ЭКОНОМИКА
И МАТЕМАТИЧЕСКИЕ МЕТОДЫ

1965
том I, вып. 3

ДВОЙСТВЕННЫЕ ЗАДАЧИ ВЫПУКЛОГО ПРОГРАММИРОВАНИЯ

Е. Г. ГОЛЬШТЕЙН
(Москва)

1. Известно, какое важное значение в линейном и нелинейном программировании имеет принцип двойственности и связанное с ним понятие разрешающих множителей (множителей Лагранжа). В данной работе приводится общая схема формирования задачи, двойственной по отношению к произвольной задаче выпуклого программирования, и устанавливаются условия, необходимые и достаточные для справедливости основной теоремы двойственности.

В отличие от геометрического принципа построения двойственных задач, предложенного в [1] (см. также [2], гл. 7), излагаемая ниже схема имеет аналитический характер и тесно связана с функцией Лагранжа. Используя эту схему, легко, например, получить известные выражения для двойственных задач применительно к линейному, квадратичному и выпуклому программированию с линейными ограничениями.

2. Рассмотрим задачу нелинейного программирования, состоящую в отыскании

$$v = \sup f(x) \tag{1}$$

при условиях

$$g(x) \geqslant 0, \tag{2}$$

$$x \in R. \tag{3}$$

Здесь $g(x) = (g_1(x), g_2(x), \ldots, g_m(x))$, $x = (x_1, x_2, \ldots, x_n)$, $f(x)$ и $g_i(x)$ — функции, определенные на множестве R. Точку x, удовлетворяющую условиям (2) — (3), назовем планом задачи (1) — (3). В дальнейшем будем иногда обозначать множество планов задачи через G. Последовательность $\{x^{(k)}\}$ планов $x^{(k)}$ называется *решением* задачи (1) — (3), если

$$\lim_{k\to\infty} f(x^{(k)}) = v,$$

где v — верхняя грань (1) при условиях (2) — (3).

В частности, решением задачи (1) — (3) может оказаться стационарная последовательность $\{x^{(k)}\} = \{x^*\}$, т. е. $f(x^*) = v$. В этом случае будем говорить, что решением задачи является план x^*, отвечающий данной стационарной последовательности.

Составим для задачи (1) — (3) функцию Лагранжа

$$F(x, \lambda) = f(x) + (\lambda, g(x)), \tag{4}$$

где $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m)$.

Пусть

$$\varphi(x) = \inf_{\lambda \geqslant 0} F(x, \lambda). \tag{5}$$

Введем в рассмотрение задачу, состоящую в отыскании

$$v' = \sup \varphi(x) \tag{6}$$

при условии

$$x \in R. \tag{7}$$

Распространим на задачу (5)—(7) введенные выше понятия плана и решения.

Лемма 1. *Любое решение задачи* (1)—(3) *является решением задачи* (5)—(7). *Если* $v' > -\infty$, *то любое решение задачи* (5)—(7) — *решение задачи* (1)—(3). *Если* $v' = -\infty$, *то система условий* (2)—(3) *противоречива.*

Доказательство. Из соотношений (4), (5), определяющих функцию $\varphi(x)$, непосредственно следует, что при $x \in R$

$$\varphi(x) = \begin{cases} f(x), & \text{если } x \in G, \\ -\infty, & \text{если } x \overline{\in} G \end{cases} \tag{8}$$

(G — множество планов задачи (1)—(3)).

Равенство (8) убеждает нас в наличии двух альтернатив: 1) множество планов G непусто и $v = v' > -\infty$; 2) множество G пусто и $v' = -\infty$.

Пусть $\{x^{(k)}\}$ — решение задачи (1)—(3). Согласно (8)

$$\lim_{k\to\infty} f(x^{(k)}) = \lim_{k\to\infty} \varphi(x^{(k)}) = v.$$

Поскольку имеет место альтернатива 1, то

$$\lim_{k\to\infty} \varphi(x^{(k)}) = v = v',$$

т. е. $\{x^{(k)}\}$ — решение задачи (5)—(7).

Пусть теперь $\{x^{(k)}\}$ — решение задачи (5)—(7), причем $v' > -\infty$. В силу (8) точки $x^{(k)} \in G$ при достаточно больших k.

Следовательно, начиная с некоторого k_0 последовательность $\{x^{(k)}\}$ состоит из планов задачи (1)—(3), причем

$$\lim_{k\to\infty} f(x^{(k)}) = \lim_{k\to\infty} \varphi(x^{(k)}) = v'.$$

Учитывая далее наличие альтернативы 1, имеем

$$\lim_{k\to\infty} f(x^{(k)}) = v' = v,$$

т. е. $\{x^{(k)}\}_{k \geqslant k_0}$ — решение задачи (1)—(3).

Наконец, при $v' = -\infty$ имеет место альтернатива 2, и, следовательно, множество планов G пусто. Лемма доказана.

Согласно лемме 1 анализ задачи (1)—(3) эквивалентен исследованию задачи (5)—(7). При этом следует подчеркнуть, что тождественность задач (1)—(3) и (5)—(7) доказана без каких бы то ни было предположений относительно функций $f(x)$, $g_i(x)$ и множества R.

Структура задачи (5)—(7), эквивалентной исходной задаче (1)—(3), подсказывает естественную формулировку для двойственной задачи.

Задача (5)—(7) заключается в том, что к функции Лагранжа $F(x, \lambda)$ вначале применяется операция «inf» по переменным $\lambda \geqslant 0$, а затем операция «sup» по переменным $x \in R$. Для того чтобы перейти к двойственной задаче, изменим порядок применения этих операций.

Пусть

$$\psi(\lambda) = \sup_{x \in R} F(x, \lambda). \tag{9}$$

Определим задачу, двойственную по отношению к задаче (1)—(3), следующим образом: требуется найти

$$\tilde{v} = \inf \psi(\lambda) \tag{10}$$

при

$$\lambda \geqslant 0. \tag{11}$$

3. Прежде чем формулировать теорему двойственности, расширим понятие плана задачи (1)—(3). Именно назовем последовательность $\{x^{(k)}\}$ точек $x^{(k)} \in R$ обобщенным планом задачи (1)—(3), если соблюдаются следующие условия:

$$\begin{aligned} &g_i(x^{(k)}) \geqslant -\varepsilon_k, \qquad && i = 1, 2, \ldots, m, \\ &\varepsilon_k \geqslant 0, && \lim_{k\to\infty} \varepsilon_k = 0, \end{aligned} \tag{12}$$

причем последовательность $f(x^{(k)})$ имеет конечный или бесконечный предел.

