ЛЕКЦИЯ 11
Глава 4. МЕТОДЫ ДЛЯ ЗАДАЧ УСЛОВНОЙ ОПТИМИЗАЦИИ И КОМПЛЕМЕНТАРНЫХ ЗАДАЧ
(продолжение)

Содержание лекции

1. Методы активного множества
   - Идентификация активных ограничений
   - Оценки расстояния до множества решений

Стратегии активного множества

Идея для задачи, включающей в себя неравенства: идентификация неравенств, активных (т.е. выполняющихся как равенства) в искомом решении, и учет их как равенств на последующих итерациях.

С равенствами проще иметь дело с вычислительной точки зрения.

Многие методы автоматически идентифицируют активные неравенства при достаточном приближении к решению, по крайней мере в случае выполнения условия строгой дополнительности. Но эти методы сами по себе не выдают надежного «сертификата» («знака»), что правильная идентификация наступила.

Здесь речь идет о намеренной и явной идентификации.

Методы, включающие в себя такие процедуры и явно использующие оценки множеств активных неравенств, — *методы активного множества*.

Стратегии активного множества

Часто используются эвристические процедуры идентификации.
Здесь: теоретически обоснованные процедуры идентификации, использующие оценки расстояния до множества решений.

В случае комплементарных задач после идентификации активных неравенств могут применяться ньютоновские методы для уравнений, а в случае задач оптимизации — методы для задач оптимизации с ограничениями-равенствами.
Ограничимся случаем задачи оптимизации с ограничениями-неравенствами; наличие ограничений-равенств не создает дополнительных трудностей в этом контексте.
Для комплементарных задач принципы те же.

Стратегии активного множества

Типичные примеры методов активного множества для задач оптимизации:

- симплекс-метод для задач линейного программирования;
- метод особых точек и другие подобные методы для задач квадратичного программирования.

Эти методы — конечные: после правильной идентификации активного множества сразу находится точное решение (как решение соответствующей системы линейных уравнений). Для более общих (нелинейных, неквадратичных) задач оптимизации методы активного множества — бесконечношаговые.

Задача оптимизации

Задача

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

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

где

- $f : \mathbb{R}^n \to \mathbb{R}$ — заданная функция;
- $G : \mathbb{R}^n \to \mathbb{R}^m$ — заданное отображение.

Стратегии активного множества

Если для искомой стационарной точки $\bar{x} \in \mathbb{R}^n$ задачи (1), (2) идентифицировать $A(\bar{x}) = \{i = 1, \ldots, m \mid g_i(\bar{x}) = 0\}$, то $\bar{x}$ будет стационарной в задаче оптимизации с ограничениями-равенствами

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

$$\bar{D} = \{x \in \mathbb{R}^n \mid g_i(x) = 0,\ i \in A(\bar{x})\}.$$

Идентифицировать $A(\bar{x})$ нужно без точного знания $\bar{x}$: можно использовать информацию, доступную в имеющемся приближении $x \in \mathbb{R}^n$ к $\bar{x}$, а также (возможно) имеющееся приближение $\mu \in \mathbb{R}^m$ к некоторому отвечающему $\bar{x}$ множителю Лагранжа.

Локальная идентификация

Система Каруша–Куна–Таккера задачи (1), (2)

$$\frac{\partial L}{\partial x}(x,\, \mu) = 0, \tag{3}$$

$$\mu \geqslant 0, \quad G(x) \leqslant 0, \quad \langle \mu,\, G(x) \rangle = 0, \tag{4}$$

где

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

— функция Лагранжа.

