ЖУРНАЛ
ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ

Том 30, 1990 — N° 1, с. 43–57

УДК 519.85

Ю.Г. ЕВТУШЕНКО, В.Г. ЖАДАН

ТОЧНЫЕ ВСПОМОГАТЕЛЬНЫЕ ФУНКЦИИ В ЗАДАЧАХ ОПТИМИЗАЦИИ

117967 Москва, ул. Вавилова, 40, ВЦ РАН

(Пересмотрена 20 февраля 2002 г.)

Вводится новое понятие точной вспомогательной функции, задача минимизации которой имеет то же множество решений, что и исходная оптимизационная задача. Даны достаточные условия для того, чтобы вспомогательная функция была точной, приведены примеры таких функций. Введение точных вспомогательных функций позволяет редуцировать решение исходной задачи к однократной минимизации вспомогательной функции. Во многих случаях задача условной минимизации сводится к задаче безусловной минимизации.

§ 1. Основные определения

Рассмотрим задачу нелинейного программирования: найти

$$f_* = \min_{x\in X} f(x), \qquad X = \{x \in \mathbb{E}^n \mid g(x) \le 0\}. \tag{1.1}$$

Здесь $\mathbb{E}^n$ есть n-мерное евклидово пространство, $f(x)$ и $g(x)$ — непрерывные функции, осуществляющие соответственно отображения $f : \mathbb{E}^n \to \mathbb{E}^1$, $g : \mathbb{E}^n \to \mathbb{E}^m$. Решением задачи (1.1) является любая точка x_* из множества

$$X_* = \{x_* \in X \mid f(x) - f(x_*) \ge 0 \quad \forall x \in X\},$$

которое всюду ниже предполагается непустым. Все результаты данной работы получены без использования условий выпуклости и дифференцируемости функций f и g по x.

Введем функцию $R(x, y)$, зависящую от исходных переменных $x \in \mathbb{E}^n$ и от некоторого вектора y из множества Y. Размерность y и вид множества Y пока не конкретизируем. Рассмотрим вспомогательную задачу минимизации

$$\min_{x\in P} R(x, y), \tag{1.2}$$

где P — некоторое замкнутое множество из $\mathbb{E}^n$, содержащее X_*. В частности, в качестве P может быть взято либо все пространство $\mathbb{E}^n$, либо само допустимое множество X, либо его часть.

Предполагая, что решение задачи (1.2) существует, определим точечно-множественное отображение

$$X(y) = \operatorname{Arg}\min_{x\in P} R(x, y).$$

Определение 1. *Функция $R(x,y)$ — точная вспомогательная функция (т.в.ф.) для задачи* (1.1) *на $P \times Y$, если $X(y) \neq \emptyset$ и $X(y) = X_*$ для любого $y \in Y$.*

Изучение т.в.ф. исключительно важно с точки зрения построения численных методов, так как знание таких функций позволяет решить исходную задачу (1.1) в результате однократной минимизации (1.2) вспомогательной функции. При этом желательно, чтобы множество Y было возможно "шире", чтобы не возникали дополнительные трудности определения точек из Y. Крайне неудобны в этом отношении функции, для которых Y состоит из одной точки.

Впервые т.в.ф. были построены в работах [1], [2], где были обнаружены так называемые точные штрафные функции, являющиеся т.в.ф. В дальнейшем этим функциям было посвящено огромное число исследований, однако, насколько известно авторам, кроме [3, 4, 5] не было предложено никаких других принципиально новых т.в.ф. Ниже развивается подход, использованный авторами в [6, 7, 8]. Даются достаточные условия т.в.ф., приведены примеры таких функций.

Удобно строить т.в.ф. в виде ограниченной снизу функции. Пусть $R(x,y)$ — т.в.ф. для задачи (1.1) на множестве $P \times Y$. Возьмем функцию

$$M(x,y) = R(x,y) - R(x_*,y), \qquad x_* \in X_* \subseteq P. \tag{1.3}$$

Для любых $x \in P$, $y \in Y$, $x_* \in X_*$ функция M неотрицательна. Множество точек x, доставляющих минимум функции $M(x,y)$ по x на P, совпадает с множеством $X(y)$ и — если $y \in Y$ — с множеством X_*. Оно является также множеством решений уравнения $M(x,y) = 0$, принадлежащих P, при каждом фиксированном векторе y из Y. Выбор конкретного элемента x_* из X_* не влияет на значения функции $M(x,y)$ при $y \in Y$, так как $R(x,y)$ принимает одно и то же значение при всех x из X_*. Поэтому

$$M(x,y) \equiv 0 \qquad \forall x \in X(y) = X_*, \qquad \forall y \in Y.$$

Если $R(x,y)$ — т.в.ф., то функция $M(x,y)$ также т.в.ф. Верно и обратное: если построенная в виде (1.3) функция $M(x,y)$ является т.в.ф., то и $R(x,y)$ т.в.ф. Поэтому вместо доказательства того, что $R(x,y)$ — т.в.ф., достаточно показать, что т.в.ф. является соответствующая ей функция $M(x,y)$.

Для того чтобы построенная в виде (1.3) функция $M(x,y)$ была т.в.ф. на множестве $P \times Y$, необходимо и достаточно, чтобы выполнялись следующие два условия:

$$M(x,y) \geq 0 \qquad \forall x \in P, \qquad \forall y \in Y, \tag{1.4}$$

$$\forall y \in Y \text{ из } M(x,y) = 0 \text{ следует, что } x \in X(y) = X_*. \tag{1.5}$$

Если функция $M(x,y)$ дифференцируема по y, множество Y открытое, то необходимо, чтобы

$$M_y(x,y) = 0 \qquad \forall x \in X(y) = X_*, \qquad \forall y \in Y.$$

Это условие можно добавить к (1.5), их совместное рассмотрение во многих случаях позволяет легко показывать, что $X(y) = X_*$.

Умножение т.в.ф. на любое положительное число оставляет вспомогательную функцию точной. Сумма точных вспомогательных функций на одном и том же множестве $P \times Y$ есть т.в.ф. на том же самом множестве. Если множества $P \times Y$ разные, то следует взять их пересечение.

Пусть функция $Q(z)$ задана на множестве $Z \subseteq \mathbb{E}^s$; полярную к ней функцию $Q^0(z_0 \mid Z)$ на множестве $Z_0 \subseteq \mathbb{E}^s$ определим с помощью соотношений

$$Q^0(z_0 \mid Z) = \inf_{\mu \in M(z_0)} \mu, \qquad M(z_0) = \{\mu \in \overline{\mathbb{E}^1} \mid \langle z, z_0 \rangle \leq \mu Q(z) \;\; \forall z \in Z\}, \tag{1.6}$$

где $z_0 \in Z_0$, $\overline{\mathbb{E}^1}$ — расширенная прямая $\mathbb{E}^1$, т.е. прямая $\mathbb{E}^1$, к которой добавлены элементы $\{+\infty\}$ и $\{-\infty\}$. Из приведенного определения следует неравенство Минковского–Малера

$$\langle z, z_0\rangle \leq Q(z)Q^0(z_0 \mid Z) \qquad \forall z \in Z, \qquad \forall z_0 \in Z_0. \tag{1.7}$$

Если функция $Q(z)$ принимает только положительные значения на Z, то наряду с (1.6) можно пользоваться следующим определением:

$$Q^0(z_0 \mid Z) = \sup_{z\in Z} \frac{\langle z, z_0\rangle}{Q(z)}. \tag{1.8}$$

Если $Q(z)$ неположительна на Z, то

$$Q^0(z_0 \mid Z) = \inf_{z\in Z} \frac{\langle z, z_0\rangle}{Q(z)}. \tag{1.9}$$

Обозначим через $\mathbb{E}^m_+$ и $\mathbb{E}^m_-$ неотрицательный и неположительный ортанты $\mathbb{E}^m$, т.е. совокупности таких векторов из $\mathbb{E}^m$, у которых все координаты соответственно неотрицательны, неположительны.

