ТОЧНЫЕ ШТРАФНЫЕ ФУНКЦИИ В НЕГЛАДКОЙ ОПТИМИЗАЦИИ*

М. В. Долгополик

maxim.dolgopolik@gmail.com

8 мая 2014 г.

Аннотация. В докладе приводится обзор некоторых результатов по теории точных штрафных функций из [1, 2].

1°. Определение штрафной функции. Рассмотрим экстремальную задачу вида

$$f(x) \to \inf_{x \in \Omega}, \tag{1}$$

где Ω — некоторое непустое подмножество пространства $\mathbb{R}^n$, а функция f определена на $\mathbb{R}^n$ и принимает вещественные значения. Множество Ω называется множеством *допустимых планов* задачи (1). Везде далее мы будем предполагать, что решение задачи (1) существует.

Пусть множество Ω задано в виде

$$\Omega = \{x \in \mathbb{R}^n \mid \varphi(x) = 0\},$$

где $\varphi\colon \mathbb{R}^n \to [0, +\infty)$ некоторая неотрицательная функция. Отметим, что любое множество Ω может быть задано в таком виде. Например, можно положить

$$\varphi(x) = \begin{cases} 1, & \text{если } x \notin \Omega, \\ 0, & \text{если } x \in \Omega. \end{cases}$$

Для любого неотрицательного λ введём функцию

$$F_\lambda(x) = f(x) + \lambda\varphi(x) \quad \forall\, x \in \mathbb{R}^n,$$

которая называется *штрафной функцией* (для заданных f и φ), а число λ называется *штрафным параметром.* Штрафная функция F_λ совпадает с функцией f на множестве Ω, а для любого $x \notin \Omega$ значение функции $F_\lambda(x)$ больше, чем значение функции $f(x)$.

*Семинар по конструктивному негладкому анализу и недифференцируемой оптимизации «CNSA & NDO»: `http://www.apmath.spbu.ru/cnsa/`

Основная идея теории точных штрафных функций заключается в том, что вместо задачи с ограничениями (1), рассматривается задача безусловной минимизации штрафной функции

$$F_\lambda(x) \to \inf_{x \in \mathbb{R}^n} \tag{2}$$

и исследуется взаимосвязь между решениями исходной задачи (1) и задачи (2). Ниже будет показано, что при некоторых предположениях относительно функций f и φ, найдётся такое $\lambda^* \geqslant 0$, что для любого $\lambda > \lambda^*$ всякое решение задачи (2) является и решением задачи (1). Таким образом, теория точных штрафных функций позволяет сводить задачи условной оптимизации к задачам оптимизации без ограничений.

2°. Свойства штрафной функции. Изучим поведение штрафной функции F_λ при изменении параметра λ. Очевидно, что если $\mu \geqslant \lambda \geqslant 0$, то

$$F_\lambda(x) \leqslant F_\mu(x) \quad \forall\, x \in \mathbb{R}^n, \quad \inf_{x \in \mathbb{R}^n} F_\lambda(x) \leqslant \inf_{x \in \mathbb{R}^n} F_\mu(x), \tag{3}$$

т. е. значения штрафной функции F_λ не убывают по λ при фиксированном x. Более того, если $\mu > \lambda$, то

$$F_\lambda(x) < F_\mu(x) \quad \forall\, x \in \mathbb{R}^n \setminus \Omega.$$

В следующем утверждении сформулировано несколько важных свойств штрафных функций.

ПРЕДЛОЖЕНИЕ 1. *Пусть существует $\lambda_0 \geqslant 0$ такое, что штрафная функция F_λ достигает глобального минимума для любого $\lambda \geqslant \lambda_0$. Тогда для всех $\mu > \lambda \geqslant \lambda_0$ и для любых*

$$x_\lambda \in \operatorname*{arg\,min}_{x \in \mathbb{R}^n} F_\lambda(x), \qquad x_\mu \in \operatorname*{arg\,min}_{x \in \mathbb{R}^n} F_\mu(x)$$

справедливы следующие утверждения:

1) *$f(x_\mu) \geqslant f(x_\lambda)$ и $\varphi(x_\mu) \leqslant \varphi(x_\lambda)$;*

2) *если $\varphi(x_\lambda) = 0$, то $\varphi(x_\mu) = 0$ и x_λ (как и x_μ) является точкой глобального минимума функции f на Ω;*

3) *если $F_\lambda(x_\lambda) = F_\mu(x_\mu)$, то для любого $\nu > \lambda$ и любой точки глобального минимума x_ν штрафной функции F_ν будет $F_\lambda(x_\lambda) = F_\nu(x_\nu)$ и $\varphi(x_\nu) = 0$.*

Доказательство. 1) По определению точек x_λ и x_μ имеем

$$F_\lambda(x_\lambda) = f(x_\lambda) + \lambda\varphi(x_\lambda) \leqslant F_\lambda(x_\mu) = f(x_\mu) + \lambda\varphi(x_\mu) \tag{4}$$

и

$$-F_\mu(x_\lambda) = -f(x_\lambda) - \mu\varphi(x_\lambda) \leqslant -F_\mu(x_\mu) = -f(x_\mu) - \mu\varphi(x_\mu).$$

Прибавляя к первому неравенству второе, получаем

$$(\lambda - \mu)\varphi(x_\lambda) \leqslant (\lambda - \mu)\varphi(x_\mu),$$

откуда

$$\varphi(x_\lambda) \geqslant \varphi(x_\mu),$$

так как $\lambda < \mu$. Следовательно, учитывая (4),

$$f(x_\lambda) \leqslant f(x_\mu) + \lambda(\varphi(x_\mu) - \varphi(x_\lambda)) \leqslant f(x_\mu).$$

2) По предыдущему свойству получаем, что $0 \leqslant \varphi(x_\mu) \leqslant \varphi(x_\lambda) = 0$, то есть $\varphi(x_\mu) = 0$. Откуда следует, что $x_\lambda, x_\mu \in \Omega$.

