\in D,$$

с весовыми множителями $l_1,l_2\in\{0,1\}$ задачи (P_p) полностью определяется дифференциальными свойствами ее функции значений (S-функции) $\beta(p)$, $p\in H$, которая в самой общей ситуации является полунепрерывной снизу. Как известно, характеристическим свойством полунепрерывных снизу функций является их субдифференцируемость в том или ином обобщенном смысле на плотном в $\operatorname{dom}\beta$ множестве (см., например, [15–17]). При этом указанная субдифференцируемость может пониматься поразному. В [6] использовались два широко применяемых в современном негладком анализе понятия субдифференцируемости, каковыми являются субдифференцируемость в смысле существования проксимального субградиента (см., например, [15–17]) и в смысле существования субдифференциала Фреше ([18, 19]). В случае существования проксимального субградиента функции $\beta(p)$ в точке p в задаче (P_p) существует обобщенный вектор Куна–Таккера $\lambda\in H$ в том смысле, что при некотором $c>0$ выполняется неравенство

$$\beta(p)\le f(z)+\langle\lambda,g(z)-p\rangle+c\|g(z)-p\|^2\ \forall z\in D.$$

В случае же существования субдифференциала Фреше функции $\beta(p)$ в точке p в задаче (P_p) обобщенный вектор Куна–Таккера $\lambda\in H$ существует в смысле неравенства

$$\beta(p)\le f(z)+\langle\lambda,g(z)-p\rangle+c\|g(z)-p\|\quad\forall z\in D,$$

при некотором $c>0$.

И в том и в другом случае существование вектора Куна–Таккера в указанных смыслах означает и существование седловых точек у соответствующих модифицированных функций Лагранжа: в первом случае это функция

$$f(z)+\langle\lambda,g(z)-p\rangle+c\|g(z)-p\|^2,$$

во втором

$$f(z)+\langle\lambda,g(z)-p\rangle+c\|g(z)-p\|.$$

В обоих вариантах понятие обобщенного вектора Куна–Таккера является более общим, чем понятие обобщенного вектора Куна–Таккера в общепринятом классическом смысле ($\beta(p)\le$ $\le L_{p,0}(z,\lambda)\ \forall z\in D$), использованного в работе [5]. В каждом из этих двух случаев метод двойственной регуляризации [6] приводит к конструированию минимизирующей последовательности в задаче (P_p). Таким образом, именно информация об обобщенной дифференцируемости функции значений естественным образом порождает конструкцию модифицированной функции Лагранжа $L_{p,c}(z,\lambda)$ $z\in D$, а центральным при устойчивом построении минимизирующей последовательности, если на это смотреть как на основу для алгоритма, является условие, в соответствии с которым минимизация модифицированной функции Лагранжа может проводиться с любой наперед заданной точностью. Если же обобщенный вектор Куна–Таккера в задаче отсутствует, то штрафной коэффициент c при конструировании минимизирующей последовательности в задаче (P_p) необходимо стремить к бесконечности (в данной работе из-за ограничений на ее объем этот случай не рассматривается). Подчеркнем одновременно, что построение элементов минимизирующей последовательности в задаче (P_p) как в первом, так и во втором случае происходит с помощью элементов последовательностей, являющихся минимизирующими для модифицированной функции Лагранжа, взятой в точках λ из соответствующей максимизирующей последовательности двойственных переменных в модифицированной двойственной задаче. В данной работе мы ограничимся рассмотрением только первого из указанных выше двух случаев.

1. Постановка задачи нелинейного программирования

Рассмотрим параметрическую задачу минимизации

$(P_{p,r})\quad f(z)\to\inf,\ g(z)=p,\ h(z)\le r,\ z\in D\subset Z,$

где $f:D\to R^1$ – непрерывный функционал, $g:D\to H$ – вполне непрерывный оператор, $h\equiv(h_1,...,h_m):D\to R^m$ – непрерывный векторный функционал, $D\subset Z$ – замкнутое ограниченное множество, Z,H – гильбертовы пространства, $p\in H,\ r\equiv(r_1,...,r_m)\in R^m$ – параметры. Будем также считать, что

$$\begin{aligned}&|f(z_1)-f(z_2)|\le L\|z_1-z_2\|,\\&\|g(z_1)-g(z_2)\|\le L\|z_1-z_2\|,\\&|h(z_1)-h(z_2)|\le L\|z_1-z_2\|\ \ \forall z_1,z_2\in D,\end{aligned}\tag{1.1}$$

где $L>0$ – некоторая постоянная. Обозначим

$$D^{\varepsilon}_{p,r}\equiv\{z\in D:\|g(z)-p\|\le\varepsilon,\min_{x\in R^m_-}|h(z)-r-x|\le\varepsilon\},\ \ \varepsilon\ge 0,$$

$$R^m_-\equiv\{x\in R^m:x\le 0\},\ R^m_+\equiv\{x\in R^m:x\ge 0\}.$$

Определим функцию значений (S-функцию) $\beta:H\times R^m\to R^1\cup\{+\infty\}$ задачи $(P_{p,r})$:

$$\beta(p,r)\equiv\beta_{+0}(p,r)\equiv\lim_{\varepsilon\to+0}\beta_{\varepsilon}(p,r),$$

$$\beta_{\varepsilon}(p,r)=\begin{cases}\inf\limits_{z\in D^{\varepsilon}_{p,r}} f(z), & D^{\varepsilon}_{p,r}\ne\varnothing;\\ +\infty, & D^{\varepsilon}_{p,r}=\varnothing.\end{cases}$$

Очевидно, что в самой общей ситуации $\beta(p,r)\le\beta_0(p,r)$, где $\beta_0(p,r)$ – классическое значение задачи $(P_{p,r})$.

Лемма 1.1. *Функция значений* $\beta:H\times R^m\to$ $\to R^1\cup\{+\infty\}$ *является полунепрерывной снизу*.

Доказательство этой леммы проводится точно так же, как и доказательство полностью аналогичного утверждения в [20, лемма 7].

Определим минимизирующую последовательность – минимизирующее приближенное решение в смысле Дж. Варги [21] в задаче $(P_{p,r})$, т.е. последовательность элементов $z^i\in D,\ i=1,2,...,$ такую, что $f(z^i)\le\beta(p,r)+\delta^i,\ z^i\in D^{\varepsilon^i}_{p,r}$ для некоторых последовательностей сходящихся к нулю неотрицательных чисел $\delta^i,\varepsilon^i, i=1,2,...$.

Пусть F – множество всевозможных наборов исходных данных $\mathrm{f}\equiv\{f,g,h\}$, каждый из которых состоит из непрерывного на множестве D функционала f , вполне непрерывного оператора g и непрерывного векторного функционала h с указанными выше свойствами (1.1). Определим наборы невозмущенных $\mathrm{f}^0\equiv\{f^0,g^0,h^0\}$ и возмущенных $\mathrm{f}^{\delta}\equiv\{f^{\delta},g^{\delta},h^{\delta}\}$ исходных данных, где $\delta\in(0,\delta_0]$, $\delta_0>0$ – некоторое число. Будем считать, что выполняются оценки

$$\begin{aligned}&|f^{\delta}(z)-f^0(z)|,\ \|g^{\delta}(z)-g^0(z)\|,\\&|h^{\delta}(z)-h^0(z)|\le K\delta\quad\forall z\in D,\end{aligned}\tag{1.2}$$

где $K>0$ – некоторая не зависящая от δ постоянная. Обозначим задачу $(P_{p,r})$, функционал f , оператор g , векторный функционал h , функцию значений β и т.п., соответствующие набору исходных данных f^{δ}, $\delta\in(0,\delta_0]$, через $(P^{\delta}_{p,r})$, f^{δ}, g^{δ}, h^{δ}, β^{δ}.