Очевидно, план x задачи (1)—(3) соответствует обобщенному плану, определяемому стационарной последовательностью $\{x^{(k)}\} = \{x\}$.

Если $X = \{x^k\}$ — обобщенный план задачи (1)—(3), то положим

$$f(X) = \lim_{k\to\infty} f(x^{(k)}). \tag{13}$$

Пусть $\overline{G}$ — совокупность обобщенных планов задачи (1)—(3). Согласно замечанию, сделанному выше, $G \subset \overline{G}$.

Расширим постановку задачи (1)—(3), потребовав, чтобы верхняя грань (1) бралась по множеству $\overline{G}$ обобщенных планов этой задачи. Значение верхней грани обозначим через v''. Очевидно,

$$v'' \geqslant v = v'. \tag{14}$$

Доопределим значение верхней грани (1) на множестве $\overline{G}$, положив $v'' = -\infty$, если $\overline{G}$ — пустое множество.

Вновь введенную задачу назовем обобщенной задачей (1)—(3).

Легко проверить, что среди обобщенных планов задачи (1)—(3) (если, конечно, множество их непусто) найдется такой обобщенный план X^*, что

$$f(X^*) = v''. \tag{15}$$

Действительно, если

$$\lim_{k\to\infty} f(X^{(k)}) = v'',$$

$X^{(k)} = \{x^{(k,t)}\}_t$, то в качестве X^* можно принять последовательность $\{\tilde{x}^{(k)}\}$, где $\tilde{x}^{(k)} = x^{(k,t_k)}$, а последовательность индексов t_k достаточно быстро стремится к бесконечности. Обобщенный план, для которого выполняется равенство (15), назовем решением обобщенной задачи (1)—(3).

Теорема 1. *Если $f(x)$ и $g_i(x)$ — выпуклые вверх функции, определенные на выпуклом множестве R, то*

$$v'' = \tilde{v}, \tag{16}$$

т. е. значение верхней грани в обобщенной исходной задаче (1)—(3) *равно значению нижней грани в двойственной задаче* (9)—(11).

Предварительно докажем вспомогательную лемму.

Лемма 2. *Справедливо соотношение*

$$v'' \leqslant \tilde{v}. \tag{17}$$

Доказательство леммы. Если $\overline{G}$ — пустое множество, то $v'' = -\infty$, и неравенство (17) очевидно. Пусть $\overline{G}$ не пусто и $\overline{X} = \{\overline{x}^{(k)}\}$ — решение обобщенной задачи (1) — (3).

Имеем для $\lambda \geqslant 0$.

$$\sup_{x \in R} F(x, \lambda) \geqslant F(\overline{x}^{(k)}, \lambda) = f(\overline{x}^{(k)}) + (\lambda, g(\overline{x}^{(k)})) \geqslant f(x^{(k)}) - \varepsilon_k \sum_{i=1} \lambda_i. \quad (18)$$

Переходя в неравенстве (18) к пределу при $k \to \infty$ и учитывая, что $\lim_{k \to \infty} \varepsilon_k = 0$, получаем

$$\psi(\lambda) = \sup_{x \in R} F(x, \lambda) \geqslant v'' \quad (19)$$

при любых $\lambda \geqslant 0$. Следовательно,

$$\widetilde{v} = \inf_{\lambda \geqslant 0} \psi(\lambda) \geqslant v'',$$

что и требовалось доказать.

Доказательство теоремы. Приводимые ниже рассуждения опираются на известную схему, применяемую обычно для доказательства теоремы Куна — Таккера (см., например, [2]).

1. Пусть вначале v'' — конечное число. Введем множество $K(\varepsilon)$, $\varepsilon > 0$, состоящее из $m+1$-мерных векторов $w = (y_0, y) = (y_0, y_1, \ldots, y_m)$ при помощи следующего условия: $w \in K(\varepsilon)$ в том и только в том случае, если при некотором $x \in R$ имеет место система неравенств

$$y_0 \leqslant f(x) + \varepsilon, \qquad y_i \leqslant g_i(x) + \varepsilon, \quad i = 1, 2, \ldots, m. \quad (20)$$

Из выпуклости вверх функций $f(x)$, $g_i(x)$ и выпуклости множества R вытекает выпуклость множества $K(\varepsilon)$.

В том же $m+1$-мерном пространстве введем луч $l(\delta)$, $\delta \geqslant 0$, состоящий из точек w, для которых $y_0 \geqslant v'' + \delta$, $y_i = 0$, $i = 1, 2, \ldots, m$.

Проверим, что при фиксированном δ и достаточно малом $\varepsilon > 0$ множества $K(\varepsilon)$ и $l(\delta)$ не имеют общих точек.

Действительно, предположение противного ведет к существованию точек $x^{(k)} \in R$ и чисел $\varepsilon_k > 0$ таких, что

$$v'' + \delta \leqslant f(x^{(k)}) + \varepsilon_k, \qquad g_i(x^{(k)}) \geqslant -\varepsilon_k, \qquad \lim_{k \to \infty} \varepsilon_k = 0.$$

Выделяя из последовательности $\{x^{(k)}\}$ подпоследовательность $\{\overline{x}^{(t)}\} = \{x^{(k_t)}\}$, для которой $f(\overline{x}^{(t)})$ имеет предел, приходим к такому обобщенному плану $\overline{X} = \{\overline{x}^{(t)}\}$, что

$$f(\overline{X}) \geqslant v'' + \delta, \qquad \delta > 0.$$

Полученное неравенство противоречит определению числа v'' и служит доказательством нашего утверждения.

Пусть теперь $\varepsilon = \varepsilon(\delta) > 0$ выбрано таким, что $K(\varepsilon) \cap l(\delta)$ — пустое множество.

