ЛЕКЦИЯ 12
Глава 4. МЕТОДЫ ДЛЯ ЗАДАЧ УСЛОВНОЙ ОПТИМИЗАЦИИ И КОМПЛЕМЕНТАРНЫХ ЗАДАЧ
(продолжение)

Содержание лекции

Задача оптимизации со смешанными ограничениями

Задача

$$f(x) \to \min, \quad x \in D, \tag{1}$$

$$D = \{x \in \mathbb{R}^n \mid F(x) = 0,\ G(x) \leqslant 0\}, \tag{2}$$

где

- $f : \mathbb{R}^n \to \mathbb{R}$ — заданная гладкая функция;
- $F : \mathbb{R}^n \to \mathbb{R}^l$ и $G : \mathbb{R}^n \to \mathbb{R}^m$ — заданные гладкие отображения.

Цель раздела: распространение полученных в пп. 4.2.2, 4.2.3 для задачи с ограничениями-равенствами результатов о штрафах и модифицированных функциях Лагранжа (и соответствующих методах) на задачу (1), (2), содержащую также ограничения-неравенства.

Задача оптимизации со смешанными ограничениями

Сведение к ограничениям-равенствам через «слэки»

$$f(x) \to \min, \quad (x, \sigma) \in \tilde{D}, \tag{3}$$

$$\tilde{D} = \{(x, \sigma) \in \mathbb{R}^n \times \mathbb{R}^m \mid F(x) = 0,\ g_i(x) + \sigma_i^2 = 0,\ i = 1, \ldots, m\}. \tag{4}$$

Здесь этот прием будет полезен. В частности, он естественным образом приведет к конструкции модифицированной функции Лагранжа для задачи (1), (2).

Если $(x, \sigma) \in \tilde{D}$, то:

- σ определяется с точностью до знаков компонент;
- эти знаки не влияют на значение целевой функции задачи (3), (4), а значит, точки $(x, \sigma^1) \in \tilde{D}$ и $(x, \sigma^2) \in \tilde{D}$, отличающиеся лишь знаками компонент σ^1 и $\sigma^2 \in \mathbb{R}^m$, можно отождествлять.

Задача оптимизации со смешанными ограничениями

С учетом этой договоренности, точка $\bar{x} \in \mathbb{R}^n$ является локальным решением задачи (1), (2) тогда и только тогда, когда точка $(\bar{x}, \bar{\sigma})$, $\bar{\sigma} = (\sqrt{-g_1(\bar{x})}, \ldots, \sqrt{-g_m(\bar{x})})$, корректно определена и является локальным решением задачи (3), (4), причем других локальных решений вида $(\bar{x}, \sigma)$, $\sigma \in \mathbb{R}^m$, у этой задачи нет.

Штрафы

Штраф для множества $D \subset \mathbb{R}^n$

Это любая функция $\psi : \mathbb{R}^n \to \mathbb{R}$ такая, что

$$\psi(x) = 0 \quad \forall\, x \in D, \quad \psi(x) > 0 \quad \forall\, x \in \mathbb{R}^n \setminus D. \tag{5}$$

Семейство _штрафных функций с параметром штрафа_ $c \geqslant 0$

$$\varphi_c : \mathbb{R}^n \to \mathbb{R}, \quad \varphi_c(x) = f(x) + c\psi(x). \tag{6}$$

Методы штрафа

Метод штрафа

Состоит в последовательном решении задач безусловной оптимизации

$$\varphi_c(x) \to \min, \quad x \in \mathbb{R}^n, \tag{7}$$

аппроксимирующих исходную задачу при $c \to +\infty$.

Методы штрафа

Алгоритм 1

Фиксируем штраф, т.е. функцию $\psi : \mathbb{R}^n \to \mathbb{R}$, удовлетворяющую (5). Выбираем последовательность $\{c_k\} \subset \mathbb{R}_+$ удовлетворяющую условию $c_k \to \infty$ $(k \to \infty)$, и полагаем $k = 0$.

1. Вычисляем $x^k \in \mathbb{R}^n$ как стационарную точку задачи (7) с целевой функцией, задаваемой формулой (6) при $c = c_k$. (Если функция ψ негладкая, то нужно уточнять понятие стационарности.)
2. Увеличиваем номер шага k на 1 и переходим к п. 1.

