В. Г. Жадан, Об одном классе итеративных методов решения задач выпуклого программирования, *Ж. вычисл. матем. и матем. физ.*, 1984, том 24, номер 5, 665–676

ЖУРНАЛ
ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ

Том 24, 1984 № 5

УДК 519.6:519.853.3

ОБ ОДНОМ КЛАССЕ ИТЕРАТИВНЫХ МЕТОДОВ РЕШЕНИЯ ЗАДАЧ ВЫПУКЛОГО ПРОГРАММИРОВАНИЯ

ЖАДАН В. Г.

(Москва)

Для решения задачи выпуклого программирования предлагается класс итеративных методов, основанных на применении модифицированных функций Лагранжа. Дано обоснование этих методов. Показано, что в случае задачи линейного программирования они сходятся за конечное число шагов.

§ 1. Введение

Рассматривается задача выпуклого программирования: найти

$$(1.1)\qquad \min_{x\in X\cap Q} f(x),\qquad X=\{x\in E_n\,|\,g^i(x)\leqslant 0,\quad i=1,2,\ldots,m\},$$

где E_n есть n-мерное евклидово пространство, $f(x)$, $g^i(x)$, $i=1, 2,\ldots, m$, выпуклые, непрерывно дифференцируемые функции на E_n, Q – выпуклое замкнутое множество в E_n.

Пусть X_* – множество решений задачи (1.1). Оно выпукло и, если $f(x)$ – строго выпуклая функция, состоит из единственной точки. Всюду считается, что X_* не пусто.

Составим для задачи (1.1) функцию Лагранжа

$$(1.2)\qquad L(x, u)=f(x)+\langle u, g(x)\rangle.$$

Здесь $u\in E_m^+$; $g(x)$ есть m-мерная вектор-функция, i-я компонента которой $g^i(x)$; E_m^+ – неотрицательный ортант E_m; $\langle\cdot,\cdot\rangle$ – евклидово скалярное произведение.

Обозначим через f_x, L_x, L_u вектор-столбцы градиентов функций $f(x)$ и $L(x, u)$, через g_x – матрицу первых производных функции g размера $n\times m$. Индекс т означает транспонирование, символ $\|\cdot\|$ – евклидову норму. Для того чтобы точка x_* была решением задачи (1.1), достаточно (см. [1]), чтобы нашелся вектор $u_*\in E_m^+$ такой, что

$$(1.3)\qquad L_x(x_*, u_*)\in K^*(Q, x_*),$$

$$(1.4)\qquad L_u(x_*, u_*)\leqslant 0,\quad \langle L_u(x_*, u_*), u_*\rangle=0.$$

Здесь $K^*(Q, x)$ – конус, двойственный к конусу допустимых направлений $K(Q, x)$ для множества Q в точке x. Если ограничения в задаче (1.1) удовлетворяют условию Слейтера, т. е. существует точка $\bar{x}\in Q$ такая, что $g^i(\bar{x})<0$ для всех $i=1, 2,\ldots, m$, то условия (1.3), (1.4) являются одновременно и необходимыми. Точки x_* и u_* вместе образуют седловую точку $z_*=[x_*, u_*]$ функции Лагранжа на множестве $Q\times E_m^+$, а именно:

$$L(x_*, u)\leqslant L(x_*, u_*)\leqslant L(x, u_*)\qquad \forall x\in Q,\quad u\in E_m^+.$$

Точки $z_*=[x_*, u_*]$, для которых выполнено (1.3), (1.4), будем называть точками Куна — Таккера. Их множество обозначим через Z_*.

Среди методов решения задачи (1.1), основанных на отыскании точек множества Z_*, за последнее время особенно интенсивно развивались методы, использующие различные модифицированные и обобщенные функции Лагранжа (см., например, [2]—[4]). В большинстве случаев при конструировании таких функций стремились улучшить свойства функции Лагранжа (1.2), сохраняя при этом неизменным множество ее седловых точек Z_*. Возможны также другие подходы. В [5]—[6] предложен класс модифицированных функций Лагранжа, для которых точки из множества Z_* уже не являются седловыми, а лишь точками локального минимакса. Было установлено, что в случае общей задачи нелинейного программирования методы, построенные на их основе, обладают локальной сходимостью. В настоящей работе применительно к задаче выпуклого программирования (1.1) уточняется вид введенных в [5] модифицированных функций Лагранжа. Показано, что при довольно естественных предположениях относительно регулярности задачи (1.1) численные методы, использующие эти модификации, сходятся из любой начальной точки, принадлежащей множеству Q, причем в случае задачи линейного программирования имеет место сходимость за конечное число шагов. Все эти утверждения обобщают соответствующие результаты, полученные в [6] для модифицированной функции Лагранжа частного вида.

§ 2. Вспомогательная функция

Ниже будем пользоваться обозначениями, обычными для выпуклого анализа [1], [7]. Пусть $h(x)$ — выпуклая функция, λ — положительное число. Тогда функция, сопряженная к $h(x)$, имеет вид

$$h^*(x)=\sup_{y\in E_n}\{\langle x,y\rangle-h(y)\};$$

$(h\lambda)(x)=\lambda h(x/\lambda)$ — произведение справа функции $h(x)$ на константу λ; $\operatorname{dom} h(x)=\{x\,|\,h(x)<+\infty\}$ — эффективная область $h(x)$; $(h0^+)(x)$ — рецессивная функция $h(x)$; если $h(x)$ — замкнутая собственная функция и $0\in\operatorname{dom} h(x)$, то

$$(h0^+)(x)=\lim_{\lambda\to+0}(h\lambda)(x).$$

Такие обозначения будут применяться и в случае, когда функция h зависит от двух аргументов, но выпукла лишь по первому. Везде имеется в виду и нигде особо не оговаривается, что они относятся только к этому первому аргументу. Так, например, $h^*(x, y)$ — функция, сопряженная к $h(x, y)$ при каждом фиксированном y. Если D — выпуклое множество, то через $\operatorname{int} D$ будем обозначать его внутренность, через $\delta(x|D)$ и $\delta^*(x|D)$ — соответственно, индикаторную и опорную функции D.