Пусть $\beta^0(p,r)<+\infty$. Нас будет интересовать процесс устойчивого конструирования в нелинейной задаче математического программирования

$$(P^0_{p,r})\quad f^0(z)\to\inf, g^0(z)=p, h^0(z)\le r,\ z\in D$$

минимизирующего приближенного решения, связанный с регуляризованным алгоритмом решения модифицированной двойственной к $(P^0_{p,r})$ задачи или, другими словами, с регуляризованным алгоритмом максимизации вогнутого недифференцируемого целевого функционала модифицированной двойственной задачи.

С учетом приближенного задания исходных данных вместо задачи $(P^0_{p,r})$ имеем семейство зависящих от ошибки δ задач

$$(P^{\delta}_{p,r})\ f^{\delta}(z)\to\inf, g^{\delta}(z)=p, h^{\delta}(z)\le r, z\in D.$$

2. Эквивалентная задача с операторным ограничением типа равенства и ее модифицированная функция Лагранжа

Следуя хорошо известному приему (см., например, [14, с. 165]), приведем задачу математического программирования с ограничениями типа равенства и неравенства $(P_{p,r})$ к виду эквивалентной задачи, содержащей лишь ограничения типа равенства. Для этого заметим, что задача $(P_{p,r})$ эквивалентна задаче

$$(\tilde P_{p,r})\quad f(z)\to\inf,\ g(z)=p,\ h(z)+y=r,$$
$$z\in D\subset Z,\ y\in R^m_+$$

в том смысле, что последовательность $z^i\in D, i=1,2,...,$ является минимизирующим приближенным решением в задаче $(P_{p,r})$ тогда и только тогда, когда последовательность $(z^i,-(h(z^i)-r)),\ i=1,2,...,$ является таковой в задаче $(\tilde P_{p,r})$. При этом функции значений задач $(P_{p,r})$ и $(\tilde P_{p,r})$ совпадают.

Определим формально модифицированную функцию Лагранжа в смысле [14, с. 165] для задачи $(\tilde P_{p,r})$

$$\tilde{L}^c_{p,r}(z,y,\lambda,\mu) \equiv f(z) + \langle \lambda, g(z)-p\rangle + \langle \mu, h(z)+y-r\rangle +$$
$$+\frac{c}{2}\|g(z)-p\|^2 + \frac{c}{2}|h(z)+y-r|^2,\ z\in D,\ y\in R^m_+,\ c\ge 0.$$

Определим далее формально и соответствующую модифицированную двойственную задачу для задачи $(\tilde{P}_{p,r})$

$$\tilde{V}^c_{p,r}(\lambda,\mu) \equiv \inf_{(z,y)\in D\times R^m_+} \tilde{L}^c_{p,r}(z,y,\lambda,\mu) \to \sup,$$
$$(\lambda,\mu)\in H\times R^m_+.$$

Функция $\tilde{V}^c_{p,r}(\cdot,\cdot)$ в силу равенства

$$\inf_{(z,y)\in D\times R^m_+} \tilde{L}^c_{p,r}(z,y,\lambda,\mu) = \inf_{z\in D}\min_{y\in R^m_+} \tilde{L}^c_{p,r}(z,y,\lambda,\mu)$$

приобретает вид

$$\tilde{V}^c_{p,r}(\lambda,\mu) = \inf_{z\in D}\min_{y\in R^m_+} \tilde{L}^c_{p,r}(z,y,\lambda,\mu) = \inf_{z\in D} L^c_{p,r}(z,\lambda,\mu),$$

где принято обозначение

$$L^c_{p,r}(z,\lambda,\mu) \equiv f(z) + \langle \lambda, g(z)-p\rangle + \frac{c}{2}\|g(z)-p\|^2 +$$
$$+\sum_{i=1}^m \min_{y_i\in R^1_+}\left(\left(\mu_i, h_i(z)+y_i-r_i\right) + \frac{c}{2}\left(h_i(z)+y_i-r_i\right)^2\right).$$

Определенная таким образом функция $L^c_{p,r}$ совпадает по своей конструкции с известной модифицированной функцией Лагранжа для исходной задачи с ограничениями типа равенства и неравенства $(P_{p,r})$ (подробности в [14, с. 167]). При этом справедливо равенство

$$L^c_{p,r}(z,\lambda,\mu) = f(z) + \langle \lambda, g(z)-p\rangle + \frac{c}{2}\|g(z)-p\|^2 +$$
$$+\frac{1}{2c}\sum_{i=1}^m \{[\max\{0, \mu_i + c(h_i(z)-r_i)\}]^2 - \mu_i^2\},$$
$$z\in D, (\lambda,\mu)\in H\times R^m, c>0.$$

Одновременно можно утверждать, что

$$L^0_{p,r}(z,\lambda,\mu) = f(z) + \langle \lambda, g(z)-p\rangle +$$
$$+\langle \mu, h(z)+y-r\rangle,\ \forall z\in D,\ \mu\ge 0.$$

3. Проксимальный субградиент полунепрерывной снизу функции значений и модифицированная функция Лагранжа

В данном разделе получим естественным путем формально введенную в разделе 2 конструкцию модифицированной функции Лагранжа на основе свойств обобщенной дифференцируемости полунепрерывной снизу функции значений β исходной задачи $(P_{p,r})$, после чего снова определим модифицированную двойственную задачу и вычислим супердифференциал модифицированной двойственной функции.

3.1. *Проксимальный субградиент полунепрерывной снизу функции в гильбертовом пространстве*

Для получения конструкции модифицированной функции Лагранжа в силу свойств обобщенной дифференцируемости функции значений нам понадобится важнейшее понятие проксимальной нормали к замкнутому множеству и соответствующее понятие проксимального субградиента полунепрерывной снизу функции (см., например, [15–17]).

Определение 3.1. (*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*

$$\langle \zeta, s-\bar{s}\rangle \le M\|s-\bar{s}\|^2\ \forall s\in S. \qquad (3.1)$$

Множество всех таких векторов ζ, 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 $\bar{x}\in \operatorname{dom} f$. 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 ζ обозначим через $\partial^P f(\bar{x})$ и назовем проксимальным субградиентом f в точке $\bar{x}$.

Лемма 3.1. *Пусть H – гильбертово пространство, $f: H\to R^1\cup\{+\infty\}$ – полунепрерывная снизу функция и $\bar{x}\in \operatorname{dom} f$. Вектор $\zeta\in H$ является проксимальным субградиентом функции f в точке $\bar{x}$, т.е. $\zeta\in\partial^P f(\bar{x})$, тогда и только тогда, когда существуют постоянные $R>0$ и $\delta>0$, такие, что*

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

3.2. *Конструкция модифицированной функции Лагранжа как следствие ее обобщенной дифференцируемости*

Вернемся далее к рассмотрению полунепрерывной снизу функции значений $\beta: H\times R^m\to$ $\to R^1\cup\{+\infty\}$ в задачах $(P_{p,r})$ и $(\tilde{P}_{p,r})$. Введем обозначение:

$$S_\delta(p,r) \equiv \{(p',r')\in H\times R^m: \|(p',r')-(p,r)\|<\delta\}.$$

Если точка $(p,r)\in\operatorname{dom}\beta$ такова, что $\partial^P\beta(p,r)\ne\varnothing$ и $\zeta\in\partial^P\beta(p,r)$, то из леммы 3.1 следует, что существуют постоянные $R>0$ и $\delta>0$ (зависящие от точки (p,r) и элемента $\zeta\equiv(\zeta_p,\zeta_r)$, $\zeta_p\in H$, $\zeta_r\in R^m$), такие, что

$$\beta(p,r)-\langle\zeta,(p,r)\rangle\le\beta(p',r')-\langle\zeta,(p',r')\rangle+ + R\|p'-p\|^2+R|r-r'|^2 \quad \forall(p',r')\in S_\delta(p,r). \quad (3.2)$$