Для каждой скалярной функции $\varphi(z)$ векторного аргумента z определим

$$\varphi_+(z) = \max[0, \varphi(z)], \qquad \varphi_-(z) = \min[0, \varphi(z)].$$

Аналогично, для векторов $z \in \mathbb{E}^s$

$$\begin{array}{ll} z_+ = [z^1_+, \ldots, z^s_+], & z^i_+ = \max[0, z^i], \\ z_- = [z^1_-, \ldots, z^s_-], & z^i_- = \min[0, z^i]. \end{array}$$

Определим p-ю гельдеровскую норму вектора z:

$$\|z\|_p = \left(\sum_{i=1}^{s} |z^i|^p\right)^{1/p}, \qquad p \geq 1. \tag{1.10}$$

Сопряженной к ней будет норма $\|z\|_{p_*}$, где $p^{-1} + p_*^{-1} = 1$.

Функцию (1.10) будем рассматривать и при ненулевых $p < 1$, а также при $p = 0$ и $p = \pm\infty$, доопределив ее следующим образом:

$$\|z\|_0 = s^{1/2}\left(\prod_{i=1}^{s} |z^i|\right)^{1/s}, \quad \|z\|_{+\infty} = \max_{1\leq i\leq s} |z^i|, \quad \|z\|_{-\infty} = \min_{1\leq i\leq s} |z^i|.$$

При $p < 0$ считается, что функция (1.10) равна нулю, если хотя бы одна координата вектора z обращается в нуль.

Приведем некоторые примеры для $Z_0 = \mathbb{E}^s_+$:

$$\begin{array}{llllll} p > 1, & p_* > 1, & Q(z) = \ \ \|z_+\|_p, & Q^0(z_0|\mathbb{E}^s) & = & \|z_0\|_{p_*}, \\ p < 1, & p_* < 1, & Q(z) = -\|z_-\|_p, & Q^0(z_0|\mathbb{E}^s_-) & = & \|z_0\|_{p_*}. \end{array}$$

Здесь p и p_* связаны между собой той же зависимостью $p^{-1} + p_*^{-1} = 1$. Укажем отдельно особые случаи:

$$\begin{array}{llllll} p = 0, & p_* = 0, & Q(z) = -\|z_-\|_0, & Q^0(z_0|\mathbb{E}^s_-) & = & \|z_0\|_0, \\ p = +\infty, & p_* = 1, & Q(z) = \ \ \|z_+\|_{+\infty}, & Q^0(z_0|\mathbb{E}^s) & = & \|z_0\|_1, \\ p = -\infty, & p_* = 1, & Q(z) = -\|z_-\|_{-\infty}, & Q^0(z_0|\mathbb{E}^s_-) & = & \|z_0\|_1. \end{array}$$

Введем вектор $w \in \mathbb{E}^m_+$ и составим функцию Лагранжа:

$$L(x,w) = f(x) + \langle w, g(x)\rangle.$$

Будем говорить, что функция Лагранжа обладает седловой точкой $[x_*, w_*] \in P \times \mathbb{E}^m_+$, если

$$L(x_*, w) \le L(x_*, w_*) \le L(x, w_*) \qquad \forall x \in P, \quad \forall w \in \mathbb{E}^m_+. \tag{1.11}$$

Известно [9], что если существует такой вектор $w_* \in \mathbb{E}^m_+$, что x_* вместе с w_* образуют седловую точку функции Лагранжа, то $x_* \in X_*$.

Введем сложную функцию $B(g(x))$. Через $g(P)$ обозначим образ множества P при отображении g. Тогда, используя неравенство (1.7), из правой части (1.11) получаем

$$f_* \le L(x, w_*) \le f(x) + B(g(x))B^0(w_* \mid g(P)) \quad \forall x \in P. \tag{1.12}$$

Если $B(g(x)) \ge 0$ для всех $x \in \mathbb{E}^n$, то можно эту оценку ослабить, введя проще вычисляемые полярные функции $B^0(w_* \mid g(\mathbb{E}^n))$ и $B^0(w_* \mid \mathbb{E}^m)$, так как согласно (1.8)

$$B^0(w_* \mid g(P)) \le B^0(w_* \mid g(\mathbb{E}^n)) \le B^0(w_* \mid \mathbb{E}^m).$$

Поэтому неравенство (1.12) в дальнейшем будем записывать в виде

$$f_* \le L(x, w_*) \le f(x) + B(g(x))B^0(w_* \mid \mathbb{E}^m) \qquad \forall x \in P. \tag{1.13}$$

Аналогично, если $B(g(x)) \le 0$ на $\mathbb{E}^n$, то, используя (1.9), получаем

$$B^0(w_* \mid g(P)) \ge B^0(w_* \mid g(\mathbb{E}^n)) \ge B^0(w_* \mid \mathbb{E}^m)$$

и неравенство (1.13) по-прежнему имеет место.

§ 2. Аддитивные т.в.ф.

Будем строить т.в.ф. в виде

$$R(x,y) = A(f(x), y) + B(g(x)), \tag{2.1}$$

где $A(f,y)$ — произвольная непрерывная функция двух аргументов, $B(g(x))$ — строго внешняя штрафная функция, т.е. $B(g(x))$ непрерывна и принимает неотрицательные значения всюду на $\mathbb{E}^n$; кроме того, $B(g(x)) = 0$ тогда и только тогда, когда $x \in X$. Хорошо известным примером строго внешней штрафной функции является

$$B(g(x)) = \|g_+(x)\|_p, \quad p \ge 1, \qquad B^0(w_* \mid \mathbb{E}^m) = \|w_*\|_{p_*}.$$

На введенные функции наложим еще два условия.

Условие А. Функции A и B таковы, что

$$A(f,y) - A(f_*,y) \ge [(f - f_*)/B^0(w_* \mid \mathbb{E}^m)]_- \qquad \forall y \in Y, \quad \forall f \in \mathbb{E}^1. \tag{2.2}$$

Условие Б. Для каждой точки y из множества Y совокупность решений системы

$$A(f(x),y) + B(g(x)) = A(f_*, y), \qquad x \in P, \tag{2.3}$$

совпадает с X_*.

В случае если A дифференцируема по y, Y — открытое множество, из (2.3) следует система

$$A_y(f(x),y) = A_y(f_*,y), \qquad x \in P,$$

решение которой часто проще, чем решение исходной системы (2.3).

Теорема 1. *Пусть в задаче* (1.1) *существует седловая точка* $[x_*,w_*] \in P \times \mathbb{E}^m_+$ *функции Лагранжа. Пусть, кроме того,* $B(g(x))$ *является строго внешней штрафной функцией,* $0 < B^0(w_* \mid \mathbb{E}^m) < +\infty$*, функции* A *и* B *удовлетворяют* **условиям А и Б**. *Тогда задаваемая* (2.1) *функция* $R(x,y)$ *является т.в.ф. для задачи* (1.1) *на множестве* $P \times Y$.