Введем на множестве $E_n\times Q$ вспомогательную функцию $M(p, q)$, на которую наложим следующее требование.

Условие А. Для любого фиксированного $q\in Q$ функция $M(p, q)$ определена и выпукла по p на E_n, достигает своего минимума по p в точках множества $K^*(Q, q)$ и $(M0^+)(p, q)=\delta^*(p|q-Q)$.

Таким образом, функция $M(p, q)$, рассматриваемая при фиксированном q как функция переменного p, является замкнутой собственной выпуклой. Более того, если множество Q не ограничено, то вдоль любого ненулевого направления, не принадлежащего конусу, двойственному к рецессивному конусу множества Q, производная функции $M(p, q)$ по этому направлению, возрастая, стремится к бесконечности. В случае когда Q совпадает со всем пространством E_n, из условия А следует, что $M(p, q)$ должна быть кофинитной, т. е. $(M0^+)(p, q)=+\infty$ для всех $p\neq 0$.

Если от функции $M(p,q)$ перейти к сопряженной ей $M^*(p,q)$, то получим двойственное к условию А

Условие А*. Для любого фиксированного $q\in Q$ функция $M^*(p,q)$ выпукла и замкнута по переменной p, достигает своего минимума по p в нуле и $\operatorname{dom} M^*(p,q)=q-Q$.

Имеет место

Лемма 1. *Пусть $M(p,q)$ — непрерывно дифференцируемая по первому аргументу функция, удовлетворяющая условию* А, *тогда сопряженная функция $M^*(p,q)$ удовлетворяет условию* А*.

Доказательство. Возьмем произвольное $q\in Q$ и зафиксируем его. Выпуклость и замкнутость функций $M^*(p,q)$ следует из основных свойств сопряженной функции. Далее, согласно теореме 13.3 из [7], опорная функция эффективного множества функции $M^*(p,q)$ совпадает с рецессивной функцией $M(p,q)$. Отсюда и из условия А получаем, что $\operatorname{dom} M^*(p,q)=q-Q$.

Осталось показать, что $M^*(p,q)>M^*(0,q)$ для любого p, отличного от нуля. Так как $M^*(p,q)$ выпукла и замкнута по p, то

$$\inf_{p\in E_n} M^*(p,q)=-M(0,q).$$

Но поскольку, какое бы $q\in Q$ ни было взято, обязательно $0\in K^*(Q,q)$, то

$$M^*(0,q)=-\min_{p\in E_n} M(p,q)=-M(0,q).$$

Если предположить, что найдется $\bar p\neq 0$ такое, что $M^*(\bar p,q)=M^*(0,q)$, то, на основании определения сопряженной функции, $\langle \bar p, p\rangle - M(p,q)\leqslant -M(0,q)$ для всех $p\in E_n$. Поэтому $\bar p$ является субградиентом функции $M(p,q)$ по p в нуле, что невозможно. Лемма доказана.

Аналогично устанавливается справедливость обратного утверждения.

Лемма 2. *Пусть $M^*(p,q)$ — удовлетворяющая условию* А* *функция, которая при каждом $q\in Q$ непрерывно дифференцируема по p внутри множества $q-Q$, тогда сопряженная функция $M(p,q)$ удовлетворяет условию* А.

Приведем примеры функций $M(p,q)$, для которых выполнено условие А. Предположим сначала, что в задаче (1.1) множество Q совпадает со всем пространством E_n. В этом случае для любого $x_*\in X_*$ конус $K^*(E_n, x_*)$ состоит лишь из нулевого вектора, условие (1.3) заменяется на $L_x(x_*, u_*)=0$. Благодаря этому обстоятельству зависимость $M(p,q)$ от второго аргумента оказывается несущественной и ее можно отбросить. Функцию $M(p,q)$ будем строить в следующем сепарабельном виде:

$$M(p)=\sum_{j=1}^{n}\varphi(p^j), \tag{2.1}$$

где $\varphi(t)$ — выпуклая кофинитная функция скалярного аргумента, достигающая своего минимума на E_1 в нуле. В качестве $\varphi(t)$ могут быть использованы, например, функции

$$\varphi_1(t)=t^2/2,\qquad \varphi_2(t)=\operatorname{ch} t,\qquad \varphi_3(t)=e^t+(t-1)^2/2.$$

К несепарабельным примерам функции $M(p)$ может быть отнесен квадрат чебышёвской нормы

$$M(p)=\max_{1\leq j\leq n}\{|p^j|^2\}.$$

Рассмотрим другой важный частный случай задачи (1.1), когда Q является множеством «параллелепипедного» типа:

$$Q=\{x\in E_n\,|\,a^j\leq x^j\leq b^j,\quad j=1,2,\ldots,n\}. \tag{2.2}$$

Наиболее простой способ построить $M(p,q)$ для данного множества (как, впрочем, и для других замкнутых выпуклых множеств) — это обратиться к сопряженной функции $M^*(p,q)$. Пусть имеется выпуклая, непрерывно дифференцируемая кофинитная функция $F(p)$, определенная на всем пространстве E_n и достигающая своего минимума на E_n в нуле. Тогда, согласно лемме 2,

$$M(p,q)=G^*(p,q), \tag{2.3}$$

где $G(p,q)=F(p)+\delta(p|q-Q)$, удовлетворяет условию А.