Поскольку функция β является функцией значений задачи с ограничениями типа равенства и неравенства $(P_{p,r})$, то можно показать, что компонента $\zeta_r\equiv(\zeta_{r,1},\ldots,\zeta_{r,m})$ элемента ζ является вектором из R^m с неположительными компонентами: $\zeta_{r,i}\le 0, i=1,2,\ldots,m.$

Так как в силу ограниченности множества D функция β ограничена на множестве $\operatorname{dom}\beta$, то в силу неравенства (3.2) можем записать для некоторой постоянной $c=c(p,r,\zeta)>0$

$$\beta(p,r)-\langle\zeta,(p,r)\rangle\le\beta(p',r')-\langle\zeta,(p',r')\rangle+ + c\|p'-p\|^2+c|r'-r|^2 \quad \forall(p',r')\in H\times R^m, \quad (3.3)$$

откуда для $\hat c>c$ в силу полунепрерывности снизу функции значений β и ее ограниченности следует, что минимизирующей последовательностью в задаче минимизации

$$\beta(p',r')-\langle\zeta,(p',r')\rangle+\hat c\|p'-p\|^2+ + \hat c|r'-r|^2\to\inf,\ (p',r')\in H\times R^m \quad (3.4)$$

является лишь любая последовательность $(p^k,r^k),\ k=1,2,\ldots,$ сходящаяся к точке (p,r), такая, что $\beta(p^k,r^k)\to\beta(p,r),\ k\to\infty$, и никакая другая последовательность. Отсюда следует, что в задаче минимизации модифицированной функции Лагранжа (МФЛ), конструкция которой возникает здесь естественным образом, как следствие дифференциальных свойств функции значений β в точке (p,r),

$$f(z)+\langle-\zeta_p,g(z)-p\rangle+\langle-\zeta_r,h(z)+y-r\rangle+\hat c\|g(z)-p\|^2+ + \hat c|h(z)+y-r|^2\to\inf,\ z\in D,\ y\in R_+^m \quad (3.5)$$

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

$$\inf_{(z,y)\in D\times R_+^m}(f(z)+\langle-\zeta_p,g(z)-p\rangle+\langle-\zeta_r,h(z)+y-r\rangle+ + \hat c\|g(z)-p\|^2+\hat c|h(z)+y-r|^2)=\beta(p,r). \quad (3.6)$$

Одновременно, если для некоторых $\zeta=(\zeta_p,\zeta_r)\in H\times R^m,\ \hat c>0$ выполняется равенство (3.6), то выполняется и неравенство (3.3) при $c=\hat c$, а следовательно, и неравенство (3.2) при $R=\hat c$ и любом $\delta>0$, т.е. $\zeta=(\zeta_p,\zeta_r)\in \in\partial^P\beta(p,r)$.

Из последних рассуждений можно вывести следующее важное для нас

Следствие 3.1. *Если, например,* $(p,r)\in\operatorname{dom}\beta$ *такая точка, что* $\partial^P\beta(p,r)\ne\varnothing$*, то при любом фиксированном* $\zeta=(\zeta_p,\zeta_r)\in\partial^P\beta(p,r)$ *определенный выше в неравенстве* (3.3) *штрафной коэффициент* $c=c(p,r,\zeta)$ *можно взять столь большим, что в задаче минимизации МФЛ* (3.5) *минимизирующей является лишь последовательность* $(z^k,y^k),\ k=1,2,\ldots,$ *такая, что* $f(z^k)\to\beta(p,r),\ g(z^k)\to p,\ h(z^k)+y^k\to r,\ k\to\infty,$ *и никакая другая последовательность.*

Определим далее и модифицированную двойственную задачу, которая, естественно, совпадает с введенной ранее в разделе 2, по аналогии с хорошо известными публикациями по методу множителей Лагранжа в задачах условной минимизации (см., например, [14]), одноименной задачей,

$$\tilde V_{p,r}^{c,\delta}(\lambda,\mu)\to\sup,\ \lambda\in H,\ \mu\in R_+^m,$$

$$\tilde V_{p,r}^{c,\delta}(\lambda,\mu)\equiv\inf_{(z,y)\in D\times R_+^m}\tilde L_{p,r}^{c,\delta}(z,y,\lambda,\mu).$$

При этом, как показано в разделе 2,

$$\tilde V_{p,r}^{c,\delta}(\lambda,\mu)=\inf_{(z,y)\in D\times R_+^m}\tilde L_{p,r}^{c,\delta}(z,y,\lambda,\mu)\equiv\inf_{z\in D}L_{p,r}^{c,\delta}(z,\lambda,\mu),$$

где

$$L_{p,r}^{c,\delta}(z,\lambda,\mu)\equiv f^\delta(z)+\langle\lambda,g(z)-p\rangle+\frac{c}{2}\|g^\delta(z)-p\|^2+ + \sum_{i=1}^m\min_{y_y\in R_+^1}\left((\mu_i,h_i^\delta(z)+y_i-r_i)+\frac{c}{2}(h_i^\delta(z)+y_i-r_i)^2\right).$$

Напомним, что, как известно [14, с.167] и как уже отмечено выше, справедливо равенство

$$L_{p,r}^{c,\delta}(z,\lambda,\mu)=f^\delta(z)+\langle\lambda,g(z)-p\rangle+\frac{c}{2}\|g^\delta(z)-p\|^2+ + \frac{1}{2c}\sum_{i=1}^m\{[\max\ \{0,\mu_i+c(h_i^\delta(z)-r_i)\}]^2-\mu_i^2\}, \quad z\in D,\ (\lambda,\mu)\in H\times R^m,\ c>0. \quad (3.7)$$

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

$$\left|\tilde V_{p,r}^{c,\delta}(\lambda,\mu)-\tilde V_{p,r}^{c,0}(\lambda,\mu)\right|\le C\delta\left(1+\|\lambda\|+|\mu|+|c|\right) \quad \forall\lambda\in H,\mu\in R^m, \quad (3.8)$$

где $C>0$ – некоторая постоянная, зависящая при фиксированных $(p,r)\in H\times R^m$ лишь от $\sup_{z\in D}\|z\|$.

Далее важное значение для нас будет иметь следующая лемма, в которой устанавливается выражение для супердифференциала $\partial\tilde V_{p,r}^{c,\delta}$ вогнутой функции $\tilde{V}^{c,\delta}_{p,r}$ при условии полной непрерывности оператора $g^\delta : D \to H$. Здесь под супердифференциалом вогнутой функции $\tilde{V}^{c,\delta}_{p,r}$ понимается субдифференциал (в смысле выпуклого анализа) с обратным знаком выпуклой функции $-\tilde{V}^{c,\delta}_{p,r}$.