«Локальная» сходимость

Теорема 1 (обобщение теоремы 4.2.3)

Пусть функция $f : \mathbb{R}^n \to \mathbb{R}$ непрерывна в окрестности точки $\bar{x} \in \mathbb{R}^n$, а $D \subset \mathbb{R}^n$ — заданное множество. Пусть точка $\bar{x} \in \mathbb{R}^n$ является строгим локальным решением задачи (1).
Тогда если функция $\psi : \mathbb{R}^n \to \mathbb{R}$ удовлетворяет (5) и непрерывна в окрестности точки $\bar{x}$, то найдется такое число $\delta > 0$, что для $c \geqslant 0$ и для любого (глобального) решения $x(c)$ задачи

$$\varphi_c(x) \to \min, \quad x \in B(\bar{x}, \delta),$$

целевая функция которой задается формулой (6), имеет место

$$x(c) \to \bar{x} \quad (c \to +\infty).$$

«Локальная» сходимость

При этом:

- из теоремы Вейерштрасса следует, что для любого c такая точка $x(c)$ существует;
- для любого достаточно большого числа c точка $x(c)$ является локальным решением задачи (7);
- ни специфика ограничений, задающих множество D, ни специфика штрафа не играют роли.

Степенные штрафы

Степенные штрафы (степени $p > 0$)

$$\psi(x) = \|\Psi(x)\|^p, \quad x \in \mathbb{R}^n, \tag{8}$$

где

$$\Psi : \mathbb{R}^n \to \mathbb{R}^l \times \mathbb{R}^m, \quad \Psi(x) = (F(x), \max\{0, G(x)\}).$$

В (8) евклидову норму удобно заменить на $\|\cdot\|_p$, либо $\|\cdot\|_\infty$:

$$\psi(x) = \|\Psi(x)\|_p^p = \|F(x)\|_p^p + \|\max\{0, G(x)\}\|_p^p, \quad x \in \mathbb{R}^n,$$

либо

$$\psi(x) = \|\Psi(x)\|_\infty^p =$$

$$= (\max\{\|F(x)\|_\infty, 0, g_1(x), \ldots, g_m(x)\})^p, \quad x \in \mathbb{R}^n.$$

Квадратичный штраф

Квадратичный штраф

Это степенной штраф с $\|\cdot\|_p$ при $p = 2$:

$$\psi(x) = \frac{1}{2}\|\Psi(x)\|^2 = \frac{1}{2}(\|F(x)\|^2 + \|\max\{0,\, G(x)\}\|^2), \quad x \in \mathbb{R}^n. \tag{9}$$

Алгоритм 1 с таким штрафом — *метод квадратичного штрафа*.

Отличие от случая ограничений-равенств: штрафная функция φ_c, определяемая в (6), (9), дифференцируема, но может не иметь второй производной в точках $x \in \mathbb{R}^n$, в которых $g_i(x) = 0$ хотя бы для одного $i \in \{1,\, \ldots,\, m\}$. Это нужно учитывать при выборе методов безусловной оптимизации для соответствующей подзадачи (7).

Метод квадратичного штрафа

Семейство штрафных функций задачи (3), (4) (с ограничениями-равенствами)

$$\tilde{\varphi}_c : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}, \quad \tilde{\varphi}_c(x, \sigma) = f(x) + c\tilde{\psi}(x, \sigma),$$

с квадратичным штрафом

$$\tilde{\psi} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}, \quad \tilde{\psi}(x, \sigma) = \frac{1}{2}\left(\|F(x)\|^2 + \sum_{i=1}^{m}(g_i(x) + \sigma_i^2)^2\right).$$

Подзадача метода квадратичного штрафа для задачи (3), (4)

$$\tilde{\varphi}_c(x, \sigma) \to \min, \quad (x, \sigma) \in \mathbb{R}^n \times \mathbb{R}^m. \tag{10}$$

Метод квадратичного штрафа

Глобальный минимум по $\sigma \in \mathbb{R}^m$ в задаче (10) достигается в (единственной с точностью до знаков компонент) точке $\sigma(x) \in \mathbb{R}^m$,

$$\sigma_i(x) = \begin{cases} 0, \text{ если } g_i(x) > 0, \\ \sqrt{-g_i(x)}, \text{ иначе}, \end{cases} \qquad i = 1, \ldots, m.$$

Поэтому

$$\min_{\sigma \in \mathbb{R}^m} \tilde{\varphi}_c(x, \sigma) = \tilde{\varphi}_c(x, \sigma(x)) =$$
$$= f(x) + \frac{c}{2} \left(\|F(x)\|^2 + \|\max\{0, G(x)\}\|^2 \right) = \varphi_c(x) \quad \forall\, x \in \mathbb{R}^n.$$

Это дает возможность выводить теоретические результаты для метода квадратичного штрафа из соответствующих результатов для задачи с ограничениями-равенствами, полученных в п. 4.2.2.

Метод квадратичного штрафа

В результате о сходимости понадобится

Условие линейной независимости в точке $x \in \mathbb{R}^n$

Строки матрицы $F'(x)$ и векторы $g_i'(x)$, $i \in A(x) \cup A^+(x)$ образуют линейно-независимую систему, где $A^+(x) = \{i = 1, \ldots, m \mid g_i(x) > 0\}$.
Если $x \in D$, то $A^+(x) = \varnothing$, и такое условие линейной независимости совпадает с обычным.