Если, подобно (2.1), строить функцию $F(p)$ в сепарабельном виде, причем считать, что $\varphi(t)$ непрерывно дифференцируема на E_1, то придем к функции

$$M(p,q)=\sum_{j=1}^{n}\psi^j(p^j,q^j).$$

Здесь

$$\psi^j(t,\tau)=\begin{cases}\varphi^*(t), & \beta_1^j(\tau)\leq t\leq\beta_2^j(\tau),\\ \varphi^*(\beta_1^j(\tau))+(\tau-b^j)[t-\beta_1^j(\tau)], & t<\beta_1^j(\tau),\\ \varphi^*(\beta_2^j(\tau))+(\tau-a^j)[t-\beta_2^j(\tau)], & t>\beta_2^j(\tau),\end{cases}$$

$\beta_1^j(\tau)$ — максимальный корень уравнения $\varphi^{*\prime}(t)=\tau-b^j$, $\beta_2^j(\tau)$ — минимальный корень уравнения $\varphi^{*\prime}(t)=\tau-a^j$, штрих означает дифференцирование. Производная $\psi_t^j(t,\tau)$ представима в виде

$$\psi_t^j(t,\tau)=\begin{cases}\varphi^{*\prime}(t) & \beta_1^j(\tau)\leq t\leq\beta_2^j(\tau),\\ \tau-b^j, & t<\beta_1^j(\tau),\\ \tau-a^j, & t>\beta_2^j(\tau).\end{cases}$$

Взяв в качестве $\varphi(t)$, например, функцию $\varphi_1(t)$, получим

$$\psi^j(t,\tau)=[\tau-(\tau-t)_{P_j}]\{t-0.5[\tau-(\tau-t)_{P_j}]\},$$

где $(\cdot)_{P_j}$ — оператор проектирования на отрезок $[a^j,b^j]$.

Ниже наряду с функцией $M(p,q)$ используется ее произведение справа на положительную константу α. Такое умножение сохраняет неизменными все основные свойства $M(p,q)$, а именно: $(M\alpha)(p,q)$ для произвольного $\alpha>0$ удовлетворяет условию А. Установим еще одно свойство функции $(M\alpha)(p,q)$. Обозначим через $\partial_p(M\alpha)(p,q)$ ее субдифференциал по первому аргументу.

Лемма 3. *Если функция $M(p,q)$ удовлетворяет условию* А, *то $\partial_p(M\alpha)(p,q)\subset q-Q$ для любых $[p,q]\in E_n\times Q$ и $\alpha>0$*.

Доказательство. Возьмем произвольное $p\in E_n$. Пусть $p^*\in\in\partial_p(M\alpha)(p,q)$. На основании теоремы 23.5 из [7], имеет место двойственная связь

(2.4) $\qquad p\in\partial_p(M\alpha)^*(p^*, q)=\alpha\partial_p M^*(p^*, q).$

Если теперь предположить, что $p^*\in q-Q$, то, согласно лемме 1, $p^*\in \in \operatorname{dom} M^*(p, q)$. Поэтому субдифференциал $\partial_p M^*(p^*, q)$ должен быть пустым множеством, что противоречит включению (2.4). Лемма доказана.

§ 3. Модифицированная функция Лагранжа

Используя вспомогательную функцию $M(p, q)$, составим модифицированную функцию Лагранжа

$$H(x, u, \alpha)=L(x, u)-(M\alpha)(L_x(x, u), x),$$

где α — некоторый положительный параметр. Непосредственной проверкой несложно убедиться, что $H(x, u, \alpha)$ вогнута по u на E_m^+ для любых фиксированных $x\in Q$ и $\alpha>0$.

Рассмотрим задачу отыскания

(3.1) $$\max_{u\in E_m^+} H(x, u, \alpha).$$

Обозначим через $U(x, \alpha)$ множество ее решений, через $u(x, \alpha)$ — произвольное решение из $U(x, \alpha)$. Так как функция $H(x, u, \alpha)$ вогнута по u для любых фиксированных $x\in Q$ и $\alpha>0$, то множество $U(x, \alpha)$ выпукло.

Определение. Ограничения в задаче (1.1) удовлетворяют условию положительной регулярности, если $g_x(x)u\neq 0$ для всех $x\in Q$ и всех ненулевых $u\in E_m^+$.

Теорема 1. *Пусть ограничения в задаче* (1.1) *удовлетворяют условию Слейтера и условию положительной регулярности. Тогда для любых $x\in Q$ и $\alpha>0$ множество $U(x, \alpha)$ не пусто и ограничено.*

Доказательство. Для произвольных фиксированных $x\in Q$ и $\alpha>0$ рассмотрим выпуклую на E_m функцию $\eta(u)=-H(x, u, \alpha)+\delta(u|E_m^+)$. Очевидно, что

$$U(x, \alpha)=\operatorname{Arg}\min_{u\in E_m} \eta(u).$$

В силу теоремы 27.1 из [7], множество минимумов функции $\eta(u)$ не пусто и ограничено тогда и только тогда, когда $0\in \operatorname{int} \operatorname{dom} \eta^*(u)$. А это включение выполняется в том и только том случае, когда $\eta(u)$ не имеет направлений рецессий, т. е.