Лемма 3.2. *Супердифференциал* (*в смысле выпуклого анализа*) $\partial\tilde{V}^{c,\delta}_{p,r}(\lambda,\mu)$ *вогнутой функции* $\tilde{V}^{c,\delta}_{p,r}(\lambda,\mu)$ *в точке* $(\lambda,\mu)\in H\times R^m$ *выражается формулой*

$$\begin{aligned}
&\partial\tilde{V}^{c,\delta}_{p,r}(\lambda,\mu)=\partial_C\tilde{V}^{c,\delta}_{p,r}(\lambda,\mu)=\\
&=\overline{\operatorname{conv}}\{\lim_{i\to\infty}(g(z^i)-p,\\
&\max\{-\mu_j/c,(h_j(z^i)-r_j)\},\\
&j=1,\dots,m):z^i\in D,\\
&L^c_{p,r}(z^i,\lambda,\mu)\to\inf_{z\in D}L^c_{p,r}(z,\lambda,\mu),i\to\infty\}\equiv\\
&\equiv\overline{\operatorname{conv}}Q_{p,r,c}(\lambda,\mu),
\end{aligned}\tag{3.9}$$

где $\partial_C\tilde{V}^{c,\delta}_{p,r}(\lambda,\mu)$ *– обобщенный градиент Кларка функции* $\tilde{V}^{c,\delta}_{p,r}(\lambda,\mu)$ *в точке* (λ,μ).

Доказательство этой леммы проводится точно так же, как и доказательство полностью аналогичного утверждения в [3].

4. Двойственная регуляризация для решения нелинейной задачи математического программирования

Укажем, прежде всего, на важные обстоятельства, вытекающие из результатов предыдущего раздела. Возможны две и только две ситуации для исходной задачи $(P^0_{p,r})$:

А) в задаче существует вектор Куна–Таккера в следующем обобщенном смысле: существует вектор $(\lambda,\mu)\in H\times R^m_+$, для которого $\beta^0(p,r)\le\inf_{z\in D}L^{c,0}_{p,r}(z,\lambda,\mu)$ для некоторого $c>0$;

В) в задаче не существует вектора Куна–Таккера в указанном смысле.

Существование вектора Куна–Таккера в указанном обобщенном смысле эквивалентно тому, что целевая функция $\tilde{V}^{c,0}_{p,r}(\lambda,\mu)$ модифицированной двойственной задачи

$$\tilde{V}^{c,0}_{p,r}(\lambda,\mu)\to\sup,\quad(\lambda,\mu)\in H\times R^m_+$$

достигает значения $\beta^0(p,r)$ в некоторой точке $(\lambda^0,\mu^0)\in H\times R^m_+$.

Лемма 3.2 дает возможность организовать поиск максимума в задаче максимизации при каждом $c>0$ сильно вогнутого функционала

$$R^{c,\delta,\alpha}_{p,r}(\lambda,\mu)\equiv\tilde{V}^{c,\delta}_{p,r}(\lambda,\mu)-\alpha\|\lambda\|^2-\alpha|\mu|^2,$$
$$(\lambda,\mu)\in H\times R^m_+,\ \delta\ge0.$$

При этом с целью конструирования минимизирующего приближенного решения в исходной задаче $(P^0_{p,r})$ будем рассматривать при $\tilde{c}>c$ задачу

$$R^{\tilde{c},\delta,\alpha}_{p,r}(\lambda,\mu)\to\max,(\lambda,\mu)\in H\times R^m_+.\tag{4.1}$$

Обозначим через $(\lambda^{c,\delta,\alpha}_{p,r},\mu^{c,\delta,\alpha}_{p,r})$ единственную точку, дающую на $H\times R^m_+$ максимум функционалу $R^{c,\delta,\alpha}_{p,r}$. Покажем, что при условии согласования

$$\frac{\delta}{\alpha(\delta)}\to0,\ \delta\to0,\ \alpha(\delta)\to0\tag{4.2}$$

регуляризованный процесс поиска максимума в модифицированной двойственной задаче (4.1) конструктивно порождает минимизирующую последовательность $z^i\in D,\ i=1,2,\dots,$ в задаче $(P^0_{p,r})$, т.е.

$$f^0(z^i)\to\beta^0(p,r),\ z^i\in D^{0,\varepsilon^i}_{p,r},\ \varepsilon^i\to0,\ i\to\infty.$$

Итак, пусть задача $(P^0_{p,r})$ обладает вектором Куна–Таккера в указанном обобщенном смысле. Рассматриваем задачу (4.1) при произвольном достаточно большом фиксированном $\tilde{c}>c$. Заметим, прежде всего, что в соответствии со сказанным выше любой такой вектор доставляет максимальное значение функции $V^{c,0}_{p,r}$ на $H\times R^m_+$, равное $\beta^0(p,r)$, и, наоборот, любой вектор, доставляющий максимум функции $V^{c,0}_{p,r}$, равный $\beta^0(p,r)$, является вектором Куна–Таккера задачи $(P^0_{p,r})$. Замкнутое множество всех таких точек максимума (λ,μ) обозначим $K_{p,r,c}$.

Для всех $(\lambda,\mu)\in K_{p,r,c}$ в силу конструкции МФЛ $L^{c,0}_{p,r}(z,\lambda,\mu)$ и выражения (3.9) для супердифференциала функции $\tilde{V}^{c,0}_{p,r}(\lambda,\mu)$ справедливы соотношения

$$\{0\}=Q^0_{p,r,\tilde{c}}(\lambda,\mu)=\partial\tilde{V}^{\tilde{c},0}_{p,r}(\lambda,\mu),\ \tilde{c}>c,\tag{4.3}$$

для всех $(\lambda,\mu)\in K_{p,r,\tilde{c}}$ справедливо включение $0\in Q^0_{p,r,\tilde{c}}(\lambda,\mu)$, причем, вообще говоря, $K_{p,r,c}\subset K_{p,r,\tilde{c}}$. Таким образом, точки множества $K_{p,r,\tilde{c}}$ характеризуются тем, что для любой точки (λ,μ) из этого множества справедливо равенство $\tilde{V}^{\tilde{c},0}_{p,r}(\lambda,\mu)=\beta^0(p,r)$ и одновременно в силу равенства (3.9) найдется такая последовательность $z^i \in D,\ i=1,2,\ldots,$ для которой справедливы предельные соотношения

$$L_{p,r}^{\tilde{c},0}(z^i,\lambda,\mu)\to\beta^0(p,r)=\inf_{z\in D}L_{p,r}^{\tilde{c},0}(z,\lambda,\mu),$$
$$f^0(z^i)\to\beta^0(p,r),\ g^0(z^i)-p\to 0, \tag{4.4}$$
$$\max\{0,h_j^0(z^i)-r_j\}\to 0,\ j=1,2,\ldots,m,\ i\to\infty.$$

В частности, такой последовательностью при $(\lambda,\mu)\in K_{p,r,\tilde{c}}$ будет являться любая последовательность $z^i\in D,\ i=1,2,\ldots,$ для которой справедливы предельные соотношения (4.4) при $(\lambda,\mu)\in K_{p,r,c}$.

Пусть далее $\tilde{c}$ столь велико, что $K_{p,r,c}\cap H\times R_+^m\neq\varnothing$. В силу оценки (3.8), условия согласования (4.2) и теоремы о сходимости метода стабилизации в [12]

$$\left\|(\lambda_{p,r,\tilde{c}}^{\delta,\alpha(\delta)},\mu_{p,r,\tilde{c}}^{\delta,\alpha(\delta)})-(\lambda_{p,r,\tilde{c}}^{0},\mu_{p,r,\tilde{c}}^{0})\right\|\to 0,\ \delta\to 0, \tag{4.5}$$

где $(\lambda_{p,r,\tilde{c}}^{0},\mu_{p,r,\tilde{c}}^{0})$ – минимальная по норме точка в множестве $K_{p,r,\tilde{c}}$.