Применим к непересекающимся выпуклым множествам $K(\varepsilon)$ и $l(\delta)$ теорему о разделяющей гиперплоскости. Согласно этой теореме найдется такой ненулевой вектор

$$(\lambda_0^{(\varepsilon)}, \lambda^{(\varepsilon)}) = (\lambda_0^{(\varepsilon)}, \lambda_1^{(\varepsilon)}, \lambda_2^{(\varepsilon)}, \ldots, \lambda_m^{(\varepsilon)}),$$

что

$$\lambda_0^{(\varepsilon)} y_0 + (\lambda^{(\varepsilon)}, y) \leqslant \lambda_0^{(\varepsilon)} (v'' + \delta) \tag{21}$$

для любой точки $(y_0, y) \in K(\varepsilon)$.

В соответствии с условиями (20) множество $K(\varepsilon)$ содержит точки w со сколь угодно малыми значениями компонент. Следовательно, неравенство (21) может соблюдаться лишь при

$$\lambda_i^{(\varepsilon)} \geqslant 0, \qquad i = 0, 1, \ldots, m. \tag{22}$$

Покажем, что $\lambda_0^{(\varepsilon)} > 0$. Пусть $\{x^{(k)}\}$ — некоторый обобщенный план. Определим компоненты точки $w^{(k)} \in K(\varepsilon)$ следующим образом:

$$y_0^{(k)} = f(x^{(k)}), \qquad y_i^{(k)} = g_i(x^{(k)}) + \varepsilon, \qquad i = 1, 2, \ldots, m.$$

Если допустить, что $\lambda_0^{(\varepsilon)} = 0$, то приходим к соотношению

$$(\lambda^{(\varepsilon)}, y^{(k)}) \leqslant 0, \tag{23}$$

верному для любого натурального k.

Подставляя в (23) выражения для $y^{(k)}$ и пользуясь тем, что

$$\lim_{k\to\infty} g(x^{(k)}) \geqslant 0,$$

получаем неравенство

$$\varepsilon \sum_{i=1}^{m} \lambda_i^{(\varepsilon)} \leqslant 0,$$

из которого с учетом (22) имеем $\lambda_i^{(\varepsilon)} = 0$ для $i = 1, 2, \ldots, m$.

Получили противоречие, поскольку вектор $(\lambda_0^{(\varepsilon)}, \lambda^{(\varepsilon)})$ по условию не равен нулевому. Следовательно,

$$\lambda_0^{(\varepsilon)} > 0. \tag{24}$$

Используя (24), положим $\tilde{\lambda}^{(\varepsilon)} = \lambda^{(\varepsilon)} / \lambda_0^{(\varepsilon)}$.

Кроме того, пусть $y_0 = f(x)$, $y = g(x)$.

Из неравенства (21) получаем

$$f(x) + (\tilde{\lambda}^{(\varepsilon)}, g(x)) \leqslant v'' + \delta$$

при любом $x \in R$.

Следовательно, $\tilde{v} \leqslant \psi(\tilde{\lambda}^{(\varepsilon)}) \leqslant v'' + \delta$.

Учитывая далее, что δ — произвольное положительное число, приходим к окончательному соотношению $\tilde{v} \leqslant v''$, которое вместе с неравенством (17) убеждает нас в справедливости равенства $\tilde{v} = v''$.

2. Предположим теперь, что $v'' = -\infty$. Пусть l_α — луч

$$y_0 \geqslant \alpha, \qquad y_i = 0, \qquad i = 1, 2, \ldots, m.$$

Рассуждения, аналогичные использованным в первой части теоремы, показывают, что при любом $\alpha > -\infty$ множества $K(\varepsilon)$ и l_α не имеют общих точек, если $\varepsilon > 0$ достаточно мало. Поэтому к этим множествам может быть применена теорема о разделяющей гиперплоскости, которая приводит к соотношению

$$\tilde{v} \leqslant \psi(\tilde{\lambda}^{(\varepsilon)}) \leqslant \alpha.$$

В силу произвольности числа α

$$\tilde{v} = -\infty.$$

Если, наконец, $v'' = \infty$, то согласно лемме 2

$$\tilde{v} \geqslant v'' = \infty.$$

Теорема полностью доказана.

4. В качестве очевидного следствия из теоремы 1 получаем следующее утверждение.

Теорема 2. *Пусть $f(x)$ и $g_i(x)$ — выпуклые вверх функции, определенные на выпуклом множестве R. Задачи* (1) — (3) *и* (9) — (11) *связаны соотношением двойственности*

$$v = \tilde{v} \tag{25}$$

в том и только в том случае, если переход от исходной задачи (1) — (3) *к обобщенной исходной задаче не ведет к возрастанию верхней грани* (1), *т. е.*

$$v = v''. \tag{26}$$

Теорема 2 сводит проверку справедливости соотношения двойственности для задач (1) — (3) и (9) — (11) к установлению равенства (26).

Приведем условия, достаточные для выполнения этого равенства.

Лемма 3. *Если в дополнение к условиям теоремы 2 функции $f(x)$ и $g_i(x)$ непрерывны, множество R замкнуто, а множество G планов задачи* (1) — (4) *непусто и ограничено, то равенство* (26) *имеет место.*

Доказательство. Пусть $X = \{x^{(k)}\}$ — произвольный обобщенный план задачи (1) — (3), т. е. последовательность точек $x^{(k)} \in R$, удовлетворяющая условиям

$$g_i(x^{(k)}) \geqslant -\varepsilon_k, \qquad i = 1, 2, \ldots, m, \tag{27}$$

где

$$\varepsilon_k > 0, \qquad k = 1, 2, \ldots, \qquad \lim \varepsilon_k = 0. \tag{28}$$

Покажем, что элементы $x^{(k)}$ последовательности X ограничены в совокупности.