Учитывая, что $x_\lambda \in \Omega$ является точкой глобального минимума штрафной функции F_λ, получаем, что

$$f(x_\lambda) = F_\lambda(x_\lambda) \leqslant F_\lambda(x) = f(x) \quad \forall\, x \in \Omega,$$

то есть x_λ является решением задачи (1).

3) Покажем сначала, что $\varphi(x_\mu) = 0$. Действительно, предположим, что $\varphi(x_\mu) > 0$. Тогда

$$F_\mu(x_\mu) = f(x_\mu) + \mu\varphi(x_\mu) = F_\lambda(x_\mu) + (\mu - \lambda)\varphi(x_\mu) > F_\lambda(x_\mu),$$

поскольку $\mu > \lambda$ и $\varphi(x_\mu) > 0$. Следовательно

$$F_\mu(x_\mu) > F_\lambda(x_\mu) \geqslant F_\lambda(x_\lambda),$$

что противоречит предположению. Значит $\varphi(x_\mu) = 0$.

Зафиксируем произвольное $\nu > \lambda$ и точку глобального минимума x_ν штрафной функции F_ν. Если $\mu \geqslant \nu$, то, учитывая что штрафная функция не убывает по λ (см. (3)), получаем

$$F_\mu(x_\mu) \geqslant F_\nu(x_\nu) \geqslant F_\lambda(x_\lambda),$$

откуда $F_\nu(x_\nu) = F_\lambda(x_\lambda)$ и следовательно $\varphi(x_\nu) = 0$.

Пусть теперь $\nu > \mu$ и предположим, что $F_\nu(x_\nu) \neq F_\lambda(x_\lambda)$. Тогда с учётом того, что штрафная функция не убывает по λ имеем

$$F_\nu(x_\nu) > F_\lambda(x_\lambda) = F_\mu(x_\mu) = f(x_\mu) + \mu\varphi(x_\mu) = f(x_\mu), \tag{5}$$

так как $\varphi(x_\mu) = 0$. С другой стороны, опять воспользовавшись равенством $\varphi(x_\mu) = 0$ и тем фактом, что x_ν это точка глобального минимума функции F_ν, получим

$$F_\nu(x_\nu) \leqslant F_\nu(x_\mu) = f(x_\mu) = F_\mu(x_\mu),$$

что противоречит (5). Значит $F_\nu(x_\nu) = F_\lambda(x_\lambda)$ и $\varphi(x_\nu) = 0$. □

СЛЕДСТВИЕ 1. *Пусть существует $\lambda > 0$ такое, что*

$$\inf_{x\in\mathbb{R}^n} F_\lambda(x) = \inf_{x\in\Omega} f(x)$$

(*в частности, можно предполагать, что существует точка глобального минимума x_λ функции F_λ такая, что $\varphi(x_\lambda) = 0$*). *Тогда для любого $\mu > \lambda$ справедливо равенство*

$$\arg\min_{x\in\mathbb{R}^n} F_\mu(x) = \arg\min_{x\in\Omega} f(x),$$

то есть множество решений задачи

$$F_\mu(x) \to \inf_{x\in\mathbb{R}^n}$$

непусто и совпадает с множеством решений исходной задачи (1).

Доказательство. Зафиксируем произвольное $\mu > \lambda$. Поскольку штрафная функция F_λ не убывает по λ, то

$$\inf_{x\in\Omega} f(x) = \inf_{x\in\mathbb{R}^n} F_\lambda(x) \leqslant \inf_{x\in\mathbb{R}^n} F_\mu(x) \leqslant \inf_{x\in\Omega} F_\mu(x) = \inf_{x\in\Omega} f(x),$$

откуда

$$\inf_{x\in\mathbb{R}^n} F_\lambda(x) = \inf_{x\in\mathbb{R}^n} F_\mu(x) = \inf_{x\in\Omega} f(x). \tag{6}$$

Пусть $y \in \Omega$ — некоторое решение задачи (1), которое существует по предположению. Тогда из (6) следует, что

$$\inf_{x\in\mathbb{R}^n} F_\mu(x) = \inf_{x\in\Omega} f(x) = f(y) = F_\mu(y),$$

поскольку $\varphi(y) = 0$. Значит y — это точка глобального минимума функции F_μ (как и функции F_λ) и справедливо включение

$$\arg\min_{x\in\Omega} f(x) \subset \arg\min_{x\in\mathbb{R}^n} F_\mu(x).$$

Справедливость обратного включения непосредственно вытекает из пункта 3 предыдущего предложения, равенства (6) и того факта, что $y \in \Omega$ — это точка глобального минимума функции F_λ. □

Укажем ещё одно полезное свойство штрафных функций.

ПРЕДЛОЖЕНИЕ 2. *Пусть существует* $\lambda_0 \geqslant 0$ *такое, что штрафная функция* F_λ *достигает глобального минимума для любого* $\lambda \geqslant \lambda_0$. *Тогда для любых* $x_\lambda \in \arg\min\limits_{x \in \mathbb{R}^n} F_\lambda(x)$, $\lambda \geqslant \lambda_0$ *будет*

$$\varphi(x_\lambda) \to 0 \text{ при } \lambda \to \infty. \tag{7}$$

Доказательство. Предположим, что (7) не выполнено, тогда существуют $m > 0$ и возрастающая последовательность λ_k такие, что $\lambda_k \to \infty$ при $k \to \infty$ и $\varphi(x_{\lambda_k}) \geqslant m$ для всех $k \in \mathbb{N}$. Из пункта 1 предложения 1 вытекает, что

$$f(x_{\lambda_k}) \geqslant f(x_{\lambda_1}) \quad \forall\, k \in \mathbb{N}.$$