Доказательство. Учитывая, что функция $B(g(x))$ равна нулю на X_*, получаем

$$M(x,y) = A(f(x),y) - A(f_*,y) + B(g(x)). \tag{2.4}$$

Условие (1.4) можно записать в данном случае в виде

$$A(f(x)) + B(g(x)) \geq A(f_*,y) \qquad \forall x \in P, \quad \forall y \in Y. \tag{2.5}$$

По условиям теоремы, $B^0(w_* \mid \mathbb{E}^m) > 0$, поэтому из (1.13) получаем

$$B(g(x)) \geq [f_* - f(x)]/B^0(w_* \mid \mathbb{E}^m) \qquad \forall x \in P.$$

Учитывая, что функция B принимает только неотрицательные значения, уточним эту нижнюю оценку:

$$B(g(x)) \geq [[f_* - f(x)]/B^0(w_* \mid \mathbb{E}^m)]_+ \qquad \forall x \in P. \tag{2.6}$$

Запишем это неравенство в виде

$$B(g(x)) \geq -[[f(x) - f_*]/B^0(w_* \mid \mathbb{E}^m)]_- \qquad \forall x \in P.$$

Применяя его вместе с (2.2) к правой части (2.4), получаем (1.4). **Условие Б** обеспечивает выполнение условия (1.5). Таким образом, функция (2.4) является т.в.ф. на множестве $P \times Y$ и, следовательно, функция (2.1) также является т.в.ф. на том же множестве $P \times Y$. Теорема доказана. □

Условиям теоремы удовлетворяет широкий класс функций. Не предполагается точное знание f_*, поэтому функция A должна строиться так, чтобы при любом f_* выполнялись **условия А и Б**. Из (2.2) следует, что для этого функция $A(f,y)$ должна быть по меньшей мере возрастающей функцией f. Кроме того, ее график должен лежать не ниже выпуклого конуса с началом в точке $N = [f_*, A(f_*,y)]$ и с двумя граничными лучами. Первый луч исходит из точки N и направлен вдоль положительного направления оси f, второй исходит из N и направлен вдоль полупрямой

$$D = (f - f_*)/B^0(w_* \mid \mathbb{E}^m), \qquad f \leq f_*.$$

Заметим также, что утверждение теоремы не изменится, если **условие А** выполнено не для всех $f \in \mathbb{E}^1$, а только для f, принадлежащих образу $f(P)$ множества P при отображении $f(x)$.

Пусть $A(f,y)$ — выпуклая функция по f при каждом $y \in Y$, тогда

$$A(f,y) - A(f_*,y) \ge \xi(f - f_*) \qquad \forall y \in Y, \quad \forall f \in \mathbb{E}^1, \quad \forall \xi \in \partial_f A(f_*,y). \tag{2.7}$$

Здесь $\partial_f A(f_*,y)$ — субдифференциал функции $A(f,y)$ по f в точке f_*. Условие (2.2) будет заведомо выполняться, если

$$\xi(f - f_*) \ge [(f - f_*)/B^0(w_* \mid \mathbb{E}^m)]_- \qquad \forall y \in Y, \quad \forall f \in \mathbb{E}^1, \quad \forall \xi \in \partial_f A(f_*,y).$$

Вместе с условием неубывания функции $A(f,y)$ по f это условие эквивалентно следующему:

$$0 \le \inf_{\xi \in \partial A_f(f_*,y)} \xi \le \sup_{\xi \in \partial A_f(f_*,y)} \xi \le 1/B^0(w_* \mid \mathbb{E}^m). \tag{2.8}$$

Приведенные неравенства можно использовать для определения множества Y, при этом дополнительно должны быть рассмотрены граничные точки Y. Заметим, что если, помимо (2.8), функция $B(g)$ выпукла и не убывает по g, функции $f(x)$ и $g(x)$ выпуклы по x, то вспомогательная функция $R(x,y)$ также выпукла по x.

В качестве примеров т.в.ф., удовлетворяющих условиям теоремы 1 и условиям (2.8) с $P = \mathbb{E}^n$, $Y \subset \mathbb{E}^1$, приведем следующие функции и соответствующие им множества Y:

$$\begin{array}{ll}
R_1(x,y) = y^{-1}f(x) + B(g(x)), & Y_1 = \{y : y > B^0(w_* \mid \mathbb{E}^m)\}, \\
R_2(x,y) = y^{-1}e^{f(x)} + B(g(x)), & Y_2 = \{y : y > B^0(w_* \mid \mathbb{E}^m)e^{f_*}\}, \\
R_3(x,y) = [y - f(x)]^{\alpha} + B(g(x)), & Y_3 = \{y : y > f_* + [-\alpha B^0(w_* \mid \mathbb{E}^m)]^{1/(1-\alpha)}\}, \\
R_4(x,y) = [f(x) - y]_+^{\beta} + B(g(x)), & Y_4 = \{y : f_* \ge y > f_* - [\beta B^0(w_* \mid \mathbb{E}^m)]^{1/(1-\beta)}\},
\end{array}$$

где $\alpha < 0$, $\beta > 1$. В случае R_3 следует считать, что $R_3(x,y) = +\infty$, когда $f(x) \ge y$. Если известно, что функция $f(x) > 0$ для любых $x \in \mathbb{E}^n$, то можно воспользоваться функциями

$$\begin{array}{ll}
R_5(x,y) = y^{-1}\mathrm{sh} f(x) + B(g(x)), & Y_5 = \{y : y > B^0(w_* \mid \mathbb{E}^m)\mathrm{ch} f_*\}, \\
R_6(x,y) = y^{-1}[f(x)]^{\gamma} + B(g(x)), & Y_6 = \{y : y > \gamma f_*^{\gamma-1} B^0(w_* \mid \mathbb{E}^m)\}.
\end{array}$$

Если функция $f(x)$ неположительна на $\mathbb{E}^n$, то можно рассмотреть функцию

$$R_7(x,y) = y^{-1}\mathrm{arctg} f(x) + B(g(x)), \quad Y_7 = \{y : y > B^0(w_* \mid \mathbb{E}^m)/(1 + f_*^2)\}.$$

Функция R_7 будет т.в.ф. и в случае, когда $f(x)$ принимает любые значения, однако множество Y_7 следует заменить на $Y_7 = \{y : y > B^0(w_* \mid \mathbb{E}^m)\}$. Отметим также, что входящие в R_3, R_4, R_6 параметры α, β и γ можно было бы рассматривать как вторые компоненты вектора y.

Первая из приведенных функций — это хорошо известная точная штрафная функция [1, 2, 9, 10]. Функция R_4 использовалась многими авторами для последовательной безусловной минимизации, при этом строилась последовательность векторов y_k, сходящаяся к f_* (см. [9, 11]). Если взять замыкание множеств Y_1 – Y_7, то функции R_1 – R_7 перестают быть т.в.ф., так как для них вместо (1.5) выполнено более слабое условие: из $M(x,y) = 0$ следует, что $X_* \subseteq X(y)$.

Обозначим $X^0 = \{x \in \mathbb{E}^n \mid g(x) < 0\}$ и рассмотрим далее случай, когда $B(g(x))$ является внутренней штрафной функцией, т.е. $B(g(x))$ непрерывна и неположительна на $\mathbb{E}^m$ и $B(g(x)) < 0$ при $x \in X^0$. Если дополнительно предположить, что $B(g(x)) < 0$ тогда и только тогда, когда $g(x) < 0$, то такую функцию будем называть строго внутренней штрафной функцией. В качестве примера внутренних штрафных функций укажем

$$B(g(x)) = -\|g_-(x)\|_p, \qquad -\infty \le p < 1. \tag{2.9}$$

При $p \le 0$ эта функция будет строго внутренней штрафной функцией.