Действительно, допущение противного означает существование подпоследовательности $\{x^{(k_\alpha)}\} = \{\bar{x}^{(\alpha)}\}$, для которой

$$\lim_{\alpha\to\infty} |\bar{x}^{(\alpha)}| = \infty. \tag{29}$$

Положим

$$z^{(\alpha)} = \frac{\bar{x}^{(\alpha)} - x'}{|\bar{x}^{(\alpha)}|}, \qquad x' \in G.$$

Выделим из ограниченной последовательности $\{z^{(\alpha)}\}$ сходящуюся подпоследовательность $\{z^{(\alpha_\gamma)}\} = \{\bar{z}^{(\gamma)}\}$:

$$\lim_{\gamma\to\infty} \bar{z}^{(\gamma)} = x^{(0)}.$$

Проверим, что

$$x' + \lambda x^{(0)} \in G \text{ при любом } \lambda \geqslant 0. \tag{30}$$

В силу выпуклости R

$$x' + \lambda\bar{z}^{(\gamma)} \in R \text{ при } 0 \leqslant \lambda \leqslant |\bar{x}^{\alpha_\gamma}|. \tag{31}$$

Далее, при тех же значениях λ из выпуклости вверх функций $g_i(x)$ и условий (27) вытекают соотношения

$$g_i(x' + \lambda \bar{z}^{(\gamma)}) \geqslant -\bar{\varepsilon}_\gamma, \qquad i = 1, 2, \ldots, m, \tag{32}$$

$$\lim_{\gamma\to\infty} \bar{\varepsilon}_\gamma = 0. \tag{33}$$

Переходя в условиях (31) и (32) к пределу при $\gamma \to \infty$ и фиксированном λ, получаем с учетом (29) и (33)

$$x' + \lambda x^{(0)} \in R, \qquad g_i(x' + \lambda x^{(0)}) \geqslant 0, \qquad 0 \leqslant \lambda < \infty$$

(здесь использованы замкнутость R и непрерывность функций $g_i(x)$), т. е. соотношение (30) действительно имеет место. Но согласно условию G — ограниченное множество и поэтому не может содержать луч. Следовательно, элементы последовательности X должны быть ограничены в совокупности.

Выделим из ограниченной последовательности $X = \{x^{(k)}\}$ подпоследовательность, сходящуюся к $x'' \in R$ (здесь использована замкнутость R). Соотношения (27) и (28) в силу непрерывности функций $g_i(x)$ приводят к неравенствам

$$g_i(x'') \geqslant 0, \qquad i = 1, 2, \ldots, m.$$

Следовательно, $x'' \in G$.

По определению

$$f(X) = \lim_{k\to\infty} f(x^{(k)}).$$

В соответствии с непрерывностью функции $f(x)$

$$f(X) = \lim_{k\to\infty} f(x^{(k)}) = f(x'').$$

Итак, для любого обобщенного плана X

$$f(X) = f(x'') \leqslant v,$$

т. е. $v'' \leqslant v$.

Учитывая, далее, очевидное неравенство $v'' \geqslant v$, получаем искомое равенство (26). Лемма доказана.

Объединим результаты теоремы 2 и леммы 3.

Теорема 3. *Если функции $f(x)$, $g_i(x)$ выпуклы вверх и непрерывны на замкнутом выпуклом множестве R, условия (2)—(3) высекают непустое ограниченное множество, то задачи (1)—(3) и (9)—(11) связаны соотношением двойственности $v = \tilde{v}$.*

Заметим, что условия теоремы 3 при $v < \infty$, очевидно, гарантируют достижимость верхней грани (1) на множестве G. Другими словами, среди решений задачи (1)—(3) обязательно имеется некоторый план. Однако двойственная задача (9)—(11) может и не обладать этим свойством. Приведем простейший пример:

$$f(x) = x - \max, \qquad g(x) = -x^2, \qquad R : x \geqslant 0.$$

Имеем $F(x, \lambda) = x - \lambda x^2$, $\psi(\lambda) = \sup\limits_{x\geqslant 0} (x - \lambda x^2) = 1/4\lambda$ $(\lambda > 0)$.

При этом $\inf\limits_{\lambda\geqslant 0} \psi(\lambda) = 0$ не достигается ни в одной точке $\lambda \geqslant 0$.

5. Приведем еще два условия, гарантирующих соблюдение соотношения двойственности (25).

Первое из них известно под названием условия Слейтера.

Будем говорить, что ограничения (2) — (3) удовлетворяют условию Слейтера, если существует план $x^{(0)}$ задачи (1) — (3), для которого

$$g_i(x^{(0)}) > 0, \qquad i = 1, 2, \ldots, m.$$

Система условий (2) задачи (1) — (3) может содержать линейные уравнения, каждое из которых, очевидно, может быть представлено в виде двух условий-неравенств:

$$-\sum_{j=1}^{n} a_{ij}x_j + b_i \geqslant 0, \qquad \sum_{j=1}^{n} a_{ij}x_j - b_i \geqslant 0. \tag{34}$$

Допустим, что условия (2) задачи (1) — (3) имеют вид

$$\begin{gathered} g_i(x) \geqslant 0, \qquad i = 1, 2, \ldots, m_1, \\ \sum_{j=1}^{n} a_{ij}x_j = b_i, \qquad i = 1, 2, \ldots, s \qquad (m_1 + 2s = m). \end{gathered} \tag{35}$$

Уравнения системы (35) предполагаются линейно независимыми.

Представляя каждое из уравнений системы (35) в виде (34), получаем следующую формулировку двойственной задачи. Пусть

$$F(x, \lambda) = f(x) + \sum_{i=1}^{m_1} \lambda_i g_i(x) + \sum_{i=1}^{s} \lambda_{m_1+i}\left(\sum_{j=1}^{n} a_{ij}x_j - b_i\right),$$

$$\psi(\lambda) = \sup_{x \in R} F(x, \lambda). \tag{36}$$

Требуется найти

$$\tilde{v} = \inf \psi(\lambda) \tag{37}$$

при

$$\lambda_i \geqslant 0, \qquad i = 1, 2, \ldots, m_1. \tag{38}$$

Применительно к задаче вида (1), (35), (3) условие Слейтера состоит в предположении существования такого плана $x^{(0)}$, что

$$g_i(x^{(0)}) > 0, \qquad i = 1, 2, \ldots, m_1;$$

$x^{(0)}$ — внутренняя точка множества R.

