ОБЗОР ПО ТОЧНЫМ ШТРАФНЫМ ФУНКЦИЯМ*

М. В. Долгополик

maxim.dolgopolik@gmail.com

2 октября 2014 г.

Аннотация. В докладе приводится обзор некоторых результатов по теории точных штрафных функций для задач нелинейного программирования.

1°. Постановка задачи. Рассмотрим задачу математического программирования

$$f(x) \to \inf, \tag{1}$$

$$g_i(x) \leqslant 0, \quad i \in I = \{1, \ldots, m\}, \tag{2}$$

$$h_j(x) = 0, \quad j \in J = \{1, \ldots, s\}. \tag{3}$$

Предполагается, что f, g_i и h_j, $i \in I$, $j \in J$ — произвольные конечные функции заданные на $\mathbb{R}^n$. Обозначим через Ω множество допустимых планов задачи (1)–(3). Обозначим также

$$r(x) = \sum_{i=1}^{m} [g_i(x)]_+ + \sum_{j=1}^{s} |h_j(x)|, \tag{4}$$

и $f^* = \inf\limits_{x \in \Omega} f(x)$.

Пусть на $\mathbb{R}^n$ задана произвольная неотрицательная функция φ такая, что $\varphi(x) = 0$ тогда и только тогда, когда точка x удовлетворяет ограничениям (2) и (3). Функция

$$F_\lambda(x) = f(x) + \lambda \varphi(x), \quad \lambda \geqslant 0,$$

называется *штрафной функцией* для задачи (1)–(3). В частности, если

$$\varphi(x) = r(x),$$

то соответствующая функция F_λ называется ℓ_1 штрафной функцией.

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

Напомним [1, 2], что штрафная функция F_λ называется *точной* в точке x^* локального минимума в задаче (1)–(3), если существует $\lambda^* \geqslant 0$ такое, что для любого $\lambda \geqslant \lambda^*$ точка x^* является точкой локального минимума штрафной функции F_λ. Штрафная функция F_λ называется *точной на множестве* $C \subset \mathbb{R}^n$, если существует $\lambda^* \geqslant 0$ такое, что

$$F_\lambda(x) \geqslant f^* \quad \forall x \in C \quad \forall \lambda \geqslant \lambda^*.$$

Наконец, штрафная функция F_λ называется *точной*, если существует λ^* такое, что для любого $\lambda \geqslant \lambda^*$ множество решений задачи (1)–(3) совпадает с множеством точек глобального минимума функции F_λ.

2°. Субаналитические функции и их приложения в теории оптимизации. Большинство результатов о точности штрафных функций для задач нелинейного программирования опирается на предположение о выполнении каких-либо условий регулярности в рассматриваемой задаче. Однако, интерес также представляются достаточные условия для точности штрафной функции, не опирающиеся на данное предположение. Такие условия можно получить в рамках теории субаналитических функций [3–6]. При этом предположения о выполнении условий регулярности ограничений заменяется на предположение о том, что функции, определяющие ограничения в задаче, обладают некоторыми аналитическими свойствами.

Введём необходимые определения. Напомним, что функция f определённая на открытом множестве U из $\mathbb{R}^n$ и принимающая вещественные значения называется *аналитической*, если для каждой точки из множества U существует окрестность, в которой функция f представима в виде сходящегося степенного ряда.

Множество $X \subset \mathbb{R}^n$ называется *полуаналитическим*, если для каждой точки $y \in \mathbb{R}^n$ существуют окрестность U точки y и конечное семейство множеств $X_{ij} \subset \mathbb{R}^n$, $i \in \{1, \ldots, p\}$, $j \in \{1, \ldots, q\}$, представимых в виде

$$X_{ij} = \{x \in U \mid f_{ij}(x) = 0\} \quad \text{или} \quad X_{ij} = \{x \in \mathbb{R}^n \mid f_{ij}(x) < 0\}$$