Ниже при рассмотрении внутренних штрафных функций предполагается, что выполнено условие регулярности ограничений, т.е. множество X^0 непусто и его замыкание совпадает с X.

Построим т.в.ф. в виде (2.1), взяв в качестве P допустимое множество X. Условие (2.2) заменим на

$$A(f, y) - A(f_*, y) \ge (f - f_*)/B^0(w_* \mid \mathbb{E}^m_-) \qquad \forall y \in Y, \quad \forall f \ge f_*. \tag{2.10}$$

Теорема 1 в данном случае может быть переформулирована следующим образом.

Теорема 2. *Пусть в задаче* (1.1) *существует седловая точка* $[x_*, w_*]$ *функции Лагранжа. Пусть, кроме того, функция* $B(g(x))$ *в* (2.1) *является внутренней штрафной функцией и* $0 < B^0(w_* \mid \mathbb{E}^m_-) < +\infty$. *Для того чтобы* $R(x, y)$ *была т.в.ф. на* $X \times Y$, *достаточно, чтобы удовлетворялось условие* (2.10) *и всякое допустимое решение* x *системы* (2.3) *было таково, что* $x \in X_*$.

Доказательство. Так как функция Лагранжа имеет седловую точку, $B^0(w_* \mid \mathbb{E}^m_-) > 0$ и $B(g(x)) \le 0$ на X, имеем

$$f_* \le f(x) + B(g(x))B^0(w_* \mid \mathbb{E}^m_-) \le f(x) \qquad \forall x \in X.$$

Отсюда, аналогично (2.6),

$$[f_* - f(x)]/B^0(w_* \mid \mathbb{E}^m_-) \le B(g(x)) \le 0 \qquad \forall x \in X. \tag{2.11}$$

Взяв в левой части произвольную точку $x \in X_*$, получим $B(g(x)) = 0$. Таким образом, всюду на X_* функция $B(g(x))$ принимает равное нулю значение. Поэтому для того, чтобы имело место (1.4), достаточно, чтобы для всех $x \in X$ выполнялось неравенство (2.5). Но, согласно (2.11), выполнение (2.5) обеспечивается (2.10).

Справедливость (1.5) следует из предположения, что всякое допустимое решение $x_* \in \in X$ системы (2.3) таково, что $x_* \in X_*$, поскольку, как было установлено, $B(g(x_*)) = 0$, когда $x_* \in X_*$. Теорема доказана. $\square$

Как видно из доказательства теоремы 2, требование $B^0(w_* \mid \mathbb{E}^m_-) > 0$ приводит к тому, что в точке $x_* \in X_*$ выполнено равенство $B(g(x)) = 0$. Таким образом, в силу условия регулярности ограничений, точка x_* обязательно должна быть в этом случае граничной точкой множества X. Если в качестве $B(g(x))$ взята функция (2.9) при $0 < p < 1$, то $B(g(x))$ является внутренней (но не строго) штрафной функцией и неравенство $B^0(w_* \mid \mid \mathbb{E}^m_-) = \|w_*\|_{p_*} > 0$ имеет место в том и только том случае, когда все компоненты вектора w_* строго положительны. Это означает, что все ограничения в точке x_* одновременно обращаются в нуль, или, другими словами, являются активными. Если (2.9) — строго внутренняя штрафная функция (что имеет место при $p \le 0$), то для справедливости неравенства $\| w_* \|_{p_*} > 0$ достаточно, чтобы активным было хотя бы одно ограничение.

Обратим внимание, что если A — выпуклая функция f при каждом $y \in Y$, то для выполнимости неравенства (2.10) достаточно, чтобы

$$\xi \ge 1/B^0(w_* \mid \mathbb{E}^m_-) \qquad \forall \xi \in \partial_f A(f_*, y), \quad \forall y \in Y.$$

Условиям теоремы 2 удовлетворяют функции R_1 – R_4, однако вместо строго внешней штрафной функции $B(g(x))$ следует взять внутреннюю штрафную функцию (например, (2.9)) и заменить множества Y_i следующими:

$$\begin{aligned}
&Y_1 = \{y : 0 < y < B^0(w_* \mid \mathbb{E}^m_-)\},\\
&Y_2 = \{y : 0 < y < e^{f_*} B^0(w_* \mid \mathbb{E}^m_-)\},\\
&Y_3 = \{y : f_* < y < f_* + [-\alpha B^0(w_* \mid \mathbb{E}^m_-)]^{1/(1-\alpha)}\}, \quad \alpha < 0,\\
&Y_4 = \{y : y < f_* - [\beta B^0(w_* \mid \mathbb{E}^m_-)]^{1/(1-\beta)}\}, \quad \beta > 1.
\end{aligned}$$

Отметим также, что поскольку при предположениях теоремы 2 функция $B(g(x))$ неположительна и на X_* принимает равное нулю значение, то вспомогательная функция $R(x,y) = f(x) - yB(g(x))$ будет также т.в.ф. на $X \times \mathbb{E}^1_+$.

Введем новый класс штрафных функций, обладающих как свойствами строго внешних, так и свойствами строго внутренних штрафных функций. Непрерывная функция $B(g(x))$ называется строго смешанной штрафной функцией, если $B(g(x)) > 0$, когда $x \notin X$, и $B(g(x)) < 0$, когда $x \in X^0$.

В качестве примера строго смешанных штрафных функций приведем следующие две функции:

$$\begin{aligned}
B(g(x)) \quad &= \quad \|g_+(x)\|_p - \|g_-(x)\|_{-p}, \quad 1 < p < +\infty, &(2.12)\\
B(g(x)) \quad &= \quad \max_{1\le i\le m} g^i(x). &(2.13)
\end{aligned}$$

Изменим условия (2.2) и (2.10), положив для любого $y \in Y$

$$A(f,y) - A(f_*,y) \ge \begin{cases} (f - f_*)/B^0(w_* \mid \mathbb{E}^m_-), & \text{если } f \ge f_*,\\ (f - f_*)/B^0(w_* \mid \mathbb{E}^m), & \text{если } f < f_*. \end{cases} \qquad (2.14)$$

Имеет место

Теорема 3. *Пусть в задаче* (1.1) *существует седловая точка* $[x_*, w_*]$ *функции Лагранжа. Пусть, кроме того, функция* $B(g(x))$ *является строго смешанной штрафной функцией и* $0 < B^0(w_* \mid \mathbb{E}^m) < B^0(w_* \mid \mathbb{E}^m_-) < +\infty$. *Для того чтобы* $R(x,y)$ *была т.в.ф. на* $P \times Y$, *достаточно, чтобы удовлетворялись условия* (2.14) *и* **Б**.