Следовательно,

$$F_{\lambda_k}(x_{\lambda_k}) = f(x_{\lambda_k}) + \lambda_k \varphi(x_{\lambda_k}) \geqslant f(x_{\lambda_1}) + \lambda_k m,$$

откуда

$$F_{\lambda_k}(x_{\lambda_k}) \xrightarrow[k \to \infty]{} \infty, \tag{8}$$

поскольку $\lambda_k \to \infty$ при $k \to \infty$. С другой стороны, для любого $x_0 \in \Omega$ будет

$$F_{\lambda_k}(x_{\lambda_k}) \leqslant F_{\lambda_k}(x_0) = f(x_0) < +\infty \quad \forall\, k \in \mathbb{N},$$

что противоречит (8). □

3°. Точные штрафные функции. Следствие 1 мотивирует нас дать следующее определение.

ОПРЕДЕЛЕНИЕ 1. Штрафная функция $F_\lambda = f + \lambda\varphi$ называется *точной штрафной*, если существует $\lambda^* \geqslant 0$ такое, что для любого $\lambda > \lambda^*$ множество точек глобального минимума функции F_λ совпадает с множеством решений задачи

$$f(x) \to \inf_{x \in \Omega}.$$

В данном разделе получим некоторые достаточные условия того, что штрафная функция является точной штрафной. Для этого потребуется вспомогательное понятие.

ОПРЕДЕЛЕНИЕ 2. Пусть функция g определена в некоторой окрестности точки $x \in \mathbb{R}^n$ и принимает вещественные значения. *Скоростью наискорейшего спуска* функции g в точке x называется величина

$$g^{\downarrow}(x) = \liminf_{y \to x} \frac{g(y) - g(x)}{\|y - x\|}.$$

Здесь $\|\cdot\|$ — евклидова норма в $\mathbb{R}^n$.

Укажем некоторые правила вычисления скорости наискорейшего спуска.

ПРЕДЛОЖЕНИЕ 3. *Пусть функция g определена в некоторой окрестности точки $x \in \mathbb{R}^n$ и принимает вещественные значения. Справедливы следующие утверждения:*

1) *если функция g дифференцируема в точке x, то $g^{\downarrow}(x) = -\|g'(x)\|$, где $g'(x)$ — градиент функции g в точке x;*

2) *если функция g дифференцируема по направлениям в смысле Адамара в точке x, то*
$$g^{\downarrow}(x) = \min_{\|h\|=1} g'(x,h),$$
где $g'(x,\cdot)$ — производная по направлениям функции g в точке x;

3) *если функция g выпукла в некоторой окрестности точки x, то*
$$g^{\downarrow}(x) = -\min_{v \in \partial g(x)} \|v\|,$$
где $\underline{\partial} g(x)$ — субдифференциал функции g в точке x.

4) *если функция g квазидифференцируема в смысле Адамара в точке x, то*
$$g^{\downarrow}(x) = -\max_{w \in \overline{\partial} g(x)} \min_{v \in \underline{\partial} g(x) + \{w\}} \|v\|,$$
где $\mathcal{D}g(x) = [\underline{\partial} g(x), \overline{\partial} g(x)]$ — квазидифференциал функции g в точке x.

Доказательство. 1) По неравенству Коши–Буняковского для любой последовательности y_k стремящейся к x будет
$$g(y_k) - g(x) = \langle g'(x), y_k - x\rangle + o(\|y_k - x\|) \geqslant -\|g'(x)\|\|y_k - x\| + o(\|y_k - x\|),$$
где $\langle \cdot,\cdot\rangle$ — скалярное произведение в $\mathbb{R}^n$ и $o(\|y_k-x\|)/\|y_k-x\| \to 0$ при $k \to \infty$. Откуда
$$\liminf_{k\to\infty} \frac{g(y_k) - g(x)}{\|y_k - x\|} \geqslant -\|g'(x)\|$$
и, следовательно, $g^{\downarrow}(x) \geqslant -\|g'(x)\|$. При этом очевидно, что если $g'(x) = 0$, то $g^{\downarrow}(x) = 0$. Если же $g'(x) \neq 0$, то, положив
$$y_k = x - \frac{1}{k} g'(x) \quad k \in \mathbb{N}$$
и воспользовавшись дифференцируемостью функции g, получим
$$\lim_{k\to\infty} \frac{g(y_k) - g(x)}{\|y_k - x\|} = \lim_{k\to\infty} \frac{-\frac{1}{k}\langle g'(x), g'(x)\rangle + o\left(\frac{1}{k}\right)}{\frac{1}{k}\|g'(x)\|} = -\|g'(x)\|.$$

Значит $g^{\downarrow}(x) = -\|g'(x)\|$.

2) Так как функция g дифференцируема по направлениям в смысле Адамара, то для любого $h \in \mathbb{R}^n$ будет

$$g'(x,h) = \lim_{[\alpha,h'] \to [+0,h]} \frac{g(x+\alpha h') - g(x)}{\alpha}.$$

Отсюда вытекает, что для любого h такого, что $\|h\| = 1$ будет $g^{\downarrow}(x) \leqslant g'(x,h)$.