При этом в соответствии с леммой 3.2 можем записать

$$\begin{aligned}\partial\tilde{V}_{p,r}^{\tilde{c},\delta}&(\lambda_{p,r,\tilde{c}}^{\delta,\alpha(\delta)},\mu_{p,r,\tilde{c}}^{\delta,\alpha(\delta)})=\partial_C\tilde{V}_{p,r}^{\tilde{c},\delta}(\lambda_{p,r,\tilde{c}}^{\delta,\alpha(\delta)},\mu_{p,r,\tilde{c}}^{\delta,\alpha(\delta)})=\\&=\overline{\mathrm{conv}}\{\lim_{i\to\infty}(g^\delta(z^i)-p,\\&\max\{-\mu_{p,r,\tilde{c},j}^{\delta,\alpha(\delta)}/\tilde{c},(h_j^\delta(z^i)-r_j)\},\\&j=1,\ldots,m):z^i\in D,\\&L_{p,r}^{\tilde{c},\delta}(z^i,\lambda_{p,r,\tilde{c}}^{\delta,\alpha(\delta)},\mu_{p,r,\tilde{c}}^{\delta,\alpha(\delta)})\to\\&\to\inf_{z\in D}L_{p,r}^{\tilde{c},\delta}(z,\lambda_{p,r,\tilde{c}}^{\delta,\alpha(\delta)},\mu_{p,r,\tilde{c}}^{\delta,\alpha(\delta)}),\\&i\to\infty\}\equiv\overline{\mathrm{conv}}Q_{p,r,\tilde{c}}^{\delta}(\lambda_{p,r,\tilde{c}}^{\delta,\alpha(\delta)},\mu_{p,r,\tilde{c}}^{\delta,\alpha(\delta)}).\end{aligned}\tag{4.6}$$

Одновременно, так как величина $\|\lambda\|^2+|\mu|^2$ конечна для всех $(\lambda,\mu)\in H\times R_+^m$, то в силу следствия (4.3) в [22]

$$\begin{aligned}&\partial R_{p,r}^{\tilde{c},\delta,\alpha(\delta)}(\lambda_{p,r,\tilde{c}}^{\delta,\alpha(\delta)},\mu_{p,r,\tilde{c}}^{\delta,\alpha(\delta)})=\\&\partial\tilde{V}_{p,r}^{\tilde{c},\delta,\alpha(\delta)}(\lambda_{p,r,\tilde{c}}^{\delta,\alpha(\delta)},\mu_{p,r,\tilde{c}}^{\delta,\alpha(\delta)})-\\&-2\alpha(\lambda_{p,r,\tilde{c}}^{\delta,\alpha(\delta)},\mu_{p,r,\tilde{c}}^{\delta,\alpha(\delta)}).\end{aligned}\tag{4.7}$$

Пусть $\delta^s,\ s=1,2,\ldots,$ – произвольная сходящаяся к нулю последовательность положительных чисел. Рассмотрим в этом случае последовательность $z_\kappa^{\tilde{c},\delta^s,i},\ i=1,2,\ldots,$ являющуюся минимизирующей для функции $L_{p,r}^{\tilde{c}+\kappa,\delta^s}(z,\lambda_{p,r,\tilde{c}}^{\delta^s,\alpha(\delta^s)},\mu_{p,r,\tilde{c}}^{\delta^s,\alpha(\delta^s)}),$ $z\in D,\ s=1,2,\ldots$ где $\kappa>0$ – не зависящая от $s=1,2,\ldots$ постоянная. Примем при этом обозначение $z_0^{\tilde{c},\delta^s,i}\equiv z^{\tilde{c},\delta^s,i}$ при $\kappa=0$. Тогда имеем при $i\to\infty$ неравенства

$$\begin{aligned}&L_{p,r}^{\tilde{c}+\kappa,\delta^s}(z_\kappa^{\tilde{c},\delta^s,i},\lambda_{p,r,\tilde{c}}^{\delta^s,\alpha(\delta^s)},\mu_{p,r,\tilde{c}}^{\delta^s,\alpha(\delta^s)})\le\\&\le\tilde{V}_{p,r}^{\tilde{c}+\kappa,\delta^s}(\lambda_{p,r,\tilde{c}}^{\delta^s,\alpha(\delta^s)},\mu_{p,r,\tilde{c}}^{\delta^s,\alpha(\delta^s)})+\varepsilon^{\tilde{c},\delta^s,i},\\&\varepsilon^{\tilde{c},\delta^s,i}\to 0,\ i=1,2,\ldots.\end{aligned}\tag{4.8}$$

В силу ограниченности множества $H\times R_+^m$ и оценок (1.2), (3.8), а также выбора подпоследовательности $i(\tilde{c},s),\ s=1,2,\ldots,$ последовательности $i=1,2,\ldots,$ такой, что $\varepsilon^{\tilde{c},\delta^s,i(\tilde{c},s)}\to 0,\ s\to\infty,$ из последней оценки выводим, что

$$\begin{aligned}&L_{p,r}^{\tilde{c}+\kappa,0}(z_\kappa^{\tilde{c},\delta^s,i(\tilde{c},s)},\lambda_{p,r,\tilde{c}}^{\delta^s,\alpha(\delta^s)},\mu_{p,r,\tilde{c}}^{\delta^s,\alpha(\delta^s)})\le\\&\le\tilde{V}_{p,r}^{\tilde{c}+\kappa,0}(\lambda_{p,r,\tilde{c}}^{\delta^s,\alpha(\delta^s)},\mu_{p,r,\tilde{c}}^{\delta^s,\alpha(\delta^s)})+\varepsilon^{\tilde{c},\delta^s,i(\tilde{c},s)}+\gamma^s,\\&\gamma^s\to 0,\ s\to\infty.\end{aligned}$$

В свою очередь, из этой оценки в силу предельного соотношения (4.5) и непрерывности функции $\tilde{V}_{p,r,c}^{0}$ получаем

$$\begin{aligned}&L_{p,r}^{\tilde{c}+\kappa,0}(z_\kappa^{\tilde{c},\delta^s,i(\tilde{c},s)},\lambda_{p,r,\tilde{c}}^{0},\mu_{p,r,\tilde{c}}^{0})\le\\&\le\tilde{V}_{p,r}^{\tilde{c}+\kappa,0}(\lambda_{p,r,\tilde{c}}^{0},\mu_{p,r,\tilde{c}}^{0})+\tilde{\gamma}^s,\ \tilde{\gamma}^s\to 0,\ s\to\infty.\end{aligned}$$

Последнее неравенство говорит о том, что последовательность $z_\kappa^s\equiv z_\kappa^{\tilde{c},\delta^s,i(\tilde{c},s)},\ s=1,2,\ldots,$ является минимизирующей в задаче

$$\tilde{L}_{p,r}^{\tilde{c}+\kappa,0}(z,\lambda_{p,r,\tilde{c}}^{0},\mu_{p,r,\tilde{c}}^{0})\to\inf,\ z\in D,$$

но такой последовательностью в силу включения $(\lambda_{p,r,\tilde{c}}^{0},\mu_{p,r,\tilde{c}}^{0})\in K_{p,r,\tilde{c}}$, независимости $\kappa>0$ от s и доказанного выше факта, что для любой точки $(\lambda,\mu)\in K_{p,r,\tilde{c}}$ найдется такая последовательность $z^i\in D$, для которой справедливы предельные соотношения (4.4), может являться лишь последовательность с указанными свойствами (4.4). Таким образом, построенная выше последовательность элементов $z_\kappa^s\equiv z_\kappa^{\tilde{c},\delta^s,i(\tilde{c},s)},$ $s=1,2,\ldots,$ будет минимизирующей в задаче $(P_{p,r}^0)$, т.е. представляет собой минимизирующее приближенное решение в этой задаче. Одновременно проведенные выше рассуждения позволяют заключить, что справедливо предельное соотношение

$$\begin{aligned}&\mathrm{Diam}\left(\partial\tilde{V}_{p,r}^{\tilde{c}+\kappa,\delta^s}(\lambda_{p,r,\tilde{c}}^{\delta^s,\alpha(\delta^s)},\mu_{p,r,\tilde{c}}^{\delta^s,\alpha(\delta^s)})\right)=\\&=\mathrm{Diam}\left(\overline{\mathrm{conv}}Q_{p,r,\tilde{c}+\kappa}^{\delta^s}(\lambda_{p,r,\tilde{c}}^{\delta,\alpha(\delta)},\mu_{p,r,\tilde{c}}^{\delta,\alpha(\delta)})\right)\to 0,\ s\to\infty,\end{aligned}$$