«Глобальная» сходимость

Теорема 2 (обобщение теоремы 4.2.2)

Пусть функция $f : \mathbb{R}^n \to \mathbb{R}$ и отображения $F : \mathbb{R}^n \to \mathbb{R}^l$ и $G : \mathbb{R}^n \to \mathbb{R}^m$ непрерывно дифференцируемы на $\mathbb{R}^n$.
Тогда если последовательность $\{x^k\}$ сгенерирована алгоритмом 1 с функцией $\psi : \mathbb{R}^n \to \mathbb{R}$, введенной в (9), то любая ее предельная точка $\bar{x} \in \mathbb{R}^n$, в которой выполнено условие линейной независимости, является стационарной точкой задачи (1), (2), и для любой сходящейся к $\bar{x}$ подпоследовательности $\{x^{k_j}\}$ справедливо

$$\{c_{k_j} F(x^{k_j})\} \to \bar{\lambda} \quad (j \to \infty),$$

$$c_{k_j} \max\{0,\, g_i(x^{k_j})\} \to \bar{\mu}_i \quad (j \to \infty) \quad \forall\, i = 1, \ldots, m,$$

где $\bar{\lambda} \in \mathbb{R}^l$ и $\bar{\mu} \in \mathbb{R}^m_+$ — (единственные) отвечающие $\bar{x}$ множители Лагранжа.

«Локальная» сходимость и скорость сходимости

В формулировке следующей теоремы:

$$L : \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m \to \mathbb{R}, \quad L(x, \lambda, \mu) = f(x) + \langle \lambda, F(x) \rangle + \langle \mu, G(x) \rangle,$$

— функция Лагранжа задачи (1), (2), а

$$C(\bar{x}) =$$

$$= \{\xi \in \mathbb{R}^n \mid h'(\bar{x})\xi = 0, \ \langle g_i'(\bar{x}), \xi \rangle \leqslant 0, \ i \in A(\bar{x}), \ \langle f'(\bar{x}), \xi \rangle \leqslant 0\}$$

— критический конус задачи (1), (2) в точке $\bar{x}$.

«Локальная» сходимость и скорость сходимости

Теорема 3 (обобщение теоремы 4.2.4)

Пусть функция $f : \mathbb{R}^n \to \mathbb{R}$ и отображения $F : \mathbb{R}^n \to \mathbb{R}^l$ и $G : \mathbb{R}^n \to \mathbb{R}^m$ дважды дифференцируемы в некоторой окрестности точки $\bar{x} \in \mathbb{R}^n$, причем их вторые производные непрерывны в этой точке. Пусть в точке $\bar{x}$ выполнено условие линейной независимости, причем $\bar{x}$ — стационарная точка задачи (1), (2), а $\bar{\lambda} \in \mathbb{R}^l$ и $\bar{\mu} \in \mathbb{R}^m_+$ — однозначно отвечающие ей множители Лагранжа. Пусть в точке $\bar{x}$ выполнено условие строгой дополнительности и достаточное условие второго порядка оптимальности

$$\left\langle \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda}, \bar{\mu})\xi, \xi \right\rangle > 0 \quad \forall \xi \in C(\bar{x}) \setminus \{0\}. \tag{11}$$

«Локальная» сходимость и скорость сходимости

Теорема 3 (продолжение)

Тогда существуют окрестность U точки $\bar{x}$ и числа $\bar{c} \geqslant 0$ и $M > 0$ такие, что для каждого $c > \bar{c}$ задача (7) с целевой функцией, задаваемой формулами (6), (9), имеет в U единственную стационарную точку $x(c)$, причем

$$\|x(c) - \bar{x}\| \leqslant M\frac{\|\bar{\lambda}\| + \|\bar{\mu}\|}{c}, \quad \|cF(x(c)) - \bar{\lambda}\| \leqslant M\frac{\|\bar{\lambda}\| + \|\bar{\mu}\|}{c},$$

$$\|c \max\{0,\ G(x(c))\} - \bar{\mu}\| \leqslant M\frac{\|\bar{\lambda}\| + \|\bar{\mu}\|}{c}.$$

Учет прямого ограничения

Допустимое множество

$$D = \{x \in P \mid F(x) = 0,\ G(x) \leqslant 0\}, \tag{12}$$

где $P \subset \mathbb{R}^n$ — «простое» множество.

Простые ограничения можно не включать в штраф, а оставлять как прямые:

Штраф для множества D из (12)

Это любая функция $\psi : P \to \mathbb{R}$ такая, что

$$\psi(x) = 0 \quad \forall\, x \in D, \qquad \psi(x) > 0 \quad \forall\, x \in P \setminus D.$$

Учет прямого ограничения

Семейство *штрафных функций*

$$\varphi_c : P \to \mathbb{R}, \quad \varphi_c(x) = f(x) + c\psi(x).$$

Семейство подзадач метода штрафа

$$\varphi_c(x) \to \min, \quad x \in P.$$

К ней применимы методы оптимизации при простых ограничениях из разд. 4.1.