По определению нижнего предела существует последовательность y_k, сходящаяся к x, и такая, что

$$g^{\downarrow}(x) = \lim_{k\to\infty} \frac{g(y_k) - g(x)}{\|y_k - x\|}.$$

Положим

$$\alpha_k = \|y_k - x\|, \quad h_k = \frac{1}{\|y_k - x\|}(y_k - x).$$

Ясно, что $\|h_k\| = 1$ для всех $k \in \mathbb{N}$. Поэтому из последовательности h_k можно извлечь подпоследовательность, сходящуюся к некоторому вектору $h^* \in \mathbb{R}^n$, $\|h^*\| = 1$. Без ограничения общности можно считать, что сама последовательность h_k сходится к h^*. Тогда

$$g^{\downarrow}(x) = \lim_{k\to\infty} \frac{g(y_k) - g(x)}{\|y_k - x\|} = \lim_{k\to\infty} \frac{g(x+\alpha_k h_k) - g(x)}{\alpha_k} = g'(x,h^*),$$

откуда получаем, что справедливо равенство

$$g^{\downarrow}(x) = \min_{\|h\|=1} g'(x,h).$$

3) Выпуклая функция g является дифференцируемой по направлениям в точке x ([3], теорема 1.4.2) и удовлетворяет условию Липшица в некоторой окрестности этой точки ([3], лемма 1.5.5). Поэтому, как нетрудно проверить, функция g дифференцируема по направлениям в смысле Адамара в точке x, причём

$$\min_{\|h\|=1} g'(x,h) = -\min_{v \in \underline{\partial} g(x)} \|v\|$$

по теореме 1.6.2 из [3]. Откуда, с учётом пункта 2, получаем справедливость требуемого равенства.

4) Справедливость данного утверждения вытекает из формулы для вычисления направления наискорейшего спуска квазидифференцируемой функции (см. [3], с. 217–218). $\square$

Введём дополнительное обозначение. Для любого $\delta > 0$ положим

$$\Omega_\delta = \{x \in \mathbb{R}^n \mid \varphi(x) < \delta\}.$$

Следующая теорема даёт достаточные условия того, что штрафная функция F_λ является точной штрафной.

ТЕОРЕМА 1. *Пусть выполнены следующие условия:*

1) *существует $\lambda_0 \geqslant 0$ такое, что для любого $\lambda \geqslant \lambda_0$ штрафная функция F_λ достигает глобального минимума;*

2) *существуют $\delta > 0$ и $a > 0$ такие, что функция φ полунепрерывна снизу на $\Omega_\delta \setminus \Omega$ и*
$$\varphi^{\downarrow}(x) \leqslant -a \quad \forall\, x \in \Omega_\delta \setminus \Omega;$$

3) *функция f удовлетворяет условию Липшица на множестве $\Omega_\delta \setminus \Omega$.*

Тогда существует такое $\lambda^* \geqslant \lambda_0$, что для любого $\lambda > \lambda^*$ и любой точки глобального минимума x_λ штрафной функции F_λ будет $\varphi(x_\lambda) = 0$, то есть точка x_λ является решением задачи (1) и штрафная функция F_λ является точной штрафной.

Доказательство. Допустим, что утверждение теоремы неверно. Тогда найдутся возрастающая последовательность λ_k и точки глобального минимума x_{λ_k} штрафной функции F_{λ_k} такие, что $\lambda_k \to \infty$ при $k \to \infty$ и $\varphi(x_{\lambda_k}) > 0$ для всех $k \in \mathbb{N}$. По предложению 2 будет $\varphi(x_{\lambda_k}) \to 0$ при $k \to \infty$. Следовательно, для любого k больше некоторого $k_0 \in \mathbb{N}$ будет

$$0 < \varphi(x_{\lambda_k}) < \delta, \quad \lambda_k > \frac{2L}{a}, \tag{9}$$

где L — это константа Липшица функции f на множестве $\Omega_\delta \setminus \Omega$.

Зафиксируем произвольное $k > k_0$. Из (9) вытекает, что $x_{\lambda_k} \in \Omega_\delta \setminus \Omega$. Следовательно, по условию теоремы будет $\varphi^{\downarrow}(x_{\lambda_k}) \leqslant -a$. Поэтому по определению скорости наискорейшего спуска получим, что существует последовательность $\{y_s\} \subset \mathbb{R}^n$, сходящаяся к x_{λ_k}, для которой

$$\varphi^{\downarrow}(x_{\lambda_k}) = \lim_{s \to \infty} \frac{\varphi(y_s) - \varphi(x_{\lambda_k})}{\|y_s - x_{\lambda_k}\|} \leqslant -a.$$

Поэтому без ограничения общности можем считать, что

$$\varphi(y_s) - \varphi(x_{\lambda_k}) \leqslant -\frac{a}{2}\|y_s - x_{\lambda_k}\| \quad \forall\, s \in \mathbb{N},$$

откуда следует, что $y_s \in \Omega_\delta$ для всех s. Учитывая полунепрерывность снизу функции φ в точке $x_{\lambda_k} \in \Omega_\delta \setminus \Omega$ и тот факт, что $\varphi(x_{\lambda_k}) > 0$, получаем, что для всех s больше некоторого $s_0 \in \mathbb{N}$ будет $\varphi(y_s) > 0$ и, следовательно, $y_s \in \Omega_\delta \setminus \Omega$.

Воспользовавшись липшицевостью функции f на множестве $\Omega_\delta \setminus \Omega$, получим, что для любых $s > s_0$ справедливо неравенство

$$F_{\lambda_k}(y_s) - F_{\lambda_k}(x_{\lambda_k}) = f(y_s) - f(x_{\lambda_k}) + \lambda_k(\varphi(y_s) - \varphi(x_{\lambda_k})) \leqslant \\ \leqslant L\|y_s - x_{\lambda_k}\| - \frac{a}{2}\lambda_k\|y_s - x_{\lambda_k}\| = \left(L - \frac{a}{2}\lambda_k\right)\|y_s - x_{\lambda_k}\|.$$

Поскольку $\lambda_k > 2L/a$ (см. (9)), то для всех $s > s_0$ будет

$$F_{\lambda_k}(y_s) - F_{\lambda_k}(x_{\lambda_k}) \leqslant \left(L - \frac{a}{2}\lambda_k\right)\|y_s - x_{\lambda_k}\| < 0,$$