из которого следует, что мы можем опять же в силу равенства (3.9) без ограничения общности одновременно считать, что

$$\begin{aligned}&\max\{-\mu_{p,r,\tilde{c},j}^{\delta^s,\alpha(\delta^s)}/(\tilde{c}+\kappa),(h_j^{\delta^s}(z_\kappa^s)-r_j)\}\to 0,\\&s\to\infty,\ j=1,\ldots,m,\end{aligned}$$

откуда, в свою очередь, следует, что

$$\max\{-\mu^{\delta^s,\alpha(\delta^s)}_{p,r,\tilde{c},j}/(\tilde{c}+\kappa),(h^0_j(z^s_\kappa)-r_j)\}\to 0,$$
$$s\to\infty,\ j=1,\dots,m.$$

Таким образом, в общей ситуации для построения минимизирующей последовательности в задаче $(P^0_{p,r})$ в случае существования вектора Куна–Таккера требуется на каждом шаге итерационного процесса решать задачу минимизации МФЛ при двух значениях штрафного коэффициента c: при $c=\tilde{c}$ и при $c=\tilde{c}+\kappa$. В то же время во многих важных частных случаях, когда известна дополнительная информация о дифференциальных свойствах функции β^0 в конкретной точке $(p,r)\in\mathrm{dom}\,\beta^0$, минимизирующей последовательностью в задаче $(P^0_{p,r})$ будет построенная выше последовательность $z^s_\kappa\equiv z^{\tilde{c},\delta^s,i(\tilde{c},s)}_\kappa$, $s=1,2,\dots$, при $\kappa=0$: $z^s_0\equiv z^s$, $s=1,2,\dots$. Рассмотрим такой частный случай.

Пусть имеется информация о том, что $\partial^P\beta^0(p,r)$ содержит минимальный по норме элемент, который мы обозначим через $\zeta_m\equiv(\zeta_{p,m},\zeta_{r,m})$. Это будет заведомо так, если, например, $\partial^P\beta^0(p,r)$ – замкнутое множество. Тогда

$$\beta^0(p,r)-\langle\zeta_{p,m},p\rangle-\langle\zeta_{r,m},r\rangle<$$
$$<\beta^0(p',r')-\langle\zeta_{p,m},p'\rangle-\langle\zeta_{r,m},r'\rangle+$$
$$+l\|p'-p\|^2+l|r'-r|^2\ \ \forall(p',r')\in H\times R^m,(p',r')\neq(p,r)$$

для достаточно большого $l>0$. Поэтому для всех $c\geq l$, так как $\zeta_m\in K_{p,r,c}$, а $K_{p,r,c}\subset\partial^P\beta^0(p,r)$, замкнутое выпуклое множество $K_{p,r,c}$ точек максимума вогнутой функции $\tilde{V}^{c,0}_{p,r}$ будет содержать минимальный по норме элемент ζ_m. Тогда при $\tilde{c}>c$ в предельном соотношении (4.5) элемент $(\lambda^0_{p,r,\tilde{c}},\mu^0_{p,r,\tilde{c}})$ будет совпадать с элементом ζ_m, и, по этой причине, минимизирующей последовательностью в задаче минимизации

$$\tilde{L}^{\tilde{c},0}_{p,r}(z,\lambda^0_{p,r,\tilde{c}},\mu^0_{p,r,\tilde{c}})\to\inf,\ \ z\in D$$

может быть лишь последовательность с указанными свойствами (4.4).

В силу следствия 3.1 штрафной коэффициент l можно считать столь большим, что в задаче минимизации МФЛ

$$f^0(z)-\langle\zeta_{p,m},g^0(z)-p\rangle-\langle\zeta_{r,m},h^0(z)+y-r\rangle+$$
$$+l\|g^0(z)-p\|^2+l|h^0(z)+y-r|^2\to\inf,\ z\in D,\ y\in R^m_+$$

равенство

$$\inf_{(z,y)\in D\times R^m_+}(f^0(z)-\langle\zeta_{p,m},g^0(z)-p\rangle-\langle\zeta_{r,m},h^0(z)+y-r\rangle+$$
$$+l\|g^0(z)-p\|^2+l|h^0(z)+y-r|^2)=\beta^0(p,r)$$

выполняется лишь при $\zeta\equiv(\zeta_{p,m},\zeta_{r,m})\in\partial^P\beta^0(p,r)$ и минимизирующей в ней является лишь последовательность (z^k,y^k), $k=1,2,\dots$, такая, что $f^0(z^k)\to\beta^0(p,r)$, $g^0(z^k)\to p$, $h^0(z^k)+y^k\to r$, $k\to\infty$, и никакая другая последовательность.

Действительно, в этом случае для всех $(\lambda,\mu)\in K_{p,r,c}$ в силу конструкции МФЛ $L^{c,0}_{p,r}(z,\lambda,\mu)$ и выражения (3.9) для супердифференциала функции $\tilde{V}^{c,0}_{p,r}(\lambda,\mu)$ при всех достаточно больших $c>0$ справедливы соотношения

$$\{0\}=Q^0_{p,r,c}(\lambda,\mu)=\partial V^{c,0}_{p,r}(\lambda,\mu)\ \ \forall(\lambda,\mu)\in K_{p,r,c}.\qquad(4.9)$$

Пусть далее определенное выше $\tilde{c}>0$ столь велико, что для него справедливо указанное обстоятельство.

Пусть δ^s, $s=1,2,\dots$, – произвольная сходящаяся к нулю последовательность положительных чисел. Практически повторяя проведенные выше рассуждения, рассмотрим последовательность $z^{\tilde{c},\delta^s,i}$, $i=1,2,\dots$, являющуюся минимизирующей для функции $L^{\tilde{c},\delta^s}_{p,r}(z,\lambda^{\delta^s,\alpha(\delta^s)}_{p,r,\tilde{c}},\mu^{\delta^s,\alpha(\delta^s)}_{p,r,\tilde{c}})$, $z\in D$, $s=1,2,\dots$. Тогда имеем опять при $i\to\infty$ неравенства

$$L^{\tilde{c},\delta^s}_{p,r}(z^{\tilde{c},\delta^s,i},\lambda^{\delta^s,\alpha(\delta^s)}_{p,r,\tilde{c}},\mu^{\delta^s,\alpha(\delta^s)}_{p,r,\tilde{c}})\leq$$
$$\leq\tilde{V}^{\tilde{c},\delta^s}_{p,r}(\lambda^{\delta^s,\alpha(\delta^s)}_{p,r,\tilde{c}},\mu^{\delta^s,\alpha(\delta^s)}_{p,r,\tilde{c}})+\varepsilon^{\tilde{c},\delta^s,i},\ i=1,2,\dots.$$