для некоторых вещественных аналитических функций f_{ij}, такое, что

$$X \cap U = \bigcup_{i=1}^{p} \bigcap_{j=1}^{q} X_{ij}.$$

Другими словами, множество X является полуаналитическим, если оно локально представимо в виде объединения множеств решений систем равенств и строгих неравенств

$$\big\{x \in U \mid f_{ij}(x) = 0 \quad \text{или} \quad f_{ij}(x) < 0, \quad j \in \{1, \ldots, q\}\big\},$$

с аналитическими левыми частями. Подмножество $X \subset \mathbb{R}^n$ называется *субаналитическим*, если для каждой точки $y \in \mathbb{R}^n$ существует окрестность U точки y такая, что множество $X \cap U$ представимо в виде проекции некоторого ограниченного полуаналитического множества $A \subset \mathbb{R}^{n+p}$ на пространство $\mathbb{R}^n$ для некоторого $p \in \mathbb{N} \cup \{0\}$. Наконец, вектор-функция F называется *полуаналитической* (*субаналитической*), если её график является *полуаналитическим* (*субаналитическим*) множеством. Теория полуаналитических и субаналитических функций восходит к работам М.С. Лоясевича (англ. M.S. Lojasiewicz) [7, 8].

До настоящего времени не существует какого-либо простого описания субаналитических множеств и функций. Однако, построено достаточно полноценное исчисление данных множеств и функций. А именно, класс субаналитических множеств замкнут относительно стандартных операций над множествами, среди которых объединение, пересечение, разность, декартово произведение, операции замыкания и внутренности, а также операция проектирования на подпространство. Простыми примерами субаналитических множеств являются любое конечное и любое открытое множества, полиэдральное множество (в частности, выпуклый многогранник) и любое множество вида

$$X = \big\{x \in \mathbb{R}^n \mid g_i(x) \leqslant 0, \quad h_j(x) = 0, \quad i \in \{1, \ldots, p\}, \quad j \in \{1, \ldots, q\}\big\},$$

где g_i и h_j аналитические функции.

Класс непрерывных субаналитических функций, определённых на компактном субаналитическом множестве, замкнут относительно всех алгебраических операций, операций взятия поточечного минимума и максимума конечного числа функций, а также операции суперпозиции. В частности, функция вида

$$f(x_1, x_2) = \min\{\operatorname{arctg} x_1, \cos x_2\} \cdot e^{\max\{x_1, x_2 \sin x_1\}}$$

является субаналитической на любом шаре. Также функция расстояния до субаналитического множества X

$$d(x, X) = \min_{y \in X} \|x - y\|,$$

где $\| \cdot \|$ — евклидова норма, является субаналитической функцией. Таким образом, субаналитические функций вовсе не обязаны быть гладкими.

Справедлива следующая теорема о точности штрафной функции для задачи нелинейного программирования (1)–(3), не опирающаяся на выполнение каких-либо условий регулярности ограничений в данной задаче.

ТЕОРЕМА 1. *[4, 5] Пусть $X \subset \mathbb{R}^n$ — компактное субаналитическое множество, функция f удовлетворяет условию Липшица на множестве X, а функции g_i и h_j являются непрерывными и субаналитическими на этом множестве, $i \in I$, $j \in J$. Тогда существует $n^* \in \mathbb{N}$ такое, что для любого $\gamma \geqslant n^*$ штрафная функция*

$$F_{\lambda,\gamma}(x) = f(x) + \lambda r(x)^{1/\gamma}$$

является точной на множестве X.