что противоречит определению точки x_{λ_k}. Следовательно, существует такое $\lambda^* \geqslant \lambda_0$, что для любого $\lambda > \lambda^*$ и для любой точки глобального минимума x_λ функции F_λ будет $\varphi(x_\lambda) = 0$, что и требовалось. □

Замечание 1. Предположим, что в условиях предыдущей теоремы функция f ограничена снизу на $\mathbb{R}^n$. Тогда можно положить

$$\lambda^* = \max\left\{\lambda_0, \frac{2L}{a}, \frac{f^*_\Omega - f^*}{\delta}\right\},$$

где

$$f^* = \inf_{x \in \mathbb{R}^n} f(x) > -\infty, \quad f^*_\Omega = \inf_{x \in \Omega} f(x).$$

Действительно, зафиксируем произвольное $\lambda > \lambda^*$ и точку глобального минимума x_λ штрафной функции $F_\lambda(x)$, и предположим, что $\varphi(x_\lambda) > 0$. Покажем сначала, что $\varphi(x_\lambda) < \delta$.

Для любого решения x_0 задачи (1) и для всех $x \in \mathbb{R}^n \setminus \Omega_\delta$ будет

$$F_\lambda(x) = f(x) + \lambda\varphi(x) > f^* + \frac{f^*_\Omega - f^*}{\delta}\delta = f^*_\Omega = f(x_0) = F_\lambda(x_0) \geqslant F_\lambda(x_\lambda),$$

откуда получаем, что $x_\lambda \in \Omega_\delta$, то есть $\varphi(x_\lambda) < \delta$. Рассуждая далее как и при доказательстве теоремы 1, нетрудно показать, что найдётся последовательность $\{y_s\}$ такая, что $y_s \to x_\lambda$ при $s \to \infty$ и для достаточно больших s справедливо неравенство

$$F(y_s) - F(x_\lambda) < 0,$$

что противоречит определению точки x_λ.

Покажем на примерах, что условия теоремы 1 близки к оптимальным и не могут быть существенно ослаблены.

ПРИМЕР 1. Пусть $n = 1$, $\Omega = (-\infty, 0]$, $\varphi(x) = x_+ = \max\{0, x\}$ и

$$f(x) = \begin{cases} 0, \text{ если } x \leqslant 0, \\ -\sqrt{x}, \text{ если } x \in [0, 1], \\ -1, \text{ если } x > 1. \end{cases}$$

Нетрудно проверить, что для любого $\lambda > 1$ будет

$$\arg\min_{x \in \mathbb{R}^n} F_\lambda(x) = \left\{ \frac{1}{4\lambda^2} \right\},$$

т. е. $\varphi(x_\lambda) > 0$ для всех $\lambda > 1$ и штрафная функция $F_\lambda(\cdot) = f(\cdot) + \lambda\varphi(\cdot)$ не является точной штрафной. Ясно, что в данном примере выполнены все условия теоремы 1, кроме условия 3 (функция f не удовлетворяет условию Липшица на множестве $\Omega_\delta \setminus \Omega = (0, \delta)$ ни для какого $\delta > 0$).

ПРИМЕР 2. Пусть $n = 1$, $\Omega = (-\infty, 0]$, $f(x) = \max\{-x, -1\}$ и $\varphi(x) = (x_+)^2$. Легко видеть, что для любого $\lambda > 1$ будет

$$\arg\min_{x \in \mathbb{R}^n} F_\lambda(x) = \left\{ \frac{1}{2\lambda} \right\},$$

т. е. $\varphi(x_\lambda) > 0$ для всех $\lambda > 1$. В данном примере выполнены все условия теоремы 1, кроме условия 2. Действительно, нетрудно проверить, что

$$\varphi^{\downarrow}(x) = -2|x| \quad \forall\, x > 0$$

и, следовательно, не существует таких $a > 0$ и $\delta > 0$, что $\varphi^{\downarrow}(x) \leqslant -a$ для всех $x \in \Omega_\delta \setminus \Omega = (0, \sqrt{\delta})$.

ПРИМЕР 3. Пусть $n = 1$, $\Omega = (-\infty, 0]$, $\varphi(x) = x_+$ и

$$f(x) = \begin{cases} 0, \text{ при } x \leqslant 0, \\ -x^2 - 1, \text{ при } x \in (0, 1], \\ -2, \text{ при } x > 1. \end{cases}$$

Нетрудно проверить, что в данном примере выполнены все условия теоремы 1, кроме условия 1. А именно, штрафная функция F_λ не достигает глобального минимума для любого $\lambda > 1$. При этом заметим, что функция f (как и F_λ) не является полунепрерывной снизу.

4°. **Точки локального минимума штрафной функции.** Сформулируем условия, при которых точки локального минимума функции f на множестве Ω являются точками локального минимума штрафной функции F_λ. Эти условия позволяют сводить исследование исходной задачу (1) к исследованию задачи минимизации штрафной функции F_λ без ограничений.

Введём дополнительные обозначения. Для любого $x \in \mathbb{R}^n$ и $r > 0$ положим $B(x, r) = \{y \in \mathbb{R}^n \mid \|x - y\| \leqslant r\}$ — замкнутый шар с центром в точке x радиуса r. Определим также

$$d(x, \Omega) = \inf_{y \in \Omega} \|x - y\|.$$

Величина $d(x, \Omega)$ называется расстоянием от точки x до множества Ω.

ТЕОРЕМА 2. *Пусть $x_0 \in \Omega$ является точкой локального минимума функции f на множестве Ω и предположим, что существуют $a > 0$, $r > 0$ и $L > 0$ такие, что*

$$\varphi(x) \geqslant a d(x, \Omega) \quad \forall\, x \in B(x_0, r),$$

а функция f удовлетворяет условию Липшица с константой L на $B(x_0, r)$. Тогда для любого $\lambda > L/a$ точка x_0 является точкой локального минимума штрафной функции F_λ.

