ГЛОБАЛЬНАЯ РЕГУЛЯРНОСТЬ В МАТЕМАТИЧЕСКОМ ПРОГРАММИРОВАНИИ*

Манлио Гаудиозо — gaudioso@deis.unical.it
В. Н. Малозёмов — malv@math.spbu.ru

28 октября 2008 г.

Делается попытка переосмыслить и дополнить содержание докладов [1, 2].

1°. Рассмотрим задачу математического программирования

$$\begin{gathered} f(x) \to \inf\,, \\ g_i(x) \leqslant 0\,, \quad i \in 1:s\,; \\ x \in P\,. \end{gathered} \tag{1}$$

Предполагается, что $P \subset \mathbb{R}^n$ — произвольное непустое множество (возможно, дискретное), $f, g_1, \dots, g_s$ — произвольные конечные функции, заданные на P. Обозначим

$$X = \big\{x \in P \mid g_i(x) \leqslant 0,\ i \in 1:s\big\}\,,$$

$$f^* = \inf\big\{f(x) \mid x \in X\big\}.$$

Введём функцию Лагранжа

$$L(x,y) = f(x) + \sum_{i=1}^{s} y_i\, g_i(x)$$

и запишем двойственную задачу

$$\varphi(y) := \inf\big\{L(x,y) \mid x \in P\big\} \to \sup_{y \in \mathbb{R}^s_+}\,. \tag{2}$$

Здесь $\mathbb{R}^s_+$ — множество векторов $y = (y_1, \dots, y_s)$ с неотрицательными компонентами. Отметим, что функция Лагранжа $L(x,y)$ при фиксированном x

*Семинар по дискретному гармоническому анализу и геометрическому моделированию «DHA & CAGD»: `http://www.dha.spb.ru/`

аффинна по y, поэтому целевая функция двойственной задачи $\varphi(y)$ *вогнута* на $\mathbb{R}^s_+$. Обозначим

$$\varphi^* = \sup\big\{\varphi(y) \mid y \in \mathbb{R}^s_+\big\}.$$

Наряду с (1), (2) рассмотрим вспомогательную параметрическую задачу

$$\begin{gathered} f(x) \to \inf\,, \\ g_i(x) \leqslant v_i\,, \quad i \in 1:s\,; \\ x \in P\,. \end{gathered} \tag{3}$$

Обозначим

$$X(v) = \big\{x \in P \mid g_i(x) \leqslant v_i\,,\ i \in 1:s\big\}.$$

Функция

$$F(v) = \inf\big\{f(x) \mid x \in X(v)\big\}, \quad v \in \mathbb{R}^s,$$

называется *функцией чувствительности* для задачи (1).

Исходная задача (1) вкладывается в задачу (3) (при $v = \mathbb{O}$). Функция $F(v)$ отражает характер этого вложения. Вместе с тем, функция чувствительности связана и с целевой функцией двойственной задачи (2). Справедливо следующее утверждение.

ЛЕММА. *При всех $y \in \mathbb{R}^s_+$ функция $\varphi(y)$ допускает представление*

$$\varphi(y) = \inf\big\{F(v) + \langle y, v\rangle \mid v \in \mathbb{R}^s\big\}.$$

При исследовании пары двойственных задач (1), (2) функция чувствительности играет важную роль.

2°. Относительно задачи (1) сделаем естественные предположения

$$X \neq \varnothing\,, \qquad f^* > -\infty\,, \tag{4}$$

которые гарантируют, что f^* — конечная величина.

Из определений следует, что $f^* \geqslant \varphi^*$. Первый вопрос, на который нужно ответить: когда выполняется *соотношение двойственности* $f^* = \varphi^*$?

ТЕОРЕМА 1. *Для того чтобы имело место соотношение двойственности, необходимо и достаточно, чтобы ε-субдифференциал функции чувствительности в нуле, $\partial_\varepsilon F(\mathbb{O})$, был непустым при всех $\varepsilon > 0$.*

3°. Условие

$$\partial F(\mathbb{O}) \neq \varnothing \tag{5}$$

назовём условием *глобальной регулярности* задачи (1). По существу, оно характеризует «регулярность» вложения задачи (1) в параметрическое семейство задач (3).