СЛЕДСТВИЕ 1. *Пусть x^* — точка локального минимума в задаче* (1)–(3). *Предположим, что функция f удовлетворяет условию Липшица в некоторой окрестности точки x^*, а функции g_i и h_j являются непрерывными и субаналитическим в некоторой окрестности этой точки, $i \in I$, $j \in J$. Тогда существует $n^* \in \mathbb{N}$ такое, что для любого $\gamma \geqslant n^*$ штрафная функция $F_{\lambda,\gamma}$ является точной в точке x^*.*

3°. Стационарные точки штрафной функции. Штрафные функции применяются с целью преобразования экстремальной задачи с ограничениями в задачу безусловной оптимизации. В рамках теории точных штрафных функций изучаются условия, гарантирующие, что задача безусловной минимизации штрафной функции эквивалентна исходной оптимизационной задаче при достаточно больших значениях штрафного параметра. При этом под эквивалентностью задач оптимизации подразумевается совпадение множеств их решений, то есть множеств точек глобального минимума. Однако, методы оптимизации для невыпуклых функций могут, в общем случае, находить лишь стационарные точки исследуемой функции. Поэтому возникает несогласованность теории точных штрафных функций и метода точных штрафных функций, как общего подхода к решению экстремальных задач с ограничениями. Теория гарантирует совпадение точек глобального минимума, а метод позволяет лишь найти стационарные точки, которые могут даже не принадлежать множеству допустимых планов исходной задачи. Для того чтобы (частично) устранить эту несогласованность необходимы условия, гарантирующие отсутствие стационарных точек штрафной функции, не принадлежащих множеству допустимых планов. Такие условия рассматривались в работах [9–12]. Здесь мы приведём условия из [10].

Напомним, что величина

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

где функция p определена в некоторой окрестности точки x и принимает вещественные значения, называется *скоростью наискорейшего спуска* функции p в точке x. Нетрудно проверить, что условие $p^{\downarrow}(x) \geqslant 0$ является необходимым условием локального минимума функции p. Поэтому точку в которой выполняется данное условие мы будем называть *стационарной* точкой функции p. Заметим, что условие $F^{\downarrow}_{\lambda}(x) \geqslant 0$, где $x \in \Omega$, эквивалентно выполнению условий Куна–Таккера [13].

Предположим, что все функции f, g_i и h_j непрерывно дифференцируемы на $\mathbb{R}^n$. Для любого $x \in \mathbb{R}^n$ введём множество

$$I(x) = \{i \in I \mid g_i(x) \geqslant 0\}.$$

Будем говорить, что в точке $x \in \mathbb{R}^n$ выполнено *условие регулярности*, если градиенты $g'_i(x)$, $i \in R(x)$ и $h'_j(x)$, $j \in J$, линейно независимы (ср. с определением регулярности ограничений в [13], где учитываются только *активные* ограничения-неравенства).

Оказывается, что выполнение условий регулярности на некотором множестве X гарантирует, что для достаточно большого значения штрафного параметра λ все стационарные точки штрафной функции на множестве X будут принадлежать множеству допустимых планов.

ТЕОРЕМА 2. *Пусть $X \subset \mathbb{R}^n$ — непустое ограниченное множество, и предположим, что условие регулярности выполняется в каждой точке $x \in X$. Тогда найдётся $\lambda^* \geqslant 0$ такое, что для любого $\lambda > \lambda^*$ не существует стационарных точек штрафной функции F_λ, принадлежащих множеству $X \setminus \Omega$.*

Замечание 2.1. Пусть выполнены условия предыдущей теоремы. Предположим также, что функция f удовлетворяет условию Липшица на множестве $X \setminus \Omega$ с константой $L > 0$, и существует $a > 0$ такое, что

$$r^{\downarrow}(x) \leqslant -a \quad \forall x \in X \setminus \Omega.$$

(существование таких констант $L > 0$ и $a > 0$ можно вывести из компактности множества X). Тогда можно показать, что утверждение теоремы выполняется при $\lambda^* = L/a$.

4°. Точные барьерные функции. Очевидно, что многие свойства штрафной функции зависят от свойств целевой функции и функций задающих ограничения в точках, расположенных сколь угодно далеко от множества допустимых планов. Поэтому возникает естественное желание каким-либо образом ограничить область задания штрафной функции. Это можно сделать с помощью введения штрафной функции специального вида [14].