Доказательство. Так как $B^0(w_* \mid \mathbb{E}^m_-) > 0$, то, аналогично тому как это делалось при доказательстве теоремы 2, убеждаемся, что $B(g(x_*)) = 0$. Поэтому согласно (2.14) выполнено (1.4). Условие (1.5) следует из (2.3). Теорема доказана. □

Условие (2.14) для гладкой функции оказывается справедливым в том и только том случае, когда $B^0(w_* \mid \mathbb{E}^m) \le B^0(w_* \mid \mathbb{E}^m_-)$. Так, например, функция $R_1(x,y)$ со строго смешанной штрафной функцией $B(g(x))$ является т.в.ф. Соответствующее множество Y_1 для нее

$$Y_1 = \{y : B^0(w_* \mid \mathbb{E}^m) < y < B^0(w_* \mid \mathbb{E}^m_-)\}.$$

В частности, если $B(g(x))$ имеет вид (2.12), получаем

$$Y_1 = \{y : \|w_*\|_{p_*} < y < \|w_*\|_{r_*}\},$$

где $p^{-1} + p_*^{-1} = 1$, $-p^{-1} + r_*^{-1} = 1$. Так как $p_* > 1$, $1 > r_* > 1/2$, то $\|w_*\|_{r_*} > \|w_*\|_{p_*}$. Для функции (2.13) множество Y_1 пусто.

§ 3. Нелинейные т.в.ф.

Предположим теперь, что вспомогательная функция $R(x,y)$ строится в виде

$$R(x,y) = H(A(f(x),y), B(g(x))), \qquad (3.1)$$

где $B(g(x))$ — строго внешняя штрафная функция, $H(t,\tau)$ — непрерывная неубывающая функция двух аргументов, т.е. $H(t_1,\tau_1) \geq H(t,\tau)$ для любых $t_1 \geq t$, $\tau_1 \geq \tau$. Относительно функции $A(f,y)$ предполагается, что она является монотонно возрастающей и выпуклой функцией первого аргумента. Считаем также, что существует такая константа D, что

$$0 < \sup_{y\in Y} \sup_{\xi\in\partial A_f(f_*,y)} \xi \leq D < +\infty. \tag{3.2}$$

Приведем условия, достаточные для того, чтобы функция (3.1) была т.в.ф. для задачи (1.1). Обозначим $N_1 = DB^0(w_* \mid \mathbb{E}^m)$.

Теорема 4. *Пусть в задаче* (1.1) *существует седловая точка* $[x_*, w_*]$ *функции Лагранжа,* $A(f,y)$ *— выпуклая, монотонно возрастающая функция первого аргумента, для которой выполнено* (3.2). *Пусть, кроме того,* $B(g(x))$ *— строго внешняя штрафная функция и* $0 < N_1 < +\infty$. *Тогда если найдется такое множество* $T \subseteq \mathbb{E}^1$, *что* $A(f_*,y) \in T$ *для любых* $y \in Y$ *и*

$$H(t,(t_*-t)_+/N_1) > H(t_*,0) \qquad \forall t_* \in T, \quad \forall t \neq t_*, \tag{3.3}$$

то (3.1) *— т.в.ф. для задачи* (1.1) *на множестве* $P \times Y$.

Доказательство. Возьмем произвольную точку $x \in P$. Так как функция $H(t,\tau)$ не убывает по второй переменной, то, согласно (2.6),

$$R(x,y) = H(A(f(x),y),B(g(x))) \geq H(A(f(x),y),[f_* - f(x)]_+/B^0(w_* \mid \mathbb{E}^m)). \tag{3.4}$$

В силу выпуклости функции $A(f,y)$ по f справедливо неравенство (2.7), из которого с учетом (3.2) получаем

$$A(f_*,y) - A(f,y) \leq D(f_* - f) \qquad \forall f \leq f_*. \tag{3.5}$$

Но $A(f,y)$ — монотонно возрастающая функция по f, поэтому наряду с (3.5) имеет место неравенство

$$(A(f_*,y) - A(f,y))_+ \leq D(f_* - f)_+ \qquad \forall f \in \mathbb{E}^1.$$

Подставляя его в (3.4) и используя условие (3.3), получаем в случае, когда $A(f(x),y) \neq$ $\neq A(f_*,y)$,

$$R(x,y) \geq H(A(f(x),y),[A(f_*,y) - A(f(x),y)]_+/N_1) > H(A(f_*,y),0) = R(x_*,y).$$

Предположим теперь, что $A(f(x),y) = A(f_*,y)$. В этом случае если $x \notin X_*$, то обязательно $x \notin X$, $B(g(x)) > 0$. Поэтому, в силу монотонности функции $H(t,\tau)$ по t и (3.3),

$$\begin{aligned} R(x,y) = H(A(f_*,y),B(g(x))) &\geq H(A(f_*,y) - N_1 B(g(x)),B(g(x))) > \\ &> H(A(f_*,y),0) = R(x_*,y). \end{aligned}$$

Таким образом, в любом случае $R(x,y) > R(x_*,y)$, если $x \notin X_*$, и, следовательно, (3.1) является т.в.ф. на множестве $P \times Y$. Теорема доказана. □

Рассмотрим два важных частных случая функции (3.1). Предположим вначале, что $A(f,y)$ имеет вид

$$A(f,y) = f - y. \tag{3.6}$$

Для этой функции условие (3.2) выполняется для любого $Y \subseteq \mathbb{E}^1$, причем $D = 1$. Поэтому согласно утверждению теоремы 4 если найдется множество $T \subseteq \mathbb{E}^1$, для которого имеет место (3.3), то функция (3.1) с функцией $A(f, y)$, представимой в виде (3.6), является т.в.ф. для задачи (1.1) на множестве $P \times Y$, где $Y = f_* - T$.

Рассмотрим другую функцию:

$$A(f, y) = y^{-1} f. \tag{3.7}$$

В этом случае если выполнено (3.3) при некотором $0 < D < +\infty$, то функция (3.1) с функцией $A(f, y)$ вида (3.7) также является т.в.ф. на $P \times Y$, однако здесь $Y = \{y \geq D^{-1} : : y^{-1} f_* \in T\}$.

Обратим внимание, что если функция $H(t, \tau)$ линейная и имеет вид $H(t, \tau) = ct + \tau$, то для выполнения условия (3.3) необходимо и достаточно, чтобы $c < 1/B^0(w_* \mid \mathbb{E}^m)$. Множество T в данном случае совпадает со всей действительной прямой $\mathbb{E}^1$. Поэтому если перейти от y к переменной $\bar{y} = y/c$, то получаем, что функция $R(x, \bar{y}) = H(y^{-1} f(x), B(g(x))) = = \bar{y}^{-1} f(x) + B(g(x))$ — т.в.ф. в том и только том случае, когда $\bar{y} > B^0(w_* \mid \mathbb{E}^m)$, что полностью согласуется с полученной выше формулой для Y_1.

Обозначим через $K_1(t_*)$ конус:

$$K_1(t_*) = \{[t, \tau] \in \mathbb{E}^2 | \tau \geq (t_* - t)_+ / N_1\}.$$

Условие (3.3) геометрически означает, что для любого $t_* \in T$ линия уровня функции $H(t, \tau)$, соответствующая значению $H(t_*, 0)$, не должна пересекать конус $K_1(t_*)$, за исключением, разумеется, самой точки $[t_*, 0]$.