ТЕОРЕМА 2. *Условие глобальной регулярности выполняется тогда и только тогда, когда $f^* = \varphi^*$ и двойственная задача (2) имеет решение.*

По ходу доказательства этой теоремы выясняется, что

- любая точка из $-\partial F(\mathbb{O})$ является решением задачи (2);
- при выполнении соотношения двойственности $f^* = \varphi^*$ любое решение двойственной задачи принадлежит множеству $-\partial F(\mathbb{O})$.

4°. Приведём одно достаточное условие, гарантирующее глобальную регулярность задачи (1).

ТЕОРЕМА 3. *Пусть $f^* = \varphi^*$ и выполнено условие Слейтера: существует точка $z \in P$, такая, что $g_i(z) < 0$ при всех $i \in 1:s$. Тогда $\partial F(\mathbb{O}) \neq \varnothing$.*

5°. Предположим, что выполнено условие глобальной регулярности (5) задачи (1). Возьмём решение двойственной задачи y^* и рассмотрим задачу Лагранжа

$$L(x, y^*) \to \inf_{x \in P} . \tag{6}$$

Её экстремальное значение равно $\varphi(y^*) = \varphi^*$.

ТЕОРЕМА 4 (принцип Лагранжа). *Любое решение x^* глобально регулярной задачи (1) является решением задачи (6). При этом выполняется условие дополнительности*

$$y_i^* \, g_i(x^*) = 0 \,, \quad i \in 1:s \,. \tag{7}$$

Приведём пример, когда задачи (1) и (6) имеют единственные, но разные решения.

ПРИМЕР. Рассмотрим экстремальную задачу

$$\begin{aligned} f(x) &:= x - 1 \to \inf \,, \\ g_1(x) &:= x^2 - 1 \leqslant 0 \,, \\ g_2(x) &:= -x^2 + 1 \leqslant 0 \,, \\ P &= \mathbb{R}_+ \,. \end{aligned}$$

Она имеет единственное решение $x^* = 1$; при этом $f^* = 0$.

Запишем функцию Лагранжа

$$\begin{aligned} L(x, y) &= x - 1 + y_1 \, (x^2 - 1) + y_2 \, (-x^2 + 1) = \\ &= x^2 \, (y_1 - y_2) + x - (y_1 - y_2 + 1) \,. \end{aligned}$$

Для целевой функции двойственной задачи получаем формулу

$$\varphi(y) := \inf\big\{L(x,y) \mid x \geqslant 0\big\} = \begin{cases} -(y_1 - y_2 + 1) & \text{при } y_1 - y_2 \geqslant 0\,, \\ -\infty & \text{при } y_1 - y_2 < 0\,. \end{cases}$$

Здесь $y \in \mathbb{R}^2_+$. Ясно, что $\varphi^* = -1$.

Решением двойственной задачи является любой вектор $y^* \geqslant \mathbb{O}$ с $y_1^* = y_2^*$. Для таких векторов задача Лагранжа принимает вид

$$L(x, y^*) := x - 1 \to \inf_{x \geqslant 0}.$$

Её единственное решение $\hat{x} = 0$ отлично от решения $x^* = 1$ исходной задачи.

В данном случае не выполняется соотношение двойственности, поскольку $f^* = 0$, $\varphi^* = -1$.

6°. Обратимся к достаточным условиям глобальной оптимальности. Говорят [3, с. 144], что пара $\{x^*, y^*\}$, где $x^* \in P$, $y^* \in \mathbb{R}^s_+$, удовлетворяет условию *глобальной оптимальности*, если

$$(\alpha) \qquad L(x^*, y^*) = \min_{x \in P} L(x, y^*)\,;$$

$$(\beta) \qquad y_i^*\, g_i(x^*) = 0\,, \qquad i \in 1:s\,;$$

$$(\gamma) \qquad g_i(x^*) \leqslant 0\,, \qquad i \in 1:s\,.$$

ТЕОРЕМА 5. *Если $\{x^*, y^*\}$ — глобально оптимальная пара, то x^* — решение задачи* (1), *y^* — решение задачи* (2) *и $f(x^*) = \varphi(y^*)$.*

Доказательство. Согласно (γ), $x^* \in X$. Для любого $x \in X$ в силу (β) и (α) имеем

$$f(x^*) = f(x^*) + \sum_{i=1}^{s} y_i^*\, g_i(x^*) = L(x^*, y^*) \leqslant L(x, y^*) = \\ = f(x) + \sum_{i=1}^{s} y_i^*\, g_i(x) \leqslant f(x)\,.$$

Это значит, что x^* — решение задачи (1), $f(x^*) = f^*$. Далее

$$\varphi^* \geqslant \varphi(y^*) = \inf\big\{L(x, y^*) \mid x \in P\big\} = L(x^*, y^*) = \\ = f(x^*) + \sum_{i=1}^{s} y_i^*\, g_i(x^*) = f(x^*) = f^* \geqslant \varphi^*.$$

Отсюда следует, что $\varphi(y^*) = \varphi^*$ и $f(x^*) = \varphi(y^*)$.