Доказательство. Так как x_0 — это точка локального минимума функции f на множестве Ω, то без ограничения общности можно считать, что

$$f(x_0) \leqslant f(x) \quad \forall\, x \in B(x_0, r) \cap \Omega. \tag{10}$$

Зафиксируем произвольное $x \in B(x_0, r/2)$. По определению точной нижней грани найдётся последовательность точек $\{x_k\} \subset \Omega$ такая, что

$$\|x_k - x\| \xrightarrow[k \to \infty]{} d(x, \Omega).$$

При этом не ограничивая общности можно считать, что $\|x_k - x\| \leqslant r/2$, так как $\|x_0 - x\| \leqslant r/2$ и $x_0 \in \Omega$. Поэтому $x_k \in B(x_0, r)$ для всех $k \in \mathbb{N}$.

Для любого $\lambda > L/a$ имеем

$$F_\lambda(x_0) - F_\lambda(x) = f(x_0) - f(x) - \lambda\varphi(x) = f(x_0) - f(x_k) + f(x_k) - f(x) - \lambda\varphi(x),$$

так как $x_0 \in \Omega$. Откуда, учитывая (10) и тот факт, что $x_k \in B(x_0, r) \cap \Omega$, имеем

$$F_\lambda(x_0) - F_\lambda(x) \leqslant f(x_k) - f(x) - \lambda\varphi(x) \leqslant L\|x_k - x\| - \lambda a d(x, \Omega).$$

Так как $\lambda > L/a$, то переходя к пределу при $k \to \infty$, получим, что

$$F_\lambda(x_0) - F_\lambda(x) \leqslant (L - \lambda a) d(x, \Omega) < 0.$$

Следовательно
$$F_\lambda(x_0) \leqslant F_\lambda(x) \quad \forall\, x \in B(x_0, r/2),$$
а это и означает, что x_0 — это точка локального минимума штрафной функции F_λ для любого $\lambda > L/a$. □

В теоремах 1 и 2 использовались различные предположения относительно функции φ. А именно, в теореме 1 требовалось выполнение неравенства $\varphi^{\downarrow}(x) \leqslant -a$, при этом в теореме 2 требовалось выполнение оценки $\varphi(x) \geqslant ad(x, \Omega)$. Укажем связь между двумя данными условиями. Для этого нам потребуется вспомогательное утверждение [4].

ТЕОРЕМА 3 (Вариационный принцип Экланда)**.** *Пусть полунепрерывная снизу функция* $f\colon \mathbb{R}^n \to \mathbb{R}$ *ограничена снизу. Пусть также для некоторого* $\varepsilon > 0$ *точка* $x_\varepsilon \in \mathbb{R}^n$ *такова, что*
$$f(x_\varepsilon) \leqslant \inf_{x \in \mathbb{R}^n} f(x) + \varepsilon.$$
Тогда для любого $\lambda > 0$ *существует точка* $y \in \mathbb{R}^n$ *такая, что*:

1) $\|y - x_\varepsilon\| \leqslant \lambda$*;*

2) $f(y) \leqslant f(x_\varepsilon)$*;*

3) для всех $x \in \mathbb{R}^n$, $x \neq y$ *выполняется неравенство*
$$f(x) + \frac{\varepsilon}{\lambda}\|x - y\| > f(y).$$

ТЕОРЕМА 4. *Пусть* $x_0 \in \Omega$*, функция* φ *непрерывна, и предположим, что существуют* $r > 0$ *и* $a > 0$ *такие, что*
$$\varphi^{\downarrow}(x) \leqslant -a \quad \forall\, x \in B(x_0, r) \setminus \Omega.$$
Тогда существует такое $\delta > 0$*, что*
$$\varphi(x) \geqslant ad(x, \Omega) \quad \forall\, x \in B(x_0, \delta). \tag{11}$$

Доказательство. Поскольку функция φ непрерывна, то существует такое $\delta > 0$, что
$$\varphi(x) < \frac{ra}{4} \quad \forall\, x \in B(x_0, \delta). \tag{12}$$
При этом без ограничения общности можно считать, что $\delta < r/2$. Пусть $x \in B(x_0, \delta)$. Если $x \in \Omega$, то $\varphi(x) = 0$ и неравенство (11) очевидно выполнено. Предположим теперь, что $x \notin \Omega$. Обозначим $\varepsilon = \varphi(x) > 0$ и зафиксируем

произвольное $t \in (a/2, a)$. По вариационному принципу Экланда найдётся такое $y \in \mathbb{R}^n$, что

$$\varphi(y) \leqslant \varphi(x), \quad \|y-x\| \leqslant \frac{\varepsilon}{t}, \quad \varphi(v) + t\|y-v\| > \varphi(y) \quad \forall\, v \in \mathbb{R}^n \setminus \{y\}. \qquad (13)$$

Покажем, что $\varphi(y) = 0$. Действительно, если $\varphi(y) > 0$, то, учитывая (12), имеем, что $y \in B(x_0, r) \setminus \Omega$, поскольку

$$x \in B(x_0, \delta) \subset B\left(x_0, \frac{r}{2}\right), \quad \|y-x\| \leqslant \frac{\varepsilon}{t} < 2\frac{\varphi(x)}{a} < \frac{r}{2}.$$

Следовательно $\varphi^{\downarrow}(y) \leqslant -a$, и поэтому существует такое $w \in \mathbb{R}^n$, $w \neq y$, что

$$\frac{\varphi(w) - \varphi(y)}{\|w-y\|} \leqslant -t.$$

Откуда $\varphi(w) + t\|w-y\| \leqslant \varphi(y)$, что противоречит (13). Значит $\varphi(y) = 0$, то есть $y \in \Omega$. Теперь воспользовавшись вторым неравенством из (13), получим

$$d(x, \Omega) \leqslant \|y-x\| \leqslant \frac{\varepsilon}{t} = \frac{\varphi(x)}{t}.$$

Поскольку $t \in (a/2, a)$ произвольно, то из последнего неравенства следует, что

$$\varphi(x) \geqslant a d(x, \Omega),$$

что и требовалось доказать. $\square$

Замечание 2. Идея доказательства предыдущей теоремы основана на доказательстве основной леммы из [5].