(3.2) $\qquad (\eta 0^+)(u)>0 \quad \forall u\neq 0.$

Вычислим $(\eta 0^+)(u)$, считая, что $u\neq 0$. Так как $g_x(x)u\neq 0$, когда $u\in E_m^+$, то, согласно определению рецессивной функции,

(3.3) $$(\eta 0^+)(u)=\begin{cases} (M0^+)(g_x(x)u, x)-\langle g(x), u\rangle, & u\in E_m^+,\\ +\infty, & u\notin E_m^+.\end{cases}$$

Пусть $u\in E_m^+$. Используя условие А, получаем

(3.4) $$\begin{aligned}&(M0^+)(g_x(x)u, x)=\delta^*(g_x(x)u|x-Q)=\\ &=\sup_{y\in x-Q}\langle y, g_x(x)u\rangle=\sup_{y\in x-Q}\langle g_x^{\mathrm{T}}(x)y, u\rangle=\\ &=-\inf_{x_1\in Q}\langle g_x^{\mathrm{T}}(x)(x_1-x), u\rangle.\end{aligned}$$

Из условия Слейтера следует, что найдется точка $\bar{x}\in Q$ такая, что $g^i(\bar{x})<$

<0, $i=1,2,\ldots,m$. Поэтому справедливо неравенство $g(x)+$ $+\langle g_x^{\mathrm{T}}(x), \bar{x}-x\rangle \leqslant g(\bar{x})<0$. Отсюда и из (3.3), (3.4) приходим к (3.2). Теорема доказана.

Множество $U(x,\alpha)$ совпадает с субдифференциалом функции $\eta^*(u)$ в нуле. Таким образом, если $\eta^*(u)$ дифференцируема в нуле, то $U(x,\alpha)$ состоит из единственной точки $u(x,\alpha)$.

Приведем два других утверждения относительно задачи (3.1).

Лемма 4. *Пусть выполнены предположения теоремы* 1, *функция* $M(p,q)$ *непрерывна на множестве* $E_n\times Q$. *Тогда для любого компактного множества* $S\subseteq Q$ *и любого* $\bar{\alpha}>0$ *можно указать константу* $C>0$ *такую, что* $\|u\|\leqslant C\alpha$ *для всех* $u\in U(x,\alpha)$, $x\in S$, $\alpha\geqslant\bar{\alpha}$.

Доказательство. Допустим противное. Тогда найдутся последовательность точек $\{x_k\}\subset S$ и последовательность $\{\alpha_k\}$ такие, что

$$u_k\in U(x_k,\alpha_k),\quad \alpha_k\geqslant\bar{\alpha}\ \text{ и } \lim_{k\to\infty}\alpha_k^{-1}\|u_k\|=+\infty.$$

Введя обозначения $\delta_k=1/\alpha_k$, $\bar{\delta}=1/\bar{\alpha}$, $v_k=\delta_k u_k$, будем иметь

$$0<\delta_k\leqslant\bar{\delta},\qquad \lim_{k\to\infty}\|v_k\|=+\infty.$$

Выберем теперь из $\{x_k\}$ и $\{\delta_k\}$ сходящиеся подпоследовательности. Не умаляя общности, можно считать, что сами последовательности $\{x_k\}$ и $\{\delta_k\}$ сходятся к $\hat{x}$ и $\hat{\delta}$ соответственно. Из вида функции $H(x,u,\alpha)$ следует, что $u_k\in U(x_k,\alpha_k)$ только тогда, когда v_k является решением задачи

$$\max_{v\in E_m^+}\langle g(x),v\rangle - M(\delta f_x(x)+g_x(x)v,x) \tag{3.5}$$

при $x=x_k$, $\delta=\delta_k$. Так как множество ее решений $V(x,\delta)$ не пусто и ограничено для любых $x\in S$ и $0\leqslant\delta\leqslant\bar{\delta}$ (в этом можно убедиться, повторяя почти дословно доказательство теоремы 1), то оно будет не пустым и ограниченным также при $x=\hat{x}$, $\delta=\hat{\delta}$. Но тогда, согласно теореме 26.2 из [8], задача (3.5), рассматриваемая как параметрическая задача выпуклого программирования, является β-корректной по x и δ в точке $[\hat{x},\hat{\delta}]\in$ $\in E_{n+1}$. Отсюда последовательность $\{v_k\}$ ограничена и ее предельные точки принадлежат $V(\hat{x},\hat{\delta})$. Полученное противоречие доказывает лемму.

Следствие. Если параметр α фиксирован, то для любого компактного множества $S\subseteq Q$ можно указать константу $C>0$ такую, что $\|u\|\leqslant C$ для всех $u\in U(x,\alpha)$, $x\in S$.

Предположим теперь, что функция $M(p,q)$ непрерывно дифференцируема по первому аргументу. Тогда функция $(M\alpha)(p,q)$ также будет дифференцируемой по этому аргументу и имеет место связь

$$(M\alpha)_p(p,q)=M_p(p|\alpha,q), \tag{3.6}$$

где через M_p и $(M\alpha)_p$ обозначены, соответственно, градиенты функций M и $(M\alpha)$ по p. В этом случае из необходимых условий экстремума для задачи (3.1) следует, что

$$g(x)-g_x^{\mathrm{T}}(x)\,(M\alpha)_p(L_x(x,u),x)\leqslant 0, \tag{3.7}$$

$$\langle u, g(x)-g_x^{\mathrm{T}}(x)\,(M\alpha)_p(L_x(x,u),x)\rangle=0 \tag{3.8}$$

для любых $x\in Q$, $\alpha>0$ и $u\in U(x,\alpha)$.