В силу оценок (1.2), (3.9), предельного соотношения (4.5) и непрерывности функции $\tilde{V}^{\tilde{c},0}_{p,r}$, выбирая подпоследовательность $i(\tilde{c},s)$, $s=1,2,\dots$, последовательности $i=1,2,\dots$ так, что $\varepsilon^{\tilde{c},\delta^s,i(\tilde{c},s)}\to 0$, $s\to\infty$, из последних оценок выводим, что

$$L^{\tilde{c},0}_{p,r}(z^{\tilde{c},\delta^s,i(\tilde{c},s)},\zeta_{p,m},\zeta_{r,m})\leq$$
$$\leq\tilde{V}^{\tilde{c},0}_{p,r}(\zeta_{p,m},\zeta_{r,m})+\tilde{\gamma}^s,\ \tilde{\gamma}^s\to 0,\ s\to\infty.$$

Последнее неравенство говорит о том, что последовательность $z^s\equiv z^{\tilde{c},\delta^s,i(\tilde{c},s)}$, $s=1,2,\dots$, является минимизирующей в задаче

$$\tilde{L}^{\tilde{c},0}_{p,r}(z,\lambda^0_{p,r,\tilde{c}},\mu^0_{p,r,\tilde{c}})\to\inf,\ z\in D,$$

но такой последовательностью в силу включения $(\lambda^0_{p,r,\tilde{c}},\mu^0_{p,r,\tilde{c}})\equiv(\zeta_{p,m},\zeta_{r,m})\in K_{p,r,\tilde{c}}$ и соотношений (4.9) может быть лишь такая последовательность $z^i\in D$, для которой справедливы предельные соотношения (4.4). Таким образом, сконструированная последовательность $z^s \equiv z^{\tilde{c},\delta^s,i(\tilde{c},s)}$, $s=1,2,\ldots,$ будет минимизирующей в задаче $(P^0_{p,r})$.

Подытоживая сказанное, можно утверждать, что для процесса построения минимизирующего приближенного решения в задаче $(P^0_{p,r})$ наиважнейшее значение имеет точность решения задачи минимизации функции Лагранжа $L^{c,0}_{p,r}(z,\lambda^{\delta,\alpha(\delta)}_{p,r,\tilde{c}},\mu^{\delta,\alpha(\delta)}_{p,r,\tilde{c}})$, $z\in D$, при каждом $\delta>0$. Если такая минимизация может быть проведена с любой наперед заданной точностью, то в задаче $(P^0_{p,r})$ конструктивно указывается минимизирующее приближенное решение. Таким образом, справедлива следующая

Теорема 4.1. *Пусть* $\delta^s, s=1,2,\ldots,$ *– произвольная сходящаяся к нулю последовательность положительных чисел. Тогда если задача* $(P^0_{p,r})$ *обладает вектором Куна–Таккера в указанном обобщенном смысле, то найдется достаточно большое* $\tilde{c}>0$, *такое, что справедливы предельные соотношения*

$$f^0(z^s)\to\beta^0(p,r),\ g^0(z^s)-p\to 0,$$
$$\max\{0,h^0_j(z^s)-r_j\}\to 0,$$
$$\max\{-\mu^{\delta^s,\alpha(\delta^s)}_{p,r,\tilde{c},j}/(\tilde{c}+\kappa),h^0_j(z^s)-r_j\},\ j=1,2,\ldots,m,$$
$$(\lambda^{\delta^s,\alpha(\delta^s)}_{p,r,\tilde{c}},\mu^{\delta^s,\alpha(\delta^s)}_{p,r,\tilde{c}})\to(\lambda^0_{\tilde{c}},\mu^0_{\tilde{c}}),$$
$$V^{\tilde{c},0}_{p,r}(\lambda^{\delta^s,\alpha(\delta^s)}_{p,r,\tilde{c}},\mu^{\delta^s,\alpha(\delta^s)}_{p,r,\tilde{c}})\to\beta^0(p,r),\ \ s\to\infty,$$

где z^s, $s=1,2,\ldots,$ *– субоптимальные элементы, минимизирующие при положительном* $\kappa>0$ *с точностью* $\varepsilon^s:\varepsilon^s\to 0,\ s\to\infty$ *МФЛ* $L^{\tilde{c}+\kappa,\delta^s}_{p,r}(z,\lambda^{\delta^s,\alpha(\delta^s)}_{p,r,\tilde{c}},\ \mu^{\delta^s,\alpha(\delta^s)}_{p,r,\tilde{c}})$, $z\in D$:

$$L^{\tilde{c}+\kappa,\delta^s}_{p,r}(z^s,\lambda^{\delta^s,\alpha(\delta^s)}_{p,r,\tilde{c}},\mu^{\delta^s,\alpha(\delta^s)}_{p,r,\tilde{c}})\le$$
$$\le\inf_{z\in D}L^{\tilde{c}+\kappa,\delta^s}_{p,r}(z,\lambda^{\delta^s,\alpha(\delta^s)}_{p,r,\tilde{c}},\mu^{\delta^s,\alpha(\delta^s)}_{p,r,\tilde{c}})+\varepsilon^s,$$

где $(\lambda^{\delta^s,\alpha(\delta^s)}_{p,r,\tilde{c}},\mu^{\delta^s,\alpha(\delta^s)}_{p,r,\tilde{c}})$ *– элементы, максимизирующие на множестве* $H\times R^m_+$ *сильно вогнутый функционал* $R^{\tilde{c},\delta^s,\alpha(\delta^s)}_{p,r}$, $\delta^s/\alpha(\delta^s)\to 0,\ s\to\infty$, $(\lambda^0_{\tilde{c}},\mu^0_{\tilde{c}})$ *– минимальный по норме в* $K_{p,r,\tilde{c}}$ *обобщенный вектор Куна–Таккера задачи* $(P^0_{p,r})$. *При этом если функция значений* β *обладает в точке* (p,r) *некоторыми дифференциальными свойствами в указанном выше смысле, то величину* κ *можно считать равной нулю.*

Сформулированное в теореме 4.1 утверждение можно переформулировать в виде регуляризованной теоремы Куна–Таккера в недифференциальной секвенциальной форме для нелинейной задачи $(P^0_{p,r})$ (см. например, [13]), но это не является предметом данной статьи. Отметим лишь, что характерным свойством такой теоремы является устойчивость по отношению к ошибкам исходных данных.

Рассмотрим в заключение простой пример задачи нелинейного программирования, показывающий, что формальный процесс, аналогичный описанному в теореме 4.1, но без регуляризации двойственной задачи, не обеспечивает устойчивого построения минимизирующей последовательности. В качестве такой задачи возьмем конечномерную задачу выпуклого программирования с ограниченным допустимым множеством. В этом случае в качестве элементов минимизирующих последовательностей можно использовать точки минимума модифицированной функции Лагранжа.

Пример 4.1. *Пусть имеется задача минимизации сильно выпуклой квадратичной функции двух переменных на множестве, задаваемом аффинным ограничением типа равенства*

$$x_1^2+x_2^2\to\min,\ x_1+x_2=1,\ 0\cdot x_1+0\cdot x_2=0,$$
$$(x_1,x_2)\in D\equiv\{0\le x_1\le 1;\ 0\le x_2\le 1\},$$

точное решение которой равно $\bar{x}=\left(\frac{1}{2},\frac{1}{2}\right)$.

В силу простоты рассматриваемой задачи можно непосредственно установить, что вектор $(\lambda_1,\lambda_2)=(-1,\alpha)$, $\alpha\in R$, *является вектором Куна–Таккера. Одновременно в силу ее выпуклости каждый такой вектор является и обобщенным вектором Куна–Таккера при любом* $c>0$.