Условие (3.3) сравнительно просто проверяется, когда $H(t, \tau)$ — квазивыпуклая функция. Действительно, пусть $H(t, \tau)$ — непрерывно дифференцируемая квазивыпуклая функция на $\mathbb{E}^2$ и можно указать такое множество $T_1 \subseteq \mathbb{E}^1$, что

$$H_t(t_*, 0) > 0, \qquad H_\tau(t_*, 0)/H_t(t_*, 0) > N_1 \qquad \forall t_* \in T_1. \tag{3.8}$$

Тогда функция одного переменного $\varphi(t) = H(t, (t_* - t)_+ / N_1)$ также является квазивыпуклой справа и слева от точки t_* и имеет в этой точке обе односторонние производные, причем, согласно (3.8), $\varphi_t^+(t_*) = H_t(t_*, 0) > 0$, $\varphi_t^-(t_*) = H_t(t_*, 0) - H_\tau(t_*, 0)/N_1 < 0$. Поэтому функция $\varphi(t)$ достигает в точке t_* своего строгого минимума по t и, следовательно, справедливо неравенство (3.3). К множеству T_1 можно добавить точки множества

$$T_2 = \{t_* \notin T_1 | t_* = \lim_{t_k \to +t} t_k, \ \ t_k \in T_1\}.$$

Тогда, если имеет место (3.8), для любого $t_* \in T = T_1 \cup T_2$ выполнено (3.3). Отметим также, что условия (3.8) сохраняются и в том случае, когда функция $H(t, \tau)$ дифференцируема только в точках некоторого множества, содержащего отрезок $\{z \in \mathbb{E}^2 \mid z^{(1)} \in T_1, \ \ z^{(2)} = 0\}$.

Приведем примеры т.в.ф., множества Y для которых могут быть найдены с помощью теоремы 4 и условий (3.8). Пусть

$$H(t, \tau) = \begin{cases} t_-/(1 - t_-\tau), & \text{если } \ 1 - t_-\tau > 0, \\ -\infty & \text{в противном случае.} \end{cases}$$

Эта функция квазивыпукла на $\mathbb{E}^2$. При $t < 0$ имеем $H_t(t, 0) = 1$, $H_\tau(t, 0) = t_-^2$. Условия (3.8) выполняются, если $t < -[B^0(w_* \mid \mathbb{E}^m)]^{1/2}$. Таким образом, $T = (-\infty, -[B^0(w_* \mid \mathbb{E}^m)]^{1/2})$ и, следовательно, вспомогательная функция

$$R_8(x, y) = [f(x) - y]_- / \{1 - [f(x) - y]_- B(g(x))\}$$

является т.в.ф. на множестве $\mathbb{E}^n \times Y_8$, где $Y_8 = \{y \in \mathbb{E}^1 \mid y > f_* + [B^0(w_* \mid \mathbb{E}^m)]^{1/2}\}$.

Если $f_* > 0$, то, применяя функцию $H(t,\tau) = \tau + (\tau^2 + 4t_+^3)^{1/2}$, приходим к следующей т.в.ф.:

$$R_9(x,y) = B(g(x)) + \{[B(g(x))]^2 + 4y^{-3}[f_+(x)]^3\}^{1/2}$$

для задачи (1.1) на множестве $\mathbb{E}^n \times Y_9$, где

$$Y_9 = \{y : y > \max\{1.9 f_* [B^0(w_* \mid \mathbb{E}^m)]^2\}\}.$$

Рассмотрим достаточные условия в случае, когда функция $B(g(x))$ в (3.1) является внутренней штрафной функцией, предполагая по-прежнему, что функция $H(t,\tau)$ неубывающая на $\mathbb{E}^2$, а функция $A(f,y)$ — выпуклая, монотонно возрастающая функция первого аргумента. Вместо (3.2) на функцию $A(f,y)$ наложим условие существования такой константы C, что

$$0 < C \le \inf_{y\in Y} \inf_{\xi \in \partial A_f(f_*,y)} \xi < +\infty. \tag{3.9}$$

Обозначим $N_2 = CB^0(w_* \mid \mathbb{E}^m_-)$.

Теорема 5. *Пусть в задаче* (1.1) *существует седловая точка* $[x_*, w_*]$ *функции Лагранжа,* $A(f,y)$ *— выпуклая, монотонно возрастающая функция первого аргумента, для которой выполнено* (3.9). *Пусть, кроме того,* $B(g(x))$ *— внутренняя штрафная функция,* $0 < N_2 < +\infty$. *Тогда если найдется такое множество* $T \subseteq \mathbb{E}^1$, *что* $A(f_*,y) \in T$ *для любых* $y \in Y$ *и*

$$H(t, (t_* - t)/N_2) > H(t_*, 0) \qquad \forall t_* \in T, \quad \forall t > t_*, \tag{3.10}$$

то (3.1) *— т.в.ф. для задачи* (1.1) *на множестве* $X \times Y$.

Доказательство. Возьмем произвольную точку $x \in X$. В случае когда $x \notin X_*$, используя неравенства (2.7), (3.9) и (3.10), получаем

$$\begin{aligned} R(x,y) &= H(A(f(x),y), B(g(x))) \ge \\ &\ge H(A(f(x),y), [f_* - f(x)]/B^0(w_* \mid \mathbb{E}^m_-)) \ge \\ &\ge H(A(f(x),y), [A(f_*,y) - A(f(x),y)]/N_2) > \\ &> H(A(f_*,y), 0) = R(x_*,y). \end{aligned}$$

Если $x \in X_*$, то, согласно (2.11), $B(g(x)) = 0$. Поэтому $R(x,y) = H(A(f(x),y),0) = R(x_*,y)$. Таким образом, $R(x,y)$ является т.в.ф. на $X \times Y$. Теорема доказана. $\square$

Если функция $A(f,y)$ представима в виде (3.6) или (3.7), то, зная T в условии (3.10), можно указать конкретный вид множества Y. Так, для первой функции получаем, что $Y = f_* - T$, для второй — что $Y = \{y \le C^{-1} : y^{-1} f_* \in T\}$.

Условие (3.10) геометрически означает, что для любого $t_* \in T$ линия уровня функции $H(t,\tau)$, соответствующая значению $H(t_*,0)$, при $t > t_*$ не может иметь общих точек с конусом

$$K_2(t_*) = \{[t,\tau] \in \mathbb{E}^2 \mid \tau \ge (t_* - t)/N_2, \;\; t \ge t_*\}.$$

Можно привести условия, аналогичные условиям (3.8), упрощающие проверку неравенства (3.10) и дающие возможность сравнительно просто устанавливать вид множества T. Предположим, что $H(t,\tau)$ — непрерывно дифференцируемая квазивыпуклая функция на $\mathbb{E}^2$ и можно указать такое множество $T_1 \subseteq \mathbb{E}^1$, что для любых $t_* \in T_1$ имеет место первое неравенство (3.8) и

$$H_\tau(t_*,0)/H_t(t_*,0) < N_2.$$

Тогда для всех $t_* \in T = T_1 \cup T_2$ выполнено неравенство (3.10), где

$$T_2 = \{t_* \notin T_1 \mid t_* = \lim_{t_k \to -t} t_k, \ \ t_k \in T_1\}.$$

В качестве примера использования теоремы 5 приведем рассмотренные ранее функции $R_8(x,y)$ и $R_9(x,y)$, в которых $B(g(x))$ — внутренняя штрафная функция. Обе эти функции — т.в.ф. соответственно на множествах $X \times Y_8$ и $X \times Y_9$, однако множества Y_8 и Y_9 в данном случае принимают следующий вид:

$$\begin{aligned} Y_8 &= \{y \in \mathbb{E}^1 \mid f_* \le y < f_* + [B^0(w_* \mid \mathbb{E}^m_-)]^{1/2}\}, \\ Y_9 &= \{y \in \mathbb{E}^1 \mid 0 < y < \min[1.9 f_* [B^0(w_* \mid \mathbb{E}^m_-)]^2]\}. \end{aligned}$$