Лемма 5. *Пусть точка* $z_*=[x_*,u_*]$ *такова, что* $u_*\in U(x_*,\alpha)$ *для не-*

которого $\alpha>0$ *и*

$$(3.9)\qquad (M\alpha)_p(L_x(x_*, u_*), x_*)=0;$$

тогда $x_*\in X_*$.

Доказательство. В силу (3.9), вектор $L_x(x_*, u_*)$ доставляет минимум функции $(M\alpha)(p, x_*)$ по p, поэтому, согласно условию А, имеет место включение (1.3). Осталось проверить выполнение условий (1.4). Но в этом легко убедиться, если подставить $x=x_*$, $u=u_*$ в (3.7), (3.8) и учесть (3.9). Лемма доказана.

§ 4. Численный метод

Ниже приводится прямой метод решения задачи (1.1), основанный на отыскании точек множества Z_*. В нем на каждом шаге пересчитываются лишь прямые переменные, а двойственные находятся из решения вспомогательной задачи максимизации (3.1).

Наложим дополнительное требование.

Условие Б. Функция $M(p, q)$ непрерывна на множестве $E_n\times Q$, непрерывно дифференцируема по p при фиксированном q, и для любых ограниченных замкнутых множеств $P_1\subset E_n$, $Q_1\subseteq Q$ можно указать константу $\gamma=\gamma(P_1, Q_1)>0$ такую, что $\langle M_p(p, q), p\rangle\geqslant\gamma\|M_p(p, q)\|^2$ для всех $p\in P_1$, $q\in Q_1$.

Итерации в предлагаемом методе ведутся по рекуррентной схеме

$$(4.1\text{а})\qquad u_k\in\operatorname{Arg}\max_{u\in E_m^+} H(x_k, u, \alpha),$$

$$(4.1\text{б})\qquad x_{k+1}=x_k-(M\alpha)_p(L_x(x_k, u_k), x_k),$$

где $\alpha>0$. Таким образом, его можно рассматривать как метод простой итерации для решения системы уравнений

$$(4.2)\qquad (M\alpha)_p(L_x(x, u(x, \alpha)), x)=0.$$

В силу утверждения леммы 5, если x_* является корнем системы (4.2), то $x_*\in X_*$.

Сформулируем и докажем теперь теоремы о сходимости процесса (4.1). Будем считать, что градиенты функций $f(x)$ и $g^i(x)$, $i=1, 2, \ldots$ $\ldots, m$, удовлетворяют на E_n условию Липшица с константой l. Тогда имеют место неравенства

$$(4.3)\qquad f(x_2)-f(x_1)-\langle f_x(x_1), x_2-x_1\rangle\leqslant\frac{l}{2}\|x_2-x_1\|^2,$$

$$(4.4)\qquad g^i(x_2)-g^i(x_1)-\langle g_x{}^i(x_1), x_2-x_1\rangle\leqslant\frac{l}{2}\|x_2-x_1\|^2,$$

верные для любых $x_1\in E_n$, $x_2\in E_n$ и $i=1, 2, \ldots, m$.

Для произвольных $x_0\in E_n$ и $\alpha>0$ обозначим через $\{x_k(x_0, \alpha)\}$ последовательность, порождаемую процессом (4.1), через $R(x_0, \alpha)$ — множество ее предельных точек. В силу утверждения леммы 3, если $x_0\in Q$, то, какое $\alpha>0$ ни брать, вся последовательность $\{x_k(x_0, \alpha)\}$ никогда не выходит за пределы множества Q.

Теорема 2. *Пусть $f(x)$ — выпуклая, непрерывно дифференцируемая функция, удовлетворяющая неравенству* (4.3), *функции $g^i(x)$, $i=1, 2, \ldots$ $\ldots, m$, линейные, в задаче* (1.1) *выполнено условие Слейтера и условие положительной регулярности ограничений, множество Q компактно. Тогда* *существует константа* $\overline{\alpha}>0$ *такая, что* $R(x_0, \alpha)\in X_*$ *для любых* $x_0\in Q$ *и* $\alpha\geqslant\overline{\alpha}$.

Доказательство. Будем рассматривать только такие α, которые не меньше некоторого $\alpha_1>0$. Из необходимых условий экстремума (3.7), (3.8) следует, что $\langle H_u(x_k, u_k, \alpha), u-u_k\rangle\leqslant 0$ для всех $u\in E_m^+$. В частности, при $u=u_{k+1}$ получаем

$$\langle H_u(x_k, u_k, \alpha), u_{k+1}-u_k\rangle\leqslant 0. \tag{4.5}$$

Неравенство (4.5) можно переписать в виде

$$\begin{aligned} L(x_k, u_k)-\langle (M\alpha)_p(L_x(x_k, u_k), x_k), L_x(x_k, u_k)\rangle &\geqslant \\ \geqslant L(x_k, u_{k+1})-\langle (M\alpha)_p(L_x(x_k, u_k), x_k), L_x(x_k, u_{k+1})\rangle &= \\ =L(x_k, u_{k+1})+\langle L_x(x_k, u_{k+1}), x_{k+1}-x_k\rangle&. \end{aligned} \tag{4.6}$$

Оценим с помощью (4.3) правую часть в (4.6):

$$L(x_k, u_{k+1})+\langle L_x(x_k, u_{k+1}), x_{k+1}-x_k\rangle\geqslant L(x_{k+1}, u_{k+1})-\frac{l}{2}\|x_{k+1}-x_k\|^2.$$