Рассмотрим возмущенную задачу при $\delta>0$:

$$x_1^2+x_2^2\to\min,\ x_1+x_2=1,\ \delta\cdot x_2=\delta^2,(x_1,x_2)\in D.$$

Ее модифицированная функция Лагранжа имеет вид

$$L^\delta_c(x_1,x_2,\lambda_1,\lambda_2)=(1+c)x_1+(1+c+\delta^2c)x_2+$$
$$+2cx_1x_2+(\lambda_1-2c)x_1+(\lambda_1-2c+\delta\lambda_2-2c\delta^3)x_2+$$
$$+(c-\lambda_1-\delta^2\lambda_2+\delta^4c).$$

Градиент модифицированной функции Лагранжа имеет вид

$$L^\delta_{c,x}=(2x_1+\lambda_1+2cx_1+2cx_2-2c;$$
$$2x_2+\lambda_1+\delta\lambda_2+2cx_1+2cx_2-2c+2c\delta^2x_2-2c\delta^3).$$

Непосредственной проверкой убеждаемся, что градиент модифицированной функции Лагранжа, взятой при значениях $\lambda_1=\lambda_1^\delta=$ $=2\delta-2$, $\lambda_2=\lambda_2^\delta=-\dfrac{4\delta-2}{\delta}$, *зануляется при* $x_1=x_1^\delta=1-\delta$, $x_2=x_2^\delta=\delta$. *Так как модифицированная функция Лагранжа является выпуклой* *по* x *и точка* $(1-\delta,\delta)$, *являясь решением возмущенной задачи* (*как единственная точка, удовлетворяющая ограничениям*) *и одновременно внутренней точкой множества* D, *зануляет ее градиент, то она доставляет минимум этой функции на* D, *т.е.* $(x_1^\delta)^2+(x_2^\delta)^2\le$ $\le L_c^\delta(x_1,x_2,\lambda_1^\delta,\lambda_2^\delta)\ \ \forall x\in D$.

Таким образом, вектор $\left(\lambda_1^\delta,\lambda_2^\delta\right)$ *является обобщенным вектором Куна–Таккера возмущенной задачи, доставляющим максимум в соответствующей двойственной задаче, а взятый с противоположным знаком, т.е. вектор* $-\left(\lambda_1^\delta,\lambda_2^\delta\right)$, *является проксимальным субградиентом функции значений возмущенной задачи, совпадающим, в силу выпуклости последней, с ее субдифференциалом в смысле выпуклого анализа. Но такой субдифференциал может состоять только из одной точки, так как непосредственной проверкой легко устанавливается, что функция значений двойственной задачи является гладкой функцией двух переменных.*

Итак, проведенные рассуждения говорят о том, что решение возмущенной модифицированной двойственной задачи при любом $c>0$ *доставляется вектором* $\lambda^\delta=(2\delta-2,\ -\dfrac{2(2\delta-1)}{\delta})$. *Заметим, что он не зависит от штрафного коэффициента* $c>0$ *в силу выпуклости исходной задачи. Соответствующая найденному решению модифицированной двойственной задачи точка минимума МФЛ* $L_c^\delta(x_1,x_2,$ $2\delta-2,\ -\dfrac{2(2\delta-1)}{\delta})$ *равна* $x^\delta=(1-\delta,\ \delta)$. *С одной, формальной, стороны, эта точка претендует, чтобы называться приближением к решению исходной невозмущенной задачи. С другой же стороны, она не сходится к ее единственному точному решению* $\bar{x}=\left(\dfrac{1}{2},\dfrac{1}{2}\right)$ *при* $\delta\to 0$.

В заключение настоящей работы выражаю благодарность своему научному руководителю профессору М.И. Сумину за постановку задачи и внимание к работе.

Работа подготовлена при финансовой поддержке РФФИ (код проекта 12-01-00199-а), а также Минобрнауки РФ в рамках государственного задания на оказание услуг в 2012–2014 гг. подведомственными высшими учебными заведениями (шифр заявки 1.1907.2011).

Список литературы

1. 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.). Екатеринбург: Изд-во Ин-та математики и механики УрО РАН, 2000. С. 66–69.
2. Сумин М.И. Регуляризованный градиентный двойственный метод решения обратной задачи финального наблюдения для параболического уравнения // Журн. вычисл. матем. и матем. физ. 2004. Т. 44. № 11. С. 2001–2019.
3. 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.
4. Sumin M.I. Parametric Dual Regularization in a Linear-Convex Mathematical Programming // Computational Optimization: New Research Developments. Chapter 10. New York: Nova Science Publishers Inc., 2010. P. 265–311.
5. Сумин М.И. Регуляризованный двойственный метод решения нелинейной задачи математического программирования // Журн. вычисл. матем. и матем. физ. 2007. Т. 47. № 5. С. 796816.
6. Sumin M.I. Parametric Dual Regularization in a Nonlinear Mathematical Programming // Advances in Mathematics Research. Volume 11. Chapter 5. New York: Nova Science Publishers Inc., 2010. P. 103–134.
7. Алексеев В.М., Тихомиров В.М., Фомин С.В. Оптимальное управление. М.: Наука, 1979.
8. Гольштейн Е.Г. Теория двойственности в математическом программировании и ее приложения. М.: Наука, 1971.
9. Гольштейн Е.Г., Третьяков Н.В. Модифицированные функции Лагранжа. Теория и методы оптимизации. М.: Наука, 1989.
10. Левитин Е.С. Теория возмущений в математическом программировании и приложения. М.: Наука, 1992.
11. Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. М.: Наука, 1986.
12. Васильев Ф.П. Методы оптимизации. М.: Факториал Пресс, 2002.
13. Сумин М.И. Регуляризованная параметрическая теорема Куна–Таккера в гильбертовом пространстве // Журн. вычисл. матем. и матем. физ. 2011. Т. 51. № 9. С. 1594–1615.
14. Бертсекас Д. Условная оптимизация и методы множителей Лагранжа. М.: Радио и связь, 1987.
15. 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. №. 2. P.431–452; Part II: Applications // Can. J. Math. 1987. V. 39. №. 2. P. 428–472.
16. Loewen P.D. Optimal Control via Nonsmooth Analysis. CRM Proceedings and Lecture Notes. Vol. 2. Providence, RI: Amer. Math. Soc., 1993.
17. 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.
18. Mordukhovich B.S. Variational Analysis and Generalized Differentiation. I: Basic Theory. Berlin: Springer, 2006.
19. Mordukhovich B.S. Variational Analysis and Generalized Differentiation. II: Applications. Berlin: Springer, 2006.

20. Сумин М.И. Субоптимальное управление системами с распределенными параметрами: минимизирующие последовательности, функция значений //Журн. вычисл. матем. и матем. физ. 1997. Т. 37. № 1. С. 23–41.

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

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

ON STABLE CONSTRUCTION OF MINIMIZING SEQUENCES IN NONLINEAR PROGRAMMING PROBLEMS WITH EQUALITY AND INEQUALITY CONSTRAINTS

A.V. Kanatov

The article describes the dual regularization approach applied to the general parametric nonlinear programming problem in a Hilbert space with an infinite-dimensional equality constraint and a finite number of functional inequality constraints. This method provides a stable (in relation to the input data errors) construction of minimizing sequence elements in the original problem from elements of sequences that minimize the modified Lagrange function taken at the values of the dual variables of the respective maximizing sequence in the modified dual problem. In particular, it is shown how generalized differentiability properties of the lower semicontinuous value functions in infinite-dimensional problems of mathematical programming generate corresponding constructions of the modified Lagrange functions. An example is given to illustrate the instability of the formal construction of minimizing sequences without the solution regularization of the modified dual problem.

Keywords: nonlinear programming, parametric problem, sequential optimization, proximal subgradient, Lagrange principle, Kuhn-Tucker vector, minimizing approximate solution, duality, regularization.