ТЕОРЕМА 5. *Пусть функция φ полунепрерывна снизу, и предположим, что существуют такие $\delta > 0$ и $a > 0$, что*

$$\varphi^{\downarrow}(x) \leqslant -a \quad \forall\, x \in \Omega_\delta \setminus \Omega.$$

Тогда

$$\varphi(x) \geqslant a d(x, \Omega) \quad \forall\, x \in \Omega_\delta.$$

Доказательство. Зафиксируем произвольное $x \in \Omega_\delta$. Если $x \in \Omega$, то, очевидно, $\varphi(x) = a d(x, \Omega) = 0$. Поэтому предположим, что $x \in \Omega_\delta \setminus \Omega$.

Положим $\varepsilon = \varphi(x)$ и зафиксируем произвольное $t \in (0, a)$. По вариационному принципу Экланда существует $y \in \mathbb{R}^n$ такое, что

$$\varphi(y) \leqslant \varphi(x), \quad \|y-x\| \leqslant \frac{\varepsilon}{t}, \quad \varphi(v) + t\|v-y\| > \varphi(y) \quad \forall\, v \in \mathbb{R}^n \setminus \{y\}. \qquad (14)$$

Ясно, что $y \in \Omega_\delta$. Покажем, что $\varphi(y) = 0$, то есть $y \in \Omega$. Действительно, если $\varphi(y) > 0$, то $y \in \Omega_\delta \setminus \Omega$ и поэтому $\varphi^{\downarrow}(y) \leqslant -a$. Следовательно, существует $w \in \mathbb{R}^n$, $w \neq y$ такое, что

$$\frac{\varphi(w) - \varphi(y)}{\|w - y\|} \leqslant -t.$$

Отсюда,

$$\varphi(w) + t\|w - y\| \leqslant \varphi(y),$$

что противоречит (14). Значит $y \in \Omega$. Откуда, воспользовавшись вторым неравенством из (14), получим, что

$$d(x, \Omega) \leqslant \|y - x\| \leqslant \frac{\varepsilon}{t} = \frac{\varphi(x)}{t}.$$

Поскольку $t \in (0, a)$ произвольно, то

$$\varphi(x) \geqslant a d(x, \Omega),$$

что и требовалось. □

Приведём пример применения теоремы 2.

ПРИМЕР 4. [6] Пусть $n = 2$, $f(x) = f(x_1, x_2) = -x_1$ и

$$\Omega = \{x = (x_1, x_2) \in \mathbb{R}^2 \mid -x_1^3 + x_2 \leqslant 0,\ x_1^4 - x_2 \leqslant 0\}$$

(см. рис. 2 ниже), то есть рассматривается экстремальная задача с ограничениями–неравенствами:

$$\begin{gathered} f(x_1, x_2) = -x_1 \to \inf, \\ -x_1^3 + x_2 \leqslant 0, \quad x_1^4 - x_2 \leqslant 0. \end{gathered}$$

Нетрудно понять, что единственным решением данной задачи является точка $\overline{x} = (1, 1)$. Исследуем эту задачу с помощью теории точных штрафных функций. Определим

$$\varphi(x) = \max\{0, -x_1^3 + x_2, x_1^4 - x_2\}$$

и зафиксируем произвольное $\delta > 0$. Покажем, что для данной функции φ не выполняются условия теоремы 1. Действительно, введём множество

$$A = \{x = (x_1, x_2) \in \mathbb{R}^2 \mid x_1 < 0,\ -x_1^3 + x_2 = x_1^4 - x_2\}.$$

Легко видеть, что для любого $r > 0$ множество $A \cap B(0, r)$ непусто. В силу непрерывности функции φ (заметим, что $\varphi(0) = 0$) найдётся такое $r > 0$, что справедливо включение $A \cap B(0, r) \subset \Omega_\delta \setminus \Omega$.

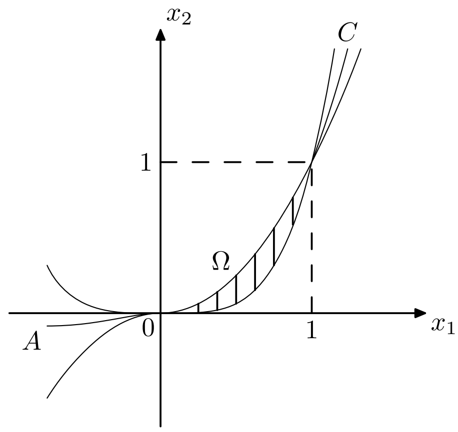

Рис. 1. Множество Ω.

Функция φ является субдифференцируемой и для любого $x \in A \cap B(0,r)$ её субдифференциал имеет вид

$$\underline{\partial}\varphi(x) = \mathrm{co}\{(-3x_1^2, 1), (4x_1^3, -1)\}. \tag{15}$$

С учётом пункта 4 предложения 3 получим, что

$$\varphi^{\downarrow}(x) \geqslant -\left|-\frac{3}{2}x_1^2 + 2x_1^3\right| = -\|y(x)\| \quad \forall\, x \in A \cap B(0,r),$$

где

$$y(x) = \frac{1}{2}(-3x_1^2, 1) + \frac{1}{2}(4x_1^3, -1) \in \underline{\partial}\varphi(x).$$

Учитывая, что $\|y(x)\| \to 0$ при $x \to 0$, получаем, что для любых $\delta > 0$ и $a > 0$ найдётся такое $x \in \Omega_\delta \setminus \Omega$, что

$$\varphi^{\downarrow}(x) > -a$$

и, следовательно, теоремой 1 воспользоваться нельзя.