Теорема доказана. □

ТЕОРЕМА 6. *Глобально оптимальная пара $\{x^*, y^*\}$ является седловой точкой функции Лагранжа, то есть при всех $x \in P$ и $y \in \mathbb{R}^s_+$*

$$L(x^*, y) \leqslant L(x^*, y^*) \leqslant L(x, y^*)\,. \tag{8}$$

Доказательство. При всех $y \in \mathbb{R}^s_+$ имеем

$$L(x^*, y) = f(x^*) + \sum_{i=1}^{s} y_i\, g_i(x^*) \leqslant f(x^*)\,,$$

поэтому

$$\sup\left\{L(x^*, y) \mid y \in \mathbb{R}^s_+\right\} \leqslant f(x^*)\,.$$

Учитывая равенство $f(x^*) = \varphi(y^*)$, получаем

$$L(x^*, y^*) \leqslant \sup\left\{L(x^*, y) \mid y \in \mathbb{R}^s_+\right\} \leqslant f(x^*) =$$

$$= \varphi(y^*) = \inf\left\{L(x, y^*) \mid x \in P\right\} \leqslant L(x^*, y^*)\,.$$

Приходим к соотношениям

$$\sup\left\{L(x^*, y) \mid y \in \mathbb{R}^s_+\right\} = L(x^*, y^*) = \inf\left\{L(x, y^*) \mid x \in P\right\},$$

равносильным (8). Теорема доказана. $\square$

7°. В качестве простейшего объекта применения общих результатов рассмотрим задачу линейного программирования

$$\begin{aligned} f(x) &:= \langle c, x\rangle \to \inf\,, \\ A\,x &\leqslant b\,, \\ P &= \mathbb{R}^n_+\,. \end{aligned} \tag{9}$$

Как задача линейного программирования она имеет двойственную

$$\begin{aligned} \psi(u) &:= \langle b, u\rangle \to \sup\,, \\ u\,A &\leqslant c\,, \quad u \leqslant \mathbb{O}\,. \end{aligned} \tag{10}$$

Множество планов задачи (10) обозначим U. В условиях (4) обе задачи (9) и (10) имеют решения и выполняется соотношение двойственности

$$f^* := \min_{x \in X} f(x) = \max_{u \in U} \psi(u) \;=:\; \psi^*.$$

Выясним, как связаны задачи (10) и (2). Имеем

$$\varphi(y) := \inf\left\{\langle c, x\rangle + \langle y, A\,x - b\rangle \mid x \in \mathbb{R}^n_+\right\} =$$

$$= -\langle b, y\rangle + \inf\left\{\langle c + y\,A, x\rangle \mid x \in \mathbb{R}^n_+\right\}.$$

Ясно, что

$$\inf\left\{\langle c+y\,A, x\rangle \mid x\in\mathbb{R}^n_+\right\} = \begin{cases} 0\,, & \text{если } c+y\,A\geqslant \mathbb{O},\\ -\infty\,, & \text{в противном случае.}\end{cases}$$

Поэтому задача (2) принимает вид

$$\begin{aligned} &\varphi(y) := -\langle b, y\rangle \to \sup\,,\\ &c+y\,A\geqslant\mathbb{O}, \quad y\geqslant\mathbb{O}. \end{aligned} \tag{11}$$

Задачи (10) и (11) эквивалентны. Множества их решений U^* и Y^* связаны соотношением $Y^*=-U^*$, и $\varphi^*=\psi^*$.

Получили, что $f^*=\psi^*=\varphi^*$ и двойственная задача (11) имеет решение. По теореме 2 в условиях (4) задача линейного программирования (9) глобально регулярна.