Отсюда и из (4.6) получаем

$$\begin{aligned} L(x_k, u_k)\geqslant L(x_{k+1}, u_{k+1})&+ \\ +\langle (M\alpha)_p(L_x(x_k, u_k), x_k), L_x(x_k, u_k)\rangle&-\frac{l}{2}\|(M\alpha)_p(L_x(x_k, u_k), x_k)\|^2. \end{aligned} \tag{4.7}$$

Поскольку последовательность $\{x_k(x_0, \alpha)\}$ принадлежит компактному множеству Q, то, в силу утверждения леммы 4, соответствующая последовательность $\{u_k(x_0, \alpha)\}$ ограничена. Поэтому существует компактное множество $P_1\subset E_n$ такое, что на каждом шаге $\alpha^{-1}L_x(x_k, u_k)\in P_1$. Но тогда, согласно условию Б, можно указать константу $\gamma=\gamma(P_1, Q)>0$ такую, что $\langle M_p(\alpha^{-1}L_x(x_k, u_k), x_k), \alpha^{-1}L_x(x_k, u_k)\rangle\geqslant\gamma\|M_p(\alpha^{-1}L_x(x_k, u_k), x_k)\|^2$.
Отсюда и из (3.6), (4.7) видно, что если положить $\overline{\alpha}=\max\{\alpha_1, l/\gamma\}$, то для всех $\alpha\geqslant\overline{\alpha}$

$$L(x_k, u_k)\geqslant L(x_{k+1}, u_{k+1})+\frac{\alpha\gamma}{2}\|(M\alpha)_p(L_x(x_k, u_k), x_k)\|^2,$$

т. е. последовательность $\{L(x_k, u_k)\}$ монотонно убывает.

Выберем теперь $\alpha\geqslant\overline{\alpha}$ и зафиксируем его. На основании следствия из леммы 4, последовательность $\{u_k(x_0, \alpha)\}$ принадлежит компактному множеству W. Непрерывная функция $L(x, u)$ ограничена снизу на компактном множестве $Q\times W$, поэтому

$$\lim_{k\to\infty}\|(M\alpha)_p(L_x(x_k, u_k), x_k)\|=0. \tag{4.8}$$

Таким образом, если $x_*\in R(x_0, \alpha)$, то $(M\alpha)_p(L_x(x_*, u_*), x_*)=0$, где u_* — предельная точка последовательности $\{u_k(x_0, \alpha)\}$. В силу утверждения леммы 5, это возможно только в случае, когда $x_*\in X_*$. Теорема доказана.

Если ограничиться рассмотрением функций $M(p, q)$ вида (2.3), причем в качестве $F(p)$ брать

$$F(p)=h(\|p\|), \tag{4.9}$$

где h — монотонно возрастающая, непрерывно дифференцируемая выпуклая функция на $[0, \infty)$, принимающая в нуле конечное значение и удовлетворяющая на любом конечном интервале $(0, r]$ неравенству

$$h'(t)/t\leqslant\gamma, \quad \gamma=\gamma(r)>0, \tag{4.10}$$

то получаем другое утверждение о сходимости процесса (4.1).

Теорема 3. *Пусть $f(x)$, $g^i(x)$, $i=1, 2, \ldots, m$,— выпуклые, непрерывно дифференцируемые функции, удовлетворяющие неравенствам* (4.3), (4.4), *в задаче* (1.1) *выполнено условие Слейтера и условие положительной регулярности ограничений, функция $M(p, q)$ строится в виде* (2.3)', (4.9). *Тогда для любого $x_0 \in Q$ можно указать $\bar{\alpha}>0$ такое, что $R(x_0, \alpha) \subseteq X_*$ для всех $\alpha \geq \bar{\alpha}$.*

Доказательство аналогично доказательству теоремы 3.2 из [6].

Возьмем $x_0 \in Q$. Так как в задаче (1.1) выполнено условие Слейтера, то существует точка Куна — Таккера $[x_*, u_*]$. Но тогда из $L_x(x_*, u_*) \in K^*(Q, x_*)$ следует неравенство

$$\langle L_x(x_*, u_*), x-x_* \rangle \geq 0, \tag{4.11}$$

справедливое для всех $x \in Q$. Кроме того, в силу теоремы 23.5 из [7],

$$\alpha^{-1} L_x(x_k, u_k) = M_p^*(x_k - x_{k+1}, x_k) = F_p(x_k - x_{k+1}),$$

где F_p — градиент функции F по p. Поэтому

$$\langle L_x(x_k, u_k) - \alpha F_p(x_k - x_{k+1}), x_* - x_{k+1} \rangle = 0. \tag{4.12}$$

Положив $x = x_{k+1}$ в (4.11) и сложив неравенство (4.11) с равенством (4.12), получим

$$\alpha \langle F_p(x_k - x_{k+1}), x_{k+1} - x_* \rangle \geq \langle L_x(x_k, u_k) - L_x(x_*, u_*), x_{k+1} - x_* \rangle. \tag{4.13}$$

Обозначим правую часть этого неравенства через Δ_k. Ее можно представить в виде $\Delta_k = \Delta_k^{(1)} + \Delta_k^{(2)} + \Delta_k^{(3)}$, где $\Delta_k^{(1)} = \langle g_x(x_k) u_k, x_{k+1} - x_* \rangle$, $\Delta_k^{(2)} = \langle g_x(x_*) u_*, x_* - x_{k+1} \rangle$, $\Delta_k^{(3)} = \langle f_x(x_k) - f_x(x_*), x_{k+1} - x_* \rangle$.