Покажем теперь, что в некоторой окрестности точки $\overline{x}$ выполнены условия теоремы 2. Для этого, с учётом теоремы 4, достаточно доказать, что найдутся $r > 0$ и $a > 0$ такие, что

$$\varphi^{\downarrow}(x) \leqslant -a \quad \forall\, x \in B(\overline{x}, r) \setminus \Omega. \tag{16}$$

Введём множество

$$C = \{x = (x_1, x_2) \in \mathbb{R}^2 \mid x_1 > 1,\ -x_1^3 + x_2 = x_1^4 - x_2\}.$$

Если $x \notin \Omega \cup C$ и $x_1 > 0$, то, как нетрудно проверить, функция φ дифференцируема в точке x и
$$\varphi^{\downarrow}(x) \leqslant -1.$$
по пункту 1 предложения 3. Если же $x \in C$, то функция φ субдифференцируема в точке x и её субдифференциал в этой точке совпадает со значением многозначного отображения Φ в точке x, где

$$\Phi(y) = \Phi(y_1, y_2) = \mathrm{co}\{(-3y_1^2, 1), (4y_1^3, -1)\} \quad \forall\, y \in \mathbb{R}^2$$

(см. (15)). С учётом пункта 4 предложения 3, справедливо равенство

$$\varphi^{\downarrow}(x) = -\min_{v \in \Phi(x)} \|v\| \quad \forall\, x \in C.$$

Ясно, что величина $\min\{\|v\| \mid v \in \Phi(x)\}$ непрерывно зависит от x_1. Поэтому, если $\min\{\|v\| \mid v \in \Phi(\overline{x})\} > 0$, то найдутся такие $r > 0$ и $a > 0$, что выполнено (16). Имеем $\Phi(\overline{x}) = \mathrm{co}\{(-3, 1), (4, -1)\}$, откуда легко получить, что

$$\min_{v \in \Phi(\overline{x})} \|v\| = \frac{1}{\sqrt{53}} > 0,$$

что и требовалось.

З а м е ч а н и е 3. В предыдущем примере в качестве функции φ можно взять функцию
$$\varphi(x) = \max\{0, -x_1^3 + x_2\} + \max\{0, x_1^4 - x_2\}.$$
Для данной функции φ условия теоремы 1 также не выполнены, однако нетрудно показать, что, как и в предыдущем примере, существует окрестность точки $\overline{x} = (1, 1)$, в которой выполнены условия теоремы 2.

5°. Точные гладкие штрафные функции. В данном разделе мы покажем, что, как правило, точная штрафная функция является негладкой, даже если исходная функция f гладкая. Иными словами, если функции f и φ гладкие, то штрафная функция F_λ не может быть точной штрафной.

ПРЕДЛОЖЕНИЕ 4. *Пусть функция f дифференцируема по направлениям, а функция φ дифференцируема на Ω, и предположим, что штрафная функция F_λ является точной штрафной. Тогда для любого решения $x_0 \in \Omega$ задачи* (1) *справедливо неравенство*

$$f'(x_0, h) \geqslant 0 \quad \forall\, h \in \mathbb{R}^n,$$

то есть в точке x_0 выполнено необходимое условие минимума дифференцируемой по направлениям функции. В частности, если функция f дифференцируема в точке x_0, то $f'(x_0) = 0$.

Доказательство. Зафиксируем произвольную точку $x_0 \in \Omega$, являющуюся точкой глобального минимума функции f на Ω. Поскольку функция F_λ является точной штрафной, то для достаточно большого λ точка x_0 является точкой глобального минимума функции F_λ. Функция F_λ, очевидно, является дифференцируемой по направлениям. Поэтому в точке x_0 выполнено необходимое условие минимума дифференцируемой по направлениям функции:

$$F'_\lambda(x_0, h) = f'(x_0, h) + \lambda\langle\varphi'(x_0), h\rangle \geqslant 0 \quad \forall\, h \in \mathbb{R}^n.$$

Так как $x_0 \in \Omega$, то x_0 является точкой глобального минимума функции φ (напомним, что функция φ неотрицательна и равна нулю на Ω по определению). Поэтому $\varphi'(x_0) = 0$. Откуда получаем, что $f'(x_0, h) \geqslant 0$ для всех h. □

ЛИТЕРАТУРА

1. Демьянов В.Ф. Условия экстремума и вариационное исчисление. М.: Высш. шк., 2004. 335 с.

2. Demyanov V.F., Di Pillo G., Facchinei F. Exact penalization via Dini and Hadamars conditional derivatives // Optimization Methods and Software, 1998. vol. 9, pp. 19–36.

3. Демьянов В.Ф., Васильев Л.В. Недифференцируема оптимизация. М.: Наука, 1981. 384 с.

4. Ekeland I. On the variational principle // Journal of Mathematical Analysis and Applications, 1974. vol. 47. pp. 324–353.

5. Иоффе А.Д. Метрическая регулярность и субдифференциальное исчисление // УМН, 2000. Т. 55, вып. 3, с. 103–162.

6. Малозёмов В.Н. Теорема Куна–Таккера в дифференциальной форме [Электронный ресурс] // Семинар «DHA & CAGD» [Офиц. сайт]. Избранные доклады. 27.02.2010. URL: `http://dha.spb.ru/reps10.shtml#1016` (дата обращения: 14.05.2014).