ТЕОРЕМА 7. *Множество решений задачи* (10) *совпадает с субдифференциалом функции чувствительности задачи* (9) *в нуле, то есть* $U^*=\partial F(\mathbb{O})$.

Доказательство. Запишем вспомогательную параметрическую задачу

$$\begin{gathered} f(x) := \langle c, x\rangle \to \inf\,,\\ A\,x\leqslant b+v\,,\\ P=\mathbb{R}^n_+\,. \end{gathered}$$

Возьмём $u^*\in U^*$ и покажем, что

$$F(v)-F(\mathbb{O})\geqslant\langle u^*, v\rangle \quad \forall\, v\in\mathbb{R}^s. \tag{12}$$

При любом $x\in X(v)$ имеем

$$\langle c,x\rangle \geqslant \langle c,x\rangle + \langle u^*, b+v-A\,x\rangle = \langle b,u^*\rangle + \langle u^*,v\rangle +$$

$$+\,\langle c-u^*A, x\rangle \geqslant \langle b,u^*\rangle + \langle u^*,v\rangle = f^*+\langle u^*,v\rangle\,.$$

Отсюда следует, что при всех $v\in\mathbb{R}^s$

$$F(v)\geqslant f^*+\langle u^*,v\rangle\,.$$

Учитывая равенство $f^*=F(\mathbb{O})$, приходим к (12).

Наоборот, пусть $u^*\in\partial F(\mathbb{O})$, так что выполняется неравенство (12). Подставляя в (12) на место v единичные орты e_i и принимая во внимание, что $F(e_i)\leqslant F(\mathbb{O})$, получаем $u^*_i\leqslant 0$. Значит, $u^*\leqslant\mathbb{O}$.

Далее, в силу (12)

$$F(v) + \langle -u^*, v\rangle \geqslant F(\mathbb{O}) \qquad \forall\ v \in \mathbb{R}^s.$$

По лемме из п. 1°, $\varphi(-u^*) \geqslant F(\mathbb{O})$. Соотношения

$$\varphi^* \geqslant \varphi(-u^*) \geqslant F(\mathbb{O}) = f^* \geqslant \varphi^*$$

убеждают нас в том, что вектор $-u^*$ является решением задачи (11). В этом случае, как отмечалось, u^* — решение задачи (10).

Теорема доказана. □

К данной теореме примыкает такой результат (см. [4]).

ТЕОРЕМА 8. *Если u — план двойственной задачи линейного программирования* (10)*, то $u \in \partial_\varepsilon F(\mathbb{O})$, где $\varepsilon = f^* - \langle b, u\rangle$.*

Доказательство. Нужно проверить, что

$$F(v) - F(\mathbb{O}) \geqslant \langle u, v\rangle - \varepsilon \qquad \forall\ v \in \mathbb{R}^s.$$

При $X(v) = \varnothing$ это очевидно. Пусть $X(v) \neq \varnothing$. Для любого $x \in X(v)$ имеем

$$\langle c, x\rangle \geqslant \langle c, x\rangle + \langle u, b + v - A\,x\rangle = \langle b, u\rangle + \langle u, v\rangle +$$
$$+ \langle c - uA, x\rangle \geqslant f^* - \varepsilon + \langle u, v\rangle = F(\mathbb{O}) + \langle u, v\rangle - \varepsilon\,.$$

Остаётся взять инфимум по $x \in X(v)$.

Теорема доказана. □

8°. Современное состояние теории чувствительности в оптимизации представлено в книге [5].

ЛИТЕРАТУРА

1. Лазарев А. В. *О соотношении двойственности в математическом программировании* // Семинар «DHA & CAGD». Избранные доклады. 17 мая 2008 г. (`http://dha.spb.ru/reps08.shtml#0517`).

2. Лазарев А. В. *Необходимые условия глобальной оптимальности* // Семинар «DHA & CAGD». Избранные доклады. 9 сентября 2008 г. (`http://dha.spb.ru/reps08.shtml#0909`).

3. Shapiro J. *Mathematical Programming: Structures and Algorithms.* J. Wiley & Sons, 1979.

4. De Leone R., Gaudioso M., Monaco M. F. *Nonsmooth optimization methods for parallel decomposition of multicommodity flow problems* // Annals of Operation Research. 1993. Vol. 44. P. 299–311.

5. Измаилов А. Ф. *Чувствительность в оптимизации.* М.: Физматлит, 2006.