Оценим последовательно все $\Delta_k^{(i)}$, $i=1, 2, 3$. Из необходимых условий (3.7), (3.8) следует

$$g(x_k) + g_x^{\mathrm{T}}(x_k)(x_{k+1} - x_k) \leq 0, \tag{4.14}$$

$$\langle u_k, g(x_k) + g_x^{\mathrm{T}}(x_k)(x_{k+1} - x_k) \rangle = 0. \tag{4.15}$$

Из (4.15) и (1.4), учитывая выпуклость функций $g^i(x)$, получаем

$$\Delta_k^{(1)} = \langle g_x(x_k) u_k, x_{k+1} - x_k + x_k - x_* \rangle \geq -\langle g(x_*), u_k \rangle \geq 0.$$

Далее, согласно (4.4) и (4.14), имеет место неравенство

$$g^i(x_{k+1}) \leq \frac{l}{2} \|x_{k+1} - x_k\|^2, \qquad i=1, 2, \ldots, m.$$

Отсюда и из (1.4) приходим к

$$\Delta_k^{(2)} = \langle g_x(x_*) u_*, x_* - x_{k+1} \rangle \geq \langle g(x_*), u_* \rangle - \langle g(x_{k+1}), u_* \rangle \geq$$

$$\geq -\frac{l}{2} \|x_{k+1} - x_k\|^2 \sum_{i=1} u_*^i.$$

Наконец, в [6] получена оценка, вытекающая из (4.3), а именно $\Delta_k^{(3)} \geq -l \|x_{k+1} - x_k\|^2/4$. Таким образом,

$$\Delta_k \geq -\frac{l}{2} \left[\sum_{i=1}^{m} u_*^i + \frac{1}{2} \right] \|x_{k+1} - x_k\|^2. \tag{4.16}$$

С другой стороны, из вида функции $F(p)$ следует, что она непрерывно дифференцируема всюду на E_n, за исключением, быть может, нуля, и

$$F_p(p) = \frac{h'(\|p\|)}{\|p\|} p.$$

Но $2\langle x_k - x_{k+1}, x_{k+1} - x_* \rangle = \|x_k - x_*\|^2 - \|x_{k+1} - x_*\|^2 - \|x_{k+1} - x_k\|^2$. Поэтому с учетом (4.10), (4.13), (4.16) получаем

$$\|x_{k+1} - x_*\|^2 \leqslant \|x_k - x_*\|^2 - \left[1 - \frac{l}{8\alpha\gamma}\left(1 + 2\sum_{i=1}^{m} u_*^i\right)\right] \|x_{k+1} - x_k\|^2,$$

где $\gamma = \gamma(\|x_0 - x_*\|)$. Отсюда видно, что если взять

$$\bar{\alpha} = l\left(1 + 2\sum_{i=1}^{m} u_*^i\right)(4\gamma)^{-1},$$

то $\|x_{k+1} - x_*\|^2 \leqslant \|x_k - x_*\|^2 - \|x_{k+1} - x_k\|^2/2$ при всех $\alpha \geqslant \bar{\alpha}$. Таким образом, последовательность $\{x_k\}$ принадлежит ограниченному множеству и выполнено (4.8). Дальнейшие рассуждения те же, что и при доказательстве теоремы 2. Теорема доказана.

Обратим внимание на одну особенность процесса (4.1). С этой целью введем функцию

$$\bar{L}_k(x, u) = L(x_k, u) + \langle L_x(x_k, u), x - x_k \rangle,$$

являющуюся линейным приближением по x функции Лагранжа в окрестности точки x_k. На основании (3.7), (3.8),

(4.17) $\qquad \bar{L}_k(x_{k+1}, u_k) \geqslant \bar{L}_k(x_{k+1}, u) \qquad \forall u \in E_m^+.$

Таким образом, метод (4.1) обладает тем свойством, что он на каждом шаге переводит текущую точку x_k в точку x_{k+1}, в которой линейная функция $\bar{L}_k(x_{k+1}, u)$ имеет максимум по u на E_m^+ и этот максимум достигается в точке u_k. Если (1.1) — задача линейного программирования, то неравенство (4.17) выполняется для самой функции Лагранжа, т. е.

(4.18) $\qquad L(x_{k+1}, u_k) \geqslant L(x_{k+1}, u) \qquad \forall u \in E_m^+.$

Рассмотрим задачу выпуклого программирования

(4.19а) $\qquad \min\limits_{y} \langle f_x(x_k), y \rangle + \alpha M^*(-y, x_k),$

(4.19б) $\qquad g^i(x_k) + \langle g_x^i(x_k), y \rangle \leqslant 0, \qquad i = 1, 2, \ldots, m.$

Из сравнения необходимых и достаточных условий для задач (3.1) и (4.19) следует, что любое направление сдвига $-(M\alpha)_p(L_x(x_k, u_k), x_k)$ в методе (4.1) является решением задачи (4.19). Если функция $M(p, q)$ симметрична относительно p, то симметричной по p будет и $M^*(p, q)$. Знак минус в выражении для целевой функции в задаче (4.19) тогда можно опустить. В частном случае, когда $M(p, q)$ строится в виде (2.3), причем в качестве $F(p)$ берется $F(p) = \|p\|^2$, задача (4.19) становится близкой к вспомогательной, используемой в методе линеаризации [9].

§ 5. Задача линейного программирования

Рассмотрим задачу линейного программирования

(5.1) $\qquad \min\limits_{x} \langle c, x \rangle, \qquad Ax \leqslant d, \qquad x \geqslant 0,$

где $c \in E_n$, $d \in E_m$, A — матрица размера $m \times n$. Двойственной к ней будет задача

$$(5.2)\qquad \max(-\langle d, u\rangle), \quad A^{\mathrm{T}}u\geqslant -c, \quad u\geqslant 0.$$

Будем считать, что решения задач (5.1) и (5.2) существуют.