Отметим, что утверждения теорем 4 и 5 не изменятся, если условия (3.3), (3.10) выполняются не для всех $t \in \mathbb{E}^1$, а только для тех, которые принадлежат некоторому подмножеству $\mathbb{E}^1$, содержащему образ множества $P \times Y$ при отображении $A(f(x), y)$. С учетом этого свойства в [8] показано, что функция

$$R_{10}(x,y) = (f(x) - y)_+ / \{1 + [f(x) - y]_+ B(g(x))\},$$

где $B(g(x))$ — внутренняя штрафная функция, также является т.в.ф. для задачи (1.1) на множестве $X \times Y_{10}$. При этом

$$Y_{10} = \{y : f_* - ([(f^* - f_*)^2 + 4B^0(w_* \mid \mathbb{E}^m_-)]^{1/2} - (f^* - f_*))/2 < y \le f_*\}, \quad f^* = \sup_{x \in X} f(x).$$

Рассмотрим т.в.ф. вида (3.1) со строго смешанными штрафными функциями $B(g(x))$. Условия на функцию $H(t,\tau)$ в данном случае получаются объединением соответствующих условий для строго внешних и внутренних штрафных функций.

Теорема 6. *Пусть в задаче* (1.1) *существует седловая точка* $[x_*, w_*]$ *функции Лагранжа,* $A(f,y)$ *— выпуклая, монотонно возрастающая функция первого аргумента, для которой выполнены условия* (3.2) *и* (3.9). *Пусть, кроме того,* $B(g(x))$ *— строго смешанная штрафная функция и* $0 < N_1 < +\infty$, $0 < N_2 < +\infty$. *Тогда если найдется такое множество* $T \subseteq \mathbb{E}^1$, *что* $A(f_*, y) \in T$ *для любых* $y \in Y$ *и*

$$H(t_*, 0) < \begin{cases} H(t, (t_* - t)/N_1), & t < t_*, \\ H(t, (t_* - t)/N_2), & t > t_*, \end{cases} \qquad \forall t_* \in T,$$

то (3.1) *— т.в.ф. для задачи* (1.1) *на множестве* $P \times Y$.

Доказательство теоремы опускается, поскольку легко может быть получено комбинированием соответствующих рассуждений, проведенных при доказательстве теорем 4 и 5.

Функция $R_8(x,y)$ со строго смешанной штрафной функцией $B(g(x))$ остается т.в.ф. для задачи (1.1) на множестве $P \times Y$, если $B(g)$ такова, что $B^0(w_* \mid \mathbb{E}^m) < B^0(w_* \mid \mathbb{E}^m_-)$. Множество Y_8 в данном случае имеет вид

$$Y_8 = \{y : f_* + [B^0(w_* \mid \mathbb{E}^m)]^{1/2} < y < f_* + [B^0(w_* \mid \mathbb{E}^m_-)]^{1/2}\}.$$

§ 4. Точные модифицированные функции Лагранжа

Перепишем правую часть неравенства (1.11) в виде

$$L(x_*, w_*) \le L(x, w) + \langle g(x), w_* - w \rangle.$$

Если $B^0(w_* - w \mid g(P)) > 0$, то, применяя неравенство Минковского–Малера (1.7), отсюда получаем

$$B(g(x)) \ge [L(x_*, w_*) - L(x, w)]/B^0(w_* - w \mid g(P)). \tag{4.1}$$

Неравенство (4.1) позволяет построить целый класс т.в.ф., основанный на использовании функции Лагранжа $L(x, w)$. Положим

$$R(x, y) = A(L(x, w), v) + B(g(x)). \tag{4.2}$$

Здесь $y = [w, v] \in Y \subseteq \mathbb{E}^m_+ \times \mathbb{E}^1$.

Обозначим

$$W_Y = \{w \in \mathbb{E}^m_+ \mid \exists v \in \mathbb{E}^1,\ \ [w, v] \in Y\}, \quad V_w = \{v \in \mathbb{E}^1 \mid [w, v] \in Y\}.$$

Множество W_Y — проекция Y на $\mathbb{E}^m_+$, множество V_w — сечение Y при фиксированном $w \in W_Y$.

Пусть функции A, B и множества P, Y таковы, что справедливы следующие условия.

Условие В.

$$A(L, v) - A(L_*, v) \ge (L - L_*)/B^0(w_* - w \mid g(P)) \quad \forall L \in \mathbb{E}^1,\ \forall w \in W_Y,\ \forall v \in V_w,$$

где $L_* = L(x_*, w_*) = f_*$.

Условие Г. Для каждой точки $y \in Y$ совокупность решений системы

$$A(L(x, w), v) + B(g(x)) = A(L(x_*, w_*), v), \quad x \in P,$$

совпадает с X_*.

К этим двум условиям добавим требование

$$L(x_*, w) = L(x_*, w_*) \qquad \forall w \in W_Y. \tag{4.3}$$

Заметим, что если $w_* \ne 0$, то для множества

$$W_Y = W_* = \{0 \le w \le w_* \mid w^i < w^i_*,\ \text{если}\ w^i_* > 0\} \tag{4.4}$$

условие (4.3) заведомо выполняется.

Теорема 7. *Пусть в задаче* (1.1) *существует седловая точка* $[x_*, w_*]$ *функции Лагранжа, выполнены* **условия В, Г** *и* (4.3). *Пусть, кроме того,* $B(g(x))$ *— строго внешняя или внутренняя штрафная функция (в последнем случае* $P \subseteq X$*) и* $0 < B^0(w_* - w \mid g(P)) < +\infty$ *для любых* $w \in W_Y$. *Тогда функция* (4.2) *— т.в.ф. для задачи* (1.1) *на множестве* $P \times Y$.

Доказательство этой теоремы почти дословно повторяет доказательства теорем 1 и 2. Заметим также, что в случае внешней штрафной функции $B(g(x))$ **условие В** можно ослабить, потребовав выполнения только неравенства

$$A(L, v) - A(L_*, v) \ge [(L - L_*)/B^0(w_* - w \mid g(P))]_- \quad \forall L \in \mathbb{E}^1,\ \forall w \in W_Y,\ \forall v \in V_w.$$

В качестве примеров функции (4.2), являющихся т.в.ф. на множестве $\mathbb{E}^n \times Y$, где $Y = \{[w,v] \mid w \in W_*,\ v \in V_w\}$, можно взять функции R_1 – R_4, заменив в них $f(x)$ на $L(x,w)$. Множества V_w для этих функций строятся так же, как соответствующие множества Y для функций R_1 – R_4. Поэтому приведем эти т.в.ф. с указанием множеств V_w только для случая, когда $B(g(x))$ — строго внешняя штрафная функция:

$$\begin{array}{rclrcl}
R_{11}(x,y) & = & v^{-1}L(x,w) + B(g(x)), & V_w & = & \{v : v > B^0(w_* - w \mid \mathbb{E}^m)\},\\
R_{12}(x,y) & = & v^{-1}e^{L(x,w)} + B(g(x)), & V_w & = & \{v : v > B^0(w_* - w \mid \mathbb{E}^m)e^{f_*}\},\\
R_{13}(x,y) & = & [v - L(x,w)]_+^{\alpha} + B(g(x)), & V_w & = & \{v : v > f_* + [-\alpha B^0(w_* - w \mid \mathbb{E}^m)]^{1/(1-\alpha)}\},\\
R_{14}(x,y) & = & [L(x,w) - v]_+^{\beta} + B(g(x)), & V_w & = & \{v : f_* \geq v > f_* - [\beta B^0(w_* - w \mid \mathbb{E}^m)]^{1/(1-\beta)}\}.
\end{array}$$

Отсюда видно, что при приближении w к w_* соответствующие множества V_w расширяются. Для т.в.ф. с внутренними штрафными функциями они, напротив, сужаются. Функция R_{11} была предложена в [3].