Зафиксируем некоторое $\sigma > 0$, и определим штрафную функцию

$$G_\lambda(x) = f(x) + \lambda \frac{\varphi(x)}{\sigma - \varphi(x)}$$

при $x \in \mathbb{R}^n$ таких, что $\varphi(x) < \sigma$ и

$$G_\lambda(x) = +\infty,$$

если $\varphi(x) \geqslant \sigma$. Напомним, что φ — это произвольная неотрицательная функция такая, что $\varphi(x) = 0$ тогда и только тогда, когда точка x удовлетворяет ограничениям (2)–(3). Отметим, что введённую таким образом штрафную функцию достаточно рассматривать лишь на множестве

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

Следуя [14], будем называть функцию G_λ *точной барьерной функцией*, если существует $\lambda^* \geqslant 0$ такое, что для любого $\lambda \geqslant \lambda^*$ множество решений задачи (1)–(3) совпадает с множеством точек глобального минимума функции G_λ. Будем также говорить, что функция G_λ *точна* в точке x^* локального минимума в задаче (1)–(3), если существует $\lambda^* \geqslant 0$ такое, что для любого $\lambda \geqslant \lambda^*$ точка x^* является точкой локального минимума функции G_λ.

Приведём некоторые результаты о взаимосвязи штрафной функции G_λ с классической штрафной функцией $F_\lambda(x) = f(x) + \lambda\varphi(x)$, рассмотренной выше.

ПРЕДЛОЖЕНИЕ 1. *Пусть x^* является точкой локального минимума в задаче (1)–(3), и предположим, что функция φ непрерывна в точке x^*. Тогда функция G_λ является точной в точке x^* тогда и только тогда, когда штрафная функция F_λ точна в этой точке.*

Доказательство. Справедливость данного утверждения непосредственно вытекает из следующих очевидных неравенств

$$\frac{1}{\sigma}\varphi(x) \leqslant \frac{\varphi(x)}{\sigma - \varphi(x)} \leqslant \frac{2}{\sigma}\varphi(x),$$

выполняющихся для любых x таких, что $\varphi(x) < \sigma/2$. □

Аналогичным образом доказываются следующие утверждения.

ПРЕДЛОЖЕНИЕ 2. *Пусть функция G_λ является точной барьерной функцией. Тогда штрафная функция F_λ является точной на множестве Ω_δ для любого $\delta < \sigma$.*

ПРЕДЛОЖЕНИЕ 3. *Пусть штрафная функция F_λ является точной на множестве Ω_δ для некоторого $\delta > 0$. Тогда функция $G_\lambda(x)$ является точной барьерной функцией для любого $\sigma \leqslant \delta$.*

Воспользовавшись предложениями 1–3 и известными результатами из теории точных штрафных функций, легко получить достаточные условия того, что функция G_λ является точной барьерной функцией. В частности, с помощью теоремы 1 из [2] нетрудно вывести следующее утверждение.

ТЕОРЕМА 3. *Пусть выполнены следующие условия:*

1) существуют $\delta > 0$ и $a > 0$ такие, что

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

2) функция f удовлетворяет условию Липшица на множестве $\Omega_\delta \setminus \Omega$;

3) функции f и φ полунепрерывны снизу на множестве Ω_δ.

Тогда функция G_λ является точной барьерной функцией для любого $\sigma \leqslant \delta$.

5°. Гладкие точные штрафные функции. В [1, 9] было показано, что, как правило, точная штрафная функция является негладкой. Данное обстоятельство не позволяет использовать различные методы классической оптимизации для минимизации точной штрафной функции. В [15] был предложен новый подход к построению точных штрафных функций, который позволяет преодолеть эту трудность. Данный подход был обобщён и усовершенствован в [16]. В статьях [17–19] изложены приложения этого подхода к решению задач оптимального управления.