Применим для решения (5.1) метод (4.1). Предположение о линейности всех функций, определяющих задачу (5.1), позволяет получить значительно более сильное утверждение о сходимости процесса (4.1) по сравнению с теоремами 2 и 3. Прежде всего отметим, что, какая бы начальная точка $x_0\in E_n^+$ ни была взята, все последующие точки последовательности $\{x_k\}$ оказываются допустимыми. Это следует из необходимого условия (3.7), если подставить в него $x=x_k$, $u=u_k$ и учесть (4.1).

Теорема 4. *Пусть в задаче* (5.1) *выполнено условие Слейтера и условие положительной регулярности ограничений. Тогда для любого* $\alpha>0$ *метод* (4.1) *сходится к решению задачи линейного программирования* (5.1) *за конечное число шагов.*

Доказательство. Неравенство (4.7) для задачи (5.1) переписывается в виде

$$(5.3)\qquad L(x_{k+1}, u_{k+1})\leqslant L(x_k, u_k)-\langle L_x(x_k, u_k), (M\alpha)_p(L_x(x_k, u_k), x_k)\rangle.$$

Далее, в силу выпуклости функции $(M\alpha)(p, q)$ по p, имеет место неравенство

$$L(x_k, u_k)\geqslant L(x_k, u_k)-\langle L_x(x_k, u_k), (M\alpha)_p(L_x(x_k, u_k), x_k)\rangle.$$

Правая часть в этом неравенстве равна $L(x_{k+1}, u_k)$. Из (4.18) следует, что $L(x_{k+1}, u_k)\geqslant L(x_{k+1}, u_*)$, где u_* — оптимальное решение в задаче (5.2). Так как x_{k+1} — допустимая точка в прямой задаче (5.1), u_* — допустимая точка в двойственной задаче (5.2), то

$$L(x_{k+1}, u_*)=\langle c+A^{\mathrm{T}}u_*, x_{k+1}\rangle-\langle d, u_*\rangle\geqslant-\langle d, u_*\rangle=\langle c, x_*\rangle.$$

Таким образом, последовательность $\{L(x_k, u_k)\}$ монотонно убывает и ограничена снизу. Поэтому, согласно (5.3),

$$\lim_{k\to\infty}\langle L_x(x_k, u_k), (M\alpha)_p(L_x(x_k, u_k), x_k)\rangle=0.$$

Последнее возможно только в случае, когда

$$\lim_{k\to\infty} L_x(x_k, u_k)=0 \text{ или } \lim_{k\to\infty} A^{\mathrm{T}}u_k=-c.$$

Составим пару двойственных задач:

$$(5.4\text{а})\qquad \min_x(-\langle A^{\mathrm{T}}u_k, x\rangle), \quad Ax\leqslant d, \quad x\geqslant 0,$$

$$(5.4\text{б})\qquad \max_u(-\langle d, u\rangle), \quad A^{\mathrm{T}}u\geqslant A^{\mathrm{T}}u_k, \quad u\geqslant 0.$$

Точки x_{k+1} и u_k, как видно из (3.7), (3.8), будут их решениями. Но вектор $-A^{\mathrm{T}}u_k$ стремится к c, поэтому, согласно лемме 4 из [10], начиная с некоторого k_0 решения прямых задач (5.4а) должны совпадать с решением задачи (5.1). Теорема доказана.

Утверждение теоремы остается в силе, если вместо неотрицательного ортанта E_n^+ в качестве Q взять все пространство E_n или параллелепипедное множество (2.2).

Литература

1. *Пшеничный Б. Н.* Выпуклый анализ и экстремальные задачи. М.: Наука, 1980.
2. *Третьяков Н. В.* Метод штрафных оценок для задач выпуклого программирования.— Экономика и матем. методы, 1973, т. 9, № 3, с. 526—540.
3. *Гольштейн Е. Г., Третьяков Н. В.* Слабые модифицированные функции Лагранжа.— Экономика и матем. методы, 1979, т. 15, № 6, с. 1180—1193.
4. *Евтушенко Ю. Г.* Применение модифицированных функций Лагранжа для решения задач нелинейного программирования.— В кн.: Иссл. операций. Вып. 7. М.: ВЦ АН СССР, 1979, с. 3—24.
5. *Голиков А. И., Жадан В. Г.* Итеративные методы решения задач нелинейного программирования с использованием модифицированных функций Лагранжа.— Ж. вычисл. матем. и матем. физ., 1980, т. 20, № 4, с. 874—888.
6. *Антипин А. С.* Методы нелинейного программирования, основанные на прямой и двойственной модификации функции Лагранжа.— Препринт ВНИИ системных иссл. М., 1979.
7. *Рокафеллар Р.* Выпуклый анализ. М.: Мир, 1973.
8. *Еремин И. И., Астафьев Н. Н.* Введение в теорию линейного и выпуклого программирования. М.: Наука, 1976.
9. *Пшеничный Б. Н., Данилин Ю. М.* Численные методы в экстремальных задачах. М.: Наука, 1975.
10. *Поляк Б. Т., Третьяков Н. В.* Об одном итеративном методе линейного программирования и его экономической интерпретации.— Экономика и матем. методы, 1972, т. 8, № 5, с. 740—751.

Поступила в редакцию 24.VI.1982
Переработанный вариант 25.VIII.1982