Дословно повторяя рассуждения, при помощи которых доказываются теоремы Куна — Таккера и Удзавы ([3], гл. III) об эквивалентности задачи выпуклого программирования и задачи об отыскании седловой точки функции Лагранжа, приходим к следующему утверждению.

Теорема 4. *Если функции $f(x)$, $g_i(x)$ выпуклы вверх, множество R выпукло и ограничения задачи* (1) — (3) *или* (1), (35), (3) *удовлетворяют условию Слейтера, то*

$$v = \tilde{v},$$

причем в случае $v < \infty$

$$\tilde{v} = \psi(\lambda^*)$$

для некоторого λ^*, удовлетворяющего условиям (11) *или* (38) *двойственной задачи.*

Итак, условие Слейтера не только гарантирует соблюдение равенства (25), но и обеспечивает при $v = \tilde{v} < \infty$ достижимость нижней грани (10) или (37) в двойственной задаче. Однако верхняя грань (1) при этом может и не достигаться.

Остановимся на частном классе задач типа (1)—(3), называемых обычно кусочно-линейными.

Пусть $f(x)$ и $g_i(x)$ — кусочно-линейные выпуклые вверх функции, т. е. функции, представимые в виде

$$g_i(x) = \min_{1 \leqslant k \leqslant N_i} \left\{ \sum_{j=1}^{n} d_{ij} x_j \right\}, \qquad i = 0, 1, 2, \ldots, m \qquad (g_0(x) \equiv f(x)).$$

Кроме того, примем, что R — выпуклое многогранное множество. Задачу (1)—(3), обладающую отмеченными свойствами, назовем кусочно-линейной.

Теорема 5. *Если* (1)—(3) *кусочно-линейная задача с непустым множеством планов, то задачи* (1)—(3) *и* (9)—(11) *связаны соотношением двойственности, причем в случае* $v = \tilde{v} < \infty$ *при некоторых* $x^* \in G$ *и* $\lambda^* \geqslant 0$

$$v = \tilde{v} = f(x^*) = \psi(\lambda^*).$$

Доказательство. Общая схема доказательства первой части утверждения теоремы похожа на доказательство теоремы 1. Если $v = \infty$, то утверждение теоремы — следствие леммы 2. Пусть $v < \infty$. Введем множества

$$K = K(0) \text{ и } l\colon\ y_0 \geqslant v;\ y_i = 0;\ i = 1, 2, \ldots, m.$$

Из условий (20), определяющих при $\varepsilon = 0$ множество K, кусочной линейности функций $f(x)$ и $g_i(x)$ и многогранности множества R вытекает, что K — выпуклое многогранное множество. В силу замкнутости K точка $w^* = (v, 0, 0, \ldots, 0) \in K$. Остальные точки луча l лежат вне K. Поэтому w^* — граничная точка K. Многогранное множество K — общая часть полупространств, порождаемых конечным числом гиперплоскостей. Среди части этих гиперплоскостей, проходящих через точку w^*, заведомо имеются гиперплоскости, не параллельные оси y_0 (в противном случае множество K содержало бы точки луча l, отличные от w^*). Любая такая гиперплоскость разделяет множества K и l, причем первая компонента λ_0 направляющего вектора (λ_0, λ) не равна нулю (она может быть сделана положительной). Рассуждения, аналогичные приведенным при доказательстве теоремы 1, показывают, что

$$\lambda^* = \lambda / \lambda_0 \geqslant 0$$

— решение задачи (9)—(11), причем

$$\tilde{v} = v = \psi(\lambda^*).$$

Для того чтобы завершить доказательство теоремы, нам осталось убедиться в достижимости верхней грани (1). Это вытекает из того, что кусочно-линейная задача (1)—(3) легко приводится к задаче линейного программирования, которая в случае ограниченности ее линейной формы на непустом множестве планов, как известно, имеет оптимальный план.

В дополнение к теореме 5 докажем следующее утверждение.

Теорема 6. *Пусть условия теоремы 5 изменены предположением о гладкости выпуклой вверх функции* $f(x)$. *Если верхняя грань* (1) *достигается на некотором плане, то задачи* (1)—(3) *и* (9)—(11) *связаны соотношением двойственности.*

Доказательство. Пусть x^* — оптимальный план задачи (1)—(3). Введем вспомогательную кусочно-линейную задачу, условия которой совпадают с (2)—(3), а геометрическим образом целевой функции является гиперплоскость $z = L(x)$, касательная к поверхности $z = f(x)$ в точке x^*.

Пусть v_1 и v_2 — верхние грани целевых функций в исходной и вспомогательной задачах; v_1'', v_2'' — верхние грани целевых функций в соответствующих обобщенных задачах.

Очевидно, x^* — решение вспомогательной задачи. Следовательно, $v_1 = v_2$.

В силу выпуклости функции $f(x)$

$$f(x) \leqslant L(x)$$

при $x \in R$. Поэтому $v_1'' \leqslant v_2''$.

Но согласно теореме 5

$$v_2 = v_2''.$$

Собирая вместе полученные соотношения, имеем

$$v_1 \leqslant v_1 \leqslant v_2'' = v_2 = v_1.$$

Итак,

$$v_1 = v_1'',$$

и утверждение теоремы 6 следует из теоремы 2.

Отметим, что требование гладкости функции $f(x)$ в теореме 6 можно ослабить: утверждение теоремы сохранит силу и для кусочно-гладких функций $f(x)$ (доказательство ничем не отличается от приведенного выше). Однако полностью снять предположение о гладкости целевой функции нельзя. Приведем соответствующий пример. Пусть требуется максимизировать выпуклую вверх функцию

$$f(x) = f(x_1, x_2) = x_1^{1/2} x_2^{1/2}$$

при условиях $-x_2 \geqslant 0$ и $x_1 \geqslant 0$, $x_2 \geqslant 0$ (R).

В данном случае G — положительная полуось $x_1 \geqslant 0$, R — положительный квадрант.