Опишем новый класс гладких точных штрафных функций. Для простоты будем считать, что ограничения типа неравенств отсутствуют, т.е. рассматривается задача вида

$$f(x) \to \inf, \tag{5}$$
$$h_j(x) = 0, \quad j \in J = \{1, \ldots, s\}. \tag{6}$$

Предположим, что функции f и h_j, $j \in J$, непрерывно дифференцируемы на $\mathbb{R}^n$. Зафиксируем ненулевой вектор $y \in \mathbb{R}^s$. Для любых $x \in \mathbb{R}^n$ и $\varepsilon \geqslant 0$ определим функцию

$$\Delta(x, \varepsilon) = \sum_{j=1}^{s} (h_j(x) - \varepsilon y_j)^2.$$

Введём штрафную функцию

$$F_\lambda(x,\varepsilon)=\begin{cases} f(x), & \text{если } \varepsilon=0 \text{ и } \Delta(x,0)=0,\\ f(x)+\frac{1}{\varepsilon}\Delta(x,\varepsilon)+\lambda\varepsilon, & \text{если } \varepsilon>0,\\ +\infty, & \text{если } \varepsilon=0 \text{ и } \Delta(x,0)\neq 0. \end{cases}$$

В отличие от описанных выше штрафных функций, функция $F_\lambda(x,\varepsilon)$ зависит от дополнительного неотрицательного параметра ε. Введение данного параметра позволяет добиться гладкости функции $F_\lambda(x,\varepsilon)$ при всех положительных ε.

Оказывается, что достаточные условия точности штрафной функции $F_\lambda(x,\varepsilon)$ совпадают с хорошо известными достаточными условиями точности ℓ_1 штрафной функции. Например, справедливы следующие утверждения [16].

ТЕОРЕМА 4. *Пусть x^* является точкой локального минимума в задаче* (5)–(6), *причём ограничения регулярны в этой точке. Тогда штрафная функция $F_\lambda(x,\varepsilon)$ является точной в точке $(x^*,0)$, то есть существует $\lambda^*\geqslant 0$ такое, что для любого $\lambda\geqslant\lambda^*$ точка $(x^*,0)$ является точкой локального минимума функции $F_\lambda(x,\varepsilon)$.*

ТЕОРЕМА 5. *Пусть $X\subset\mathbb{R}^n$ — открытое ограниченное множество, и предположим, что условие регулярности выполнено в каждой точке множества X. Тогда существует такое $\lambda^*\geqslant 0$, что для любого $\lambda\geqslant\lambda^*$ справедливы следующие утверждения:*

1) *если (x^*,ε^*) — точка локального минимума штрафной функции $F_\lambda(x,\varepsilon)$ и $x^*\in X$, то $\varepsilon^*=0$, то есть x^* — точка локального минимума в задаче* (5)–(6).

2) *если $x^*\in X$ — точка локального минимума в задаче* (5)–(6), *то точка $(x^*,0)$ является точкой локального минимума штрафной функции $F_\lambda(x,\varepsilon)$.*

В работе [16] также была предложена следующая принципиальная схема алгоритма минимизации штрафной функции $F_\lambda(x,\varepsilon)$. Зафиксируем $\delta_0>0$, $\lambda_0\geqslant 1$ и положим $k:=0$. Также зададим точность $\sigma>0$.

Шаг 1. Найдём точку (x_k,ε_k) такую, что

$$F_{\lambda_k}(x_k,\varepsilon_k)\leqslant\inf\big\{F_{\lambda_k}(x,\varepsilon)\mid(x,\varepsilon)\in\mathbb{R}^n\times[0,+\infty)\big\}+\delta_k.$$

Шаг 2. Если $\varepsilon_k=0$ и $\delta_k<\sigma$, то **Стоп**. Иначе перейдём на **Шаг 3**.

Шаг 3. Определим

$$\delta_{k+1} = \frac{1}{2}\delta_k, \quad \lambda_{k+1} = \begin{cases} \lambda_k, \text{ если } \varepsilon_k = 0, \\ \rho\lambda_k, \text{ если } \varepsilon_k > 0, \end{cases}$$

где $\rho > 1$ — константа.

Шаг 4. Положим $k := k + 1$ и перейдём на **Шаг 1**.