Рассмотрим теперь нелинейные т.в.ф., составленные с применением функции Лагранжа:

$$R(x,y) = H(A(L(x,w),v), B(g(x))), \tag{4.5}$$

где, как и прежде, $H(t,\tau)$ — непрерывная неубывающая функция двух переменных, $A(L,v)$ — монотонно возрастающая выпуклая функция первого аргумента. Предположим также, что функция A и множество Y таковы, что для любого $w \in W_Y$ существуют константы $C(w)$ и $D(w)$, для которых

$$0 < \sup_{y \in Y} \sup_{\xi \in \partial A_L(L_*,v)} \xi \leq D(w) < +\infty, \tag{4.6}$$

$$0 < C(w) < \sup_{y \in Y} \sup_{\xi \in \partial A_L(L_*,v)} \xi < +\infty. \tag{4.7}$$

Кроме того, считаем, что

$$L(x,w) > L(x_*,w_*) = L_* \qquad \forall w \in W_Y, \quad \forall x \in X \backslash X_*, \tag{4.8}$$

а вместо (3.3), (3.10) на функцию $H(t,\tau)$ наложены условия: для каждого $w \in W_Y$ существует такое множество $T(w) \subseteq \mathbb{E}^1$, что $A(L_*,v) \in T(w)$ для всех $v \in V_w$ и для любых $t_* \in T(w)$ будет

$$H(t,(t_* - t)_+/D(w)B^0(w_* - w \mid \mathbb{E}^m)) > H(t_*,0) \qquad \forall t \neq t_*, \tag{4.9}$$

$$H(t,(t_* - t)/C(w)B^0(w_* - w \mid \mathbb{E}^m_-)) > H(t_*,0) \qquad \forall t > t_*. \tag{4.10}$$

Отметим, что если множество Y таково, что соответствующее множество W_Y имеет вид (4.4), то при выполнении (4.3) условие (4.8) также выполнено. Действительно, если $x \in X$ и $f(x) > f_*$, то в случае, когда $\langle g(x), w\rangle = 0$, получаем $L(x,w) = f(x) > f_* = L_*$. Если же $\langle g(x), w\rangle < 0$, то $\langle g(x), w\rangle > \langle g(x), w_*\rangle$ и поэтому $L(x,w) > L(x,w_*) \geq L_*$.

Теорема 8. *Пусть в задаче* (1.1) *существует седловая точка* $[x_*, w_*]$ *функции Лагранжа,* $A(L,v)$ *— выпуклая, монотонно возрастающая функция первого аргумента, имеет место* (4.3), (4.4). *Тогда если* $B(g(x))$ *— строго внешняя штрафная функция,* $0 < B^0(w_* - w \mid \mathbb{E}^m) < +\infty$ *для любого* $w \in W_Y$ *и выполнено* (4.6), (4.9), *то* (4.5) *— т.в.ф. для задачи* (1.1) *на множестве* $P \times Y$. *Если* $B(g(x))$ *— внутренняя штрафная функция,* $0 < B^0(w_* - w \mid \mathbb{E}^m_-) < +\infty$ *для любого* $w \in W_Y$ *и выполнено* (4.7), (4.10), *то* (4.5) *— т.в.ф. для задачи* (1.1) *на множестве* $X \times Y$.

Доказательство этой теоремы аналогично доказательству теорем 4 и 5.

Предположим, что функция $A(L,v)$ имеет вид $A(L,v) = L - v$. Тогда из утверждения теоремы 8 следует, что $V_w = f_* - T(w)$ для любого $w \in W_Y$. На основании этого свойства приходим к следующей т.в.ф.:

$$R_{15}(x,y) = [L(x,w) - v]_-/\{1 - [L(x,w) - v]_- B(g(x))\}.$$

Если $B(g(x))$ — строго внешняя штрафная функция, то R_{15} — т.в.ф. на множестве $P \times Y_{15}$, где

$$Y_{15} = \{[w,v] : w \in W_*,\ \ v > f_* + [B^0(w_* - w \mid \mathbb{E}^m)]^{1/2}\}. \tag{4.11}$$

В случае когда $B(g(x))$ — внутренняя штрафная функция, R_{15} будет т.в.ф. на множестве $X \times Y_{15}$, однако множество (4.11) следует теперь заменить на

$$Y_{15} = \{[w,v] : w \in W_*,\ \ f_* \le v < f_* + [B^0(w_* - w \mid \mathbb{E}^m_-)]^{1/2}\}.$$

Можно рассмотреть также т.в.ф. вида (4.5) со строго смешанными штрафными функциями $B(g(x))$. Достаточные условия при этом остаются прежними, за исключением условия (4.9), выполнение которого требуется только при $t < t_*$.

В заключение отметим, что функции (2.1), (3.1) являются частными случаями функций (4.2), (4.5). Чтобы их получить, достаточно в (4.2), (4.5) положить множители $w = 0$.

Отметим также, что если в задаче (1.1) существует более чем одна седловая точка $[x_*, w_*]$, то при оценке множеств Y с целью их расширения следует взять верхние или нижние границы по всем возможным значениям множителей Лагранжа w_*.

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

1. *Еремин И.И.* Метод “штрафов” в выпуклом программировании // Докл. АН СССР, 1967. Т. 173. N° 4. С. 748–751.

2. *Zangwill W.* Non-linear programming via penalty function // Manag. Sci., 1967. Pp. 344–358.

3. *Скарин В.Д.* Об одной модификации метода штрафных функций в выпуклом программировании // Нелинейная оптимизация и приложения в планировании. Свердловск: УНЦ АН СССР, 1973. С. 51–62.

4. *Evans J.P., Gould F.J., Tolle J.W.* Exact penalty functions in nonlinear programming // Math. Program., 1973. V.4. P. 72–97.

5. *Голиков А.И.* Метод внешних центров и негладкие штрафные функции // Системы программного обеспечения решения задач оптимального планирования. VII Всес. симпозиум. М., 1982. С. 125–126.

6. *Евтушенко Ю.Г.* Оценки точности в методах штрафных функций // Пробл. прикл. матем. и информатики. М.: Наука, 1987. С. 199–208.

7. *Жадан В.Г.* О некоторых оценках коэффициента штрафа в методах точных штрафных функций // ЖВМ и МФ, 1984. Т.24. N° 8. С. 1164–1171.

8. *Евтушенко Ю.Г., Жадан В.Г.* К вопросу о систематизации численных методов нелинейного программирования. Методы последовательной безусловной минимизации // Сообщ. по прикл. матем. М.: ВЦ АН СССР, 1988.

9. *Евтушенко Ю.Г.* Методы решения экстремальных задач и их применение в системах оптимизации. М.: Наука, 1982.

10. *Han S.-P., Mangasarian O.L.* Exact penalty functions in nonlinear programming // Math. Program., 1979. V.17. P. 251–269.

11. *Morrison D.D.* Optimization by least squares // SIAM J. Numer. Analys., 1968. V.5. N° 1. P. 83–88.