Имеем

$$f(x_1, 0) = 0 \text{ при } x_1 \geqslant 0, \text{ т. е. } v = 0.$$

С другой стороны,

$$\psi(\lambda_1) = \sup_{x_1 \geqslant 0,\ x_2 \geqslant 0} [x_1^{1/2} x_2^{1/2} - \lambda_1 x_2] = \infty$$

при любом λ_1.

Следовательно,

$$\tilde{v} = \inf_{\lambda_1 \geqslant 0} \psi(\lambda_1) = \infty,$$

т. е. $v < \tilde{v}$.

Пара двойственных задач (1) — (3) и (9) — (11) тесно связана с задачей об отыскании седловой точки функции Лагранжа $F(x, \lambda)$.

Теорема 7. *Для существования седловой точки функции Лагранжа $F(x, \lambda)$ при $x \in R$, $\lambda \geqslant 0$ необходимо и достаточно, чтобы задачи (1) — (3), (9) — (11) были связаны соотношением двойственности и имели в качестве решений некоторые точки $x^* \in G$, $\lambda^* \geqslant 0$. При этом любая пара $x^* \in G$, $\lambda^* \geqslant 0$ решений двойственных задач составляет седловую точку функции $F(x, \lambda)$ и, обратно, седловая точка (x^*, λ^*) функции $F(x, \lambda)$ определяет решения x^* и λ^* задач (1) — (3) и (9) — (11) соответственно.*

Доказательство. Пусть

$$v = \tilde{v} = f(x^*) = \psi(\lambda^*).$$

Согласно (8)

$$f(x^*) = \varphi(x^*) = \inf_{\lambda \geqslant 0} F(x^*, \lambda).$$

Следовательно,

$$\sup_{x\in R} F(x, \lambda^*) = \inf_{\lambda\geqslant 0} F(x^*, \lambda) = \alpha. \tag{39}$$

Используя (39), имеем

$$F(x^*, \lambda^*) \leqslant \sup_{x\in R} F(x, \lambda^*) = \alpha,$$

$$F(x^*, \lambda^*) \geqslant \inf_{\lambda\geqslant 0} F(x^*, \lambda) = \alpha,$$

т. е.

$$F(x^*, \lambda^*) = \alpha. \tag{40}$$

Равенства (39) и (40) означают, что $(x^*; \lambda^*)$ — седловая точка функции $F(x, \lambda)$ при $x \in R$, $\lambda \geqslant 0$.

Пусть теперь (x^*, λ^*) — седловая точка функции Лагранжа $F(x, \lambda)$ при $x \in R$, $\lambda \geqslant 0$, т. е.

$$F(x, \lambda^*) \leqslant F(x^*, \lambda^*) \leqslant F(x^*, \lambda) \tag{41}$$

для $x \in R$, $\lambda \geqslant 0$.

Из (41) вытекает, что

$$\varphi(x^*) = \inf_{\lambda\geqslant 0} F(x^*, \lambda) \geqslant \psi(\lambda^*) = \sup_{x\in R} F(x, \lambda^*),$$

или согласно лемме 1

$$f(x^*) = \varphi(x^*) \geqslant \psi(\lambda^*).$$

С другой стороны, по лемме 2

$$f(x^*) \leqslant v \leqslant v'' \leqslant \tilde{v} \leqslant \psi(\lambda^*).$$

Следовательно,

$$f(x^*) = \psi(\lambda^*) = v = \tilde{v},$$

т. е. задачи (1)—(3) и (9)—(11) связаны соотношением двойственности, причем x^* и λ^* — решения этих задач. Теорема доказана.

6. Остановимся на технике составления двойственных задач. Прежде чем приступить к формулировке двойственной задачи, необходимо разделить условия исходной задачи на две группы. Первую группу составляют условия (2), вторую — условия, высекающие в n-мерном пространстве множество R. Например, в задачах линейного программирования во вторую группу обычно включают условия $x_j \geqslant 0$ или $\alpha_j \leqslant x_j \leqslant \beta_j$.

Основная операция при составлении двойственной задачи заключается в построении функции $\psi(\lambda)$ из соотношения (9) или (36).

Пусть G^* — множество точек λ, для которых $\psi(\lambda) < \infty$.

Если $\lambda^{(1)}$, $\lambda^{(2)}$ — любая пара точек m-мерного пространства, то при $0 \leqslant \alpha \leqslant 1$

$$\psi(\alpha\lambda^{(1)} + (1-\alpha)\lambda^{(2)}) = \sup_{x\in R} [\alpha F(x, \lambda^{(1)}) + (1-\alpha) F(x, \lambda^{(2)})] \leqslant$$

$$\leqslant \alpha \sup_{x\in R} F(x, \lambda^{(1)}) + (1-\alpha) \sup_{x\in R} F(x, \lambda^{(2)}) = \alpha\psi(\lambda^{(1)}) + (1-\alpha)\psi(\lambda^{(2)}). \tag{42}$$

Из соотношения (42) вытекает, что, во-первых, G^* — выпуклое множество (если оно не пусто), а, во-вторых, функция $\psi(\lambda)$, определенная на множестве G^*, выпукла вниз.

Таким образом, двойственная задача (9) — (11) заключается в минимизации выпуклой вниз функции, определенной на выпуклом множестве, т. е. представляет собой задачу выпуклого программирования. Подчеркнем, что этот вывод не требует никаких предположений относительно исходной задачи. Поэтому определение величины $\tilde{v}$, которая согласно лемме 2 мажорирует верхнюю грань v исходной задачи, всегда сводится к анализу задачи выпуклого программирования. Это обстоятельство иногда позволяет получить оценку сверху для глобального максимума многоэкстремальной задачи.

В некоторых случаях множество G^* и функция $\psi(\lambda)$ могут быть найдены в явном виде.

Допустим, что исходная задача состоит в максимизации выпуклой вверх функции

$$f(x) \tag{43}$$

при условиях

$$Ax = b, \tag{44}$$

$$x \in R, \tag{45}$$

где $A = \|a_{ij}\|_{m,\,n}$, $b = (b_1, b_2, \ldots, b_m)^T$.

Предполагается, что ранг матрицы A равен m.