Поскольку штрафная функция $F_\lambda(x, \varepsilon)$ является непрерывно дифференцируемой на множестве $\mathbb{R}^n \times (0, +\infty)$, то на Шаге 1 приведённого алгоритма можно использовать любой метод оптимизации гладких функций. Можно показать, что при некоторых предположения описанный алгоритм сходится к решению задачи (5)–(6).

ЛИТЕРАТУРА

1. Долгополик М. В. *Точные штрафные функции в негладкой оптимизации* [Электронный ресурс] // Семинар «CNSA & NDO» [Официальный сайт]. Избранные доклады. 8.05.2014. URL: `http://www.apmath.spbu.ru/cnsa/pdf/DolgopolikM_ExactPenalty.pdf` (дата обращения: 6.11.2014).

2. Долгополик М. В. *Необходимые и достаточные условия для точности штрафной функции* [Электронный ресурс] // Семинар «CNSA & NDO» [Официальный сайт]. Избранные доклады. 18.09.2014. URL: `http://www.apmath.spbu.ru/cnsa/pdf/DolgopolikMV_18Sept2014.pdf` (дата обращения: 6.11.2014).

3. Warga J. *A necessary and sufficient condition for a constrained minimum* // SIAM Journal on Optimization, 1992. vol. 2. pp. 665–667.

4. Dedieu J. P. *Penalty Functions in Subanalytic Optimization* // Optim., 1992. vol. 26. pp. 27–32.

5. Luo Z.-Q., Pang J.-S., Ralph D., Wu S.-Q. *Exact Penalization and Stationarity Conditions of Mathematical Programs with Equilibrium Constraints* // Math. Program., 1996. vol. 75. pp. 19–76.

6. Luo Z.-Q., Pang J.-S., Ralph D. *Mathematical Programs with Equilibrium Constraints.* Cambridge University Press, Cambridge, United Kingdom, 1996. 401 p.

7. Lojasiewicz M. S. *Sur le promléme de la division* // Studia Mathematica, 1959. vol. 18. pp. 87–136.

8. Lojasiewicz M. S. *Ensembles semi-analytiques.* Institut des Hautes Etudes Scientifiques, Bures-sur-Yvette. 1964.

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

10. Di Pillo, G., Grippo L. *On the Exactness of a Class of Nondifferentiable Penalty Functions* // J. Optim. Theory Appl., 1988. vol. 57. pp. 399–410.

11. Di Pillo G., Grippo L. *Exact Penalty Functions in Constrained Optimization* // SIAM J. Control Optim., 1989. vol. 27. pp. 1333–1360.

12. Ye J. J. *The Exact Penalty Principle* Nonlinear Anal., 2012. vol. 75. pp. 1642–1654.

13. 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).

14. Di Pillo G., Facchinei F. *Exact Barrier Function Methods for Lipschitz Programs* // Appl. Math. Optim., 1995. vol. 32. pp. 1–31.

15. Huyer W., Neumaier A. *A New Exact Penalty function* // SIAM J. Optim., 2003. vol. 13. pp. 1141–1158.

16. Wang C., Ma C., Zhou J. *A New Class of Exact Penalty Functions and Penalty Algorithms* // J. Glob. Optim., 2014. vol. 58. pp. 51–73.

17. Li B., Yu C. J., Teo K. L., Duan G. R. *An Exact Penalty Function Method for Continuous Inequality Constrained Optimal Control Problem* // J. Optim. Theory Appl., 2011. vol. 151. pp. 260–291.

18. Jiang C., Lin Q., Yu C., Teo K. L., Duan G.-R. *An Exact Penalty Method for Free Terminal Time Optimal Control Problem with Continuous Inequality Constraints* // J. Optim. Theory Appl., 2012. vol. 154. pp. 30–53.

19. Lin Q., Loxton R., Teo K. L., Wu Y. H. *Optimal Feedback Control for Dynamic Systems with State Constraints: An Exact Penalty Approach* // Optim. Lett., 2014. vol. 8. pp. 1535–1551.