Пусть

$$\tilde{f}(p) = \sup_{x \in R}\left[f(x) - \sum_{j=1}^{n} x_j p_j\right], \qquad p = (p_1, p_2, \ldots, p_n),$$

$\tilde{R}$ — множество точек p, для которых $\tilde{f}(p) < \infty$.

Функция $\tilde{f}(p)$, определенная на множестве $\tilde{R}$, называется сопряженной к выпуклой функции $f(x)$, заданной на выпуклом множестве R (см., например, [2], гл. 7). Легко проверить, что $\tilde{R}$ — непустое выпуклое множество, $\tilde{f}(p)$ — выпуклая вниз функция. Задача, двойственная по отношению к задаче (43) — (45), может быть сформулирована с помощью сопряженной функции $\tilde{f}(p)$.

Образуем функцию Лагранжа

$$F(x, \lambda) = f(x) - \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij}\lambda_i + \sum_{i=1}^{m} b_i\lambda_i.$$

Имеем

$$\psi(\lambda) = \sup_{x \in R} F(x, \lambda) = \sup_{x \in R}[f(x) - (x, p)] + \sum_{i=1}^{m} b_i\lambda_i = \tilde{f}(p) + (b, \lambda),$$

где $p = \lambda A$.

Таким образом, задача, двойственная по отношению к (43) — (45), заключается в минимизации функции $\tilde{f}(\lambda A) + (b, \lambda)$ на всем пространстве точек λ или, что то же самое, на множестве G^*, состоящем из точек λ, для которых $\lambda A \in \tilde{R}$.

В данном случае для соблюдения соотношения двойственности достаточно потребовать, чтобы либо система (44) имела в качестве решения одну из внутренних точек множества R (теорема 4), либо функция $f(x)$ была непрерывна на замкнутом множестве R, а непустое множество планов G исходной задачи было ограниченным (теорема 3).

Допустим, что

$$f(x) = f_1(x^{(1)}) + (c^{(2)}, x^{(2)}),$$

где $x = (x^{(1)}, x^{(2)})$, причем $f_1(x^{(1)})$ — строго выпуклая гладкая функция, определенная на всем пространстве точек $x^{(1)}$.

Потребуем, кроме того, чтобы

$$1/_r \max_{|x^{(1)}|=r} f(x^{(1)}) \to -\infty \quad \text{при } r \to \infty. \tag{46}$$

Исходная задача состоит в максимизации функции

$$f(x) = f_1(x^{(1)}) + (c^{(2)}, x^{(2)}) \tag{47}$$

при условиях

$$A^{(1)}x^{(1)} + A^{(2)}x^{(2)} = b, \tag{48}$$

$$x^{(1)} \geqslant 0,\ x^{(2)} \geqslant 0. \tag{49}$$

В качестве R примем множество точек $x = (x^{(1)}, x^{(2)})$, у которых $x^{(2)} \geqslant 0$.

Пусть $p = (p_1, p_2)$.

Очевидно,

$$\tilde{f}(p) = \begin{cases} \tilde{f}_1(p^{(1)}), & \text{если } p^{(2)} \geqslant c^{(2)}, \\ \infty & \text{в противном случае.} \end{cases} \tag{50}$$

Переменными двойственной задачи являются

$$\lambda = (\lambda^{(1)}, \lambda^{(2)}),\ \lambda^{(1)} = (\lambda_1^{(1)}, \lambda_2^{(1)}, \ldots, \lambda_m^{(1)}),\ \lambda^{(2)} = (\lambda_1^{(2)}, \lambda_2^{(2)}, \ldots, \lambda_k^{(2)}),$$

где $\lambda^{(1)}$ отвечает условиям (48). $\lambda^{(2)} \geqslant 0$ соответствует условиям $x^{(1)} \geqslant 0$, k — число компонент вектора $x^{(1)}$.

Учитывая (50), приходим к следующей формулировке двойственной задачи.

Минимизировать

$$\tilde{f}_1(\lambda^{(1)}A^{(1)} - \lambda^{(2)}E) + (\lambda^{(1)}, b) \tag{51}$$

при условиях

$$\lambda^{(1)}A^{(2)} \geqslant c^{(2)}, \tag{52}$$

$$\lambda^{(2)} \geqslant 0, \tag{53}$$

где E — единичная матрица порядка k.

Уточним вид сопряженной функции $\tilde{f}(p^{(1)})$.

Из условия (46) вытекает, что

$$\tilde{f}_1(p^{(1)}) = \sup_{x^{(1)}} [f_1(x^{(1)}) - (x^{(1)}, p)] = \max_{x^{(1)}} [f_1(x^{(1)}) - (x^{(1)}, p^{(1)})] =$$

$$= f(\bar{x}^{(1)}) - (\bar{x}^{(1)}, p^{(1)}).$$

В силу гладкости функции $f_1(x^{(1)})$ экстремальная точка $\bar{x}^{(1)} = \bar{x}^{(1)}(p)$ определяется из уравнения

$$\operatorname{grad} f_1(x^{(1)}) = p^{(1)}. \tag{54}$$

Пользуясь тем, что уравнение (54) имеет единственное решение при любом $p^{(1)}$ ($f_1(x^{(1)})$ — строго выпуклая функция), введем функцию $t(p^{(1)})$, обратную по отношению к левой части (54).

В таком случае

$$\tilde{f}_1(p^{(1)}) = f_1(t(p^{(1)})) - (t(p^{(1)}),\ p^{(1)}), \tag{55}$$

причем $\operatorname{grad} f_1(t(p^{(1)})) \equiv p^{(1)}$.

Соотношения (55) определяют так называемое преобразование Лежандра.

Соотношения (51)—(53) и (55) составляют задачу, двойственную относительно задачи (47)—(49). Формулировка задачи (51)—(53), (55) имеется в [4].

Если в соотношениях (51)—(53), (55) положить

$$f_1(x^1) = x^{(1)T}Qx^{(1)} + (c^{(1)}, x^{(1)}),$$

где $Q = \| g_{ij} \|_k$ — матрица отрицательно определенной квадратичной формы, $c^{(1)}$ — k-мерный вектор, то получим двойственную задачу квадратичного программирования, сформулированную в [4].

Заметим, что задачи (47)—(49) и (51)—(53), (55) всегда связаны соотношением двойственности, так как исходная задача имеет оптимальный план (см. теорему 6).

Приведем еще один пример формирования двойственной задачи.

Исходная задача состоит в максимизации функции

$$f(x) = \sum_{j=1}^{n} f_j(x_j) \tag{56}$$

при условиях

$$\sum_{j=1}^{n} a_{ij}x_j = b_i, \quad i = 1, 2, \ldots, m, \tag{57}$$

$$0 \leqslant x_j \leqslant d_j. \tag{58}$$

Здесь $f_j(x_j)$ — кусочно-линейная выпуклая вверх функция, заданная на $[0, d_j]$.

Пусть

$$d_{0j} = 0 < d_{1j} < \ldots < d_{l_j j} < d_j = d_{l_j+1,\, j}$$

— значения x_j, при которых функция $f_j(x_j)$ терпит излом; c_{ij} — угловой коэффициент функции $f_j(x_j)$ на отрезке

$$[d_{i-1,\, j},\ d_{ij}], \qquad i = 1, 2, \ldots, l_{j+1}.$$

Приняв в качестве R множество с условиями (58), займемся построением функции $\tilde{f}(p)$.

Имеем

$$\tilde{f}(p) = \sup_{x\in R}\left[\sum_{j=1}^{n} f_j(x_j) - \sum_{j=1}^{n} p_j x_j\right] = \sum_{j=1}^{n} \sup_{0\leqslant x_j\leqslant d_j} [f_j(x_j) - x_j p_j] = \sum_{j=1}^{n} \tilde{f}_j(p_j).$$

Для определения каждой из функций $\tilde{f}_j(p_j)$ заметим, что

$$\tilde{f}_j(p_j) = \sup_{0\leqslant x_j\leqslant d_j} [f_j(x_j) - x_j p_j] =$$

$$= \begin{cases} f(d_{ij}) - d_{ij}p_j, & \text{если } c_{i+1,\, j} \leqslant p_j \leqslant c_{ij} \quad (i = 1, 2, \ldots, l_j); \\ f_j(0), & \text{если } p_j \geqslant c_{1j}; \\ f_j(d_j) - d_j p_j, & \text{если } p_j \leqslant c_{l_j+1,\, j}, \quad d_j < \infty; \\ \infty, & \text{если } p_j \leqslant c_{l_j+1,\, j}, \quad d_j = \infty. \end{cases}$$

Из этого соотношения вытекает, что $\tilde{f}_j(p_j)$ — кусочно-линейная выпуклая вниз функция, которая терпит излом в точках

$$c_l \quad < c_{l\,j} < \ldots < c_{1j},$$

причем на отрезке $[c_{ij}, c_{i-1,j}]$, $i = 2, 3, \ldots, l_j + 1$ угловым коэффициентом функции $\tilde{f}_j(p_j)$ является число $d_{i-1,j}$ на луче $[c_{1j}, \infty]$ ее угловой коэффициент — $d_{0j} = 0$. Что касается луча $[-\infty, c_{l_j+1,j}]$, то на нем функция $\tilde{f}_j(p_j)$ определена только в случае $d_j < \infty$ и при выполнении этого условия d_j — ее угловой коэффициент. Итак, при переходе от функции $f_j(x)$ к сопряженной функции $\tilde{f}_j(p_j)$ параметры d_{ij} и c_{ij} меняются ролями: c_{ij} — точки излома функции $\tilde{f}_j(p_j)$, d_{ij} — угловые коэффициенты участков линейности этой функции. Отметим, что в данном случае множество точек p, в которых $f(p) < \infty$, задается условиями $p_j \geqslant c_{l_j+1,j}$, если $d_j = \infty$.

Основываясь на общей схеме, рассмотренной при анализе задачи (43) — (45), получаем формулировку задачи, двойственной относительно (56) — (58).

Требуется минимизировать

$$\sum_{j=1}^{n} \tilde{f}_j\left(\sum_{i=1}^{m} a_{ij}\lambda_i\right) + \sum_{i=1}^{m} \lambda_i b_i \tag{59}$$

при условии

$$\sum_{i=1}^{m} a_{ij}\lambda_i \geqslant c_{l_j+1,j}, \quad \text{если} \quad d_j = \infty. \tag{60}$$

Расширим формулировку задачи (56) — (58), положив

$$f(x) = \sum_{j=1}^{n} f_j(x_j) + \sum_{i=m+1}^{m+k} f_{i+n-m}\left(\sum_{j=1}^{n} a_{ij}x_j\right).$$

Если ввести дополнительные переменные

$$x_{i+n-m} = -\sum_{j=1}^{n} a_{ij}x_j, \qquad i = m+1, m+2, \ldots, m+k,$$

то новая задача приводится к задаче (56) — (58). Учитывая это, легко получить следующую расширенную формулировку двойственной задачи.

Минимизировать

$$\sum_{j=1}^{n} \tilde{f}_j\left(\sum_{i=1}^{m+k} a_{ij}\lambda_i\right) + \sum_{j=n+1}^{n+k} \tilde{f}_j(\lambda_{j+m-n}) + \sum_{i=1}^{m} b_i\lambda_i$$

при условиях

$$\sum_{i=1}^{m+k} a_{ij}\lambda_i \geqslant c_{l_j+1,j}, \quad \text{если} \quad d_j = \infty \ (j = 1, 2, \ldots, n),$$

$$c_{l_j+1,j} \leqslant \lambda_{j+m-n} \leqslant c_{1j}, \qquad j = n+1, \ldots, n+k.$$

ЛИТЕРАТУРА

1. W. Fenchel. Convex Cones, sets and functions. Office of naval research logistics project report, Depart. of mathemat. Princeton University, 1953.
2. С. Карлин. Математические методы в теории игр, программировании и экономике. М., Изд-во «Мир», 1964.
3. К. Дж. Эрроу, Л. Гурвиц, Х. Удзава. Исследования по линейному и нелинейному программированию. М., Изд-во иностр. лит., 1962.
4. Д. Деннис. Математическое программирование и электрические цепи. М., Изд-во иностр. лит., 1961.

Поступила в редакцию
26 II 1965