Множество отвечающих стационарной точке $\bar{x}$ множителей Лагранжа

$$\mathcal{M}(\bar{x}) = \{\mu \in \mathbb{R}^m \mid (\bar{x},\, \mu) \text{ удовлетворяет } (3),\, (4)\}.$$

Локальная идентификация

Если для данного $\bar{\mu} \in \mathcal{M}(\bar{x})$ выполнено условие строгой дополнительности

$$\bar{\mu}_i > 0 \quad \forall\, i \in A(\bar{x}),$$

то локальная идентификация не составляет труда: найдется окрестность U точки $(\bar{x},\, \bar{\mu})$ такая, что

$$A(\bar{x}) = A(x,\, \mu) \quad \forall\, (x,\, \mu) \in U,$$

где

$$A(x,\, \mu) = \{i = 1,\, \ldots,\, m \mid \mu_i \geqslant -g_i(x)\}, \quad x \in \mathbb{R}^n,\, \mu \in \mathbb{R}^m.$$

Локальная идентификация

Доказательство

Если $i \in A(\bar{x})$, то $g_i(\bar{x}) = 0$ и $\bar{\mu}_i > 0$, и поэтому для любого (x, μ), достаточно близкого к $(\bar{x}, \bar{\mu})$, имеет место $\mu_i \geqslant -g_i(x)$, а значит, $i \in A(x, \mu)$, т.е. $A(\bar{x}) \subset A(x, \mu)$.

Если $i \in \{1, \ldots, m\} \setminus A(\bar{x})$, то $g_i(\bar{x}) < 0$ и $\bar{\mu}_i = 0$, и поэтому для любого (x, μ), достаточно близкого к $(\bar{x}, \bar{\mu})$, имеет место $\mu_i < -g_i(x)$, а значит, $i \notin A(x, \mu)$, т.е. $A(\bar{x}) \supset A(x, \mu)$.

Локальная идентификация

Без строгой дополнительности можно гарантировать лишь

$$A(x, \mu) \subset A(\bar{x}) \quad \forall\, (x, \mu) \in U.$$

Если $\mathcal{M}(\bar{x}) = \{\bar{\mu}\}$ (т.е. выполнено строгое условие Мангасариана–Фромовица), то

$$A_{+}(\bar{x}) \subset A(x, \mu) \subset A(\bar{x}) \quad \forall\, (x, \mu) \in U,$$

где

$$A_{+}(\bar{x}) = \{i \in A(\bar{x}) \mid \bar{\mu}_i > 0\}.$$

Далее ни строгая дополнительность, ни строгое условие Мангасариана–Фромовица не предполагаются.

Оценки расстояния

Идея: сравнивать значения $-g_i(x)$ не с μ_i, $i = 1, \ldots, m$, а со значениями «идентифицирующей функции», в основе определения которой лежит некоторая оценка расстояния до $\{\bar{x}\} \times \mathcal{M}(\bar{x})$.

Функция $r : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}_+$ *оценивает расстояние* до $\{\bar{x}\} \times \mathcal{M}(\bar{x})$ вблизи $\bar{\mu} \in \mathcal{M}(\bar{x})$, если $r(\bar{x}, \mu) = 0$ для всех $\mu \in \mathcal{M}(\bar{x})$, r непрерывна на $\{\bar{x}\} \times \mathcal{M}(\bar{x})$, и существуют $c > 0$, $\nu > 0$ и окрестность U точки $(\bar{x}, \bar{\mu})$ такие, что

$$\operatorname{dist}((x, \mu), \{\bar{x}\} \times \mathcal{M}(\bar{x})) \leqslant c(r(x, \mu))^\nu \quad \forall\, (x, \mu) \in U. \tag{5}$$

Константы c и ν обычно неизвестны, но их знание для идентификации и не требуется.

Оценки расстояния

Оценки расстояния до решений и их использование

- условия квадратичного роста и связанные с ними оценки в п. 3.1.2;
- теорема 4.4.2 об оценке расстояния до изолированного нуля негладкого отображения;
- теорема 4.6.6 будет использовать оценку расстояния до допустимого множества для анализа скорости сходимости методов штрафа.

Локальная идентификация

Положим

$$A(x,\, \mu) = \{i = 1,\, \ldots,\, m \mid \rho(r(x,\, \mu)) \geqslant -g_i(x)\}, \quad x \in \mathbb{R}^n,\, \mu \in \mathbb{R}^m, \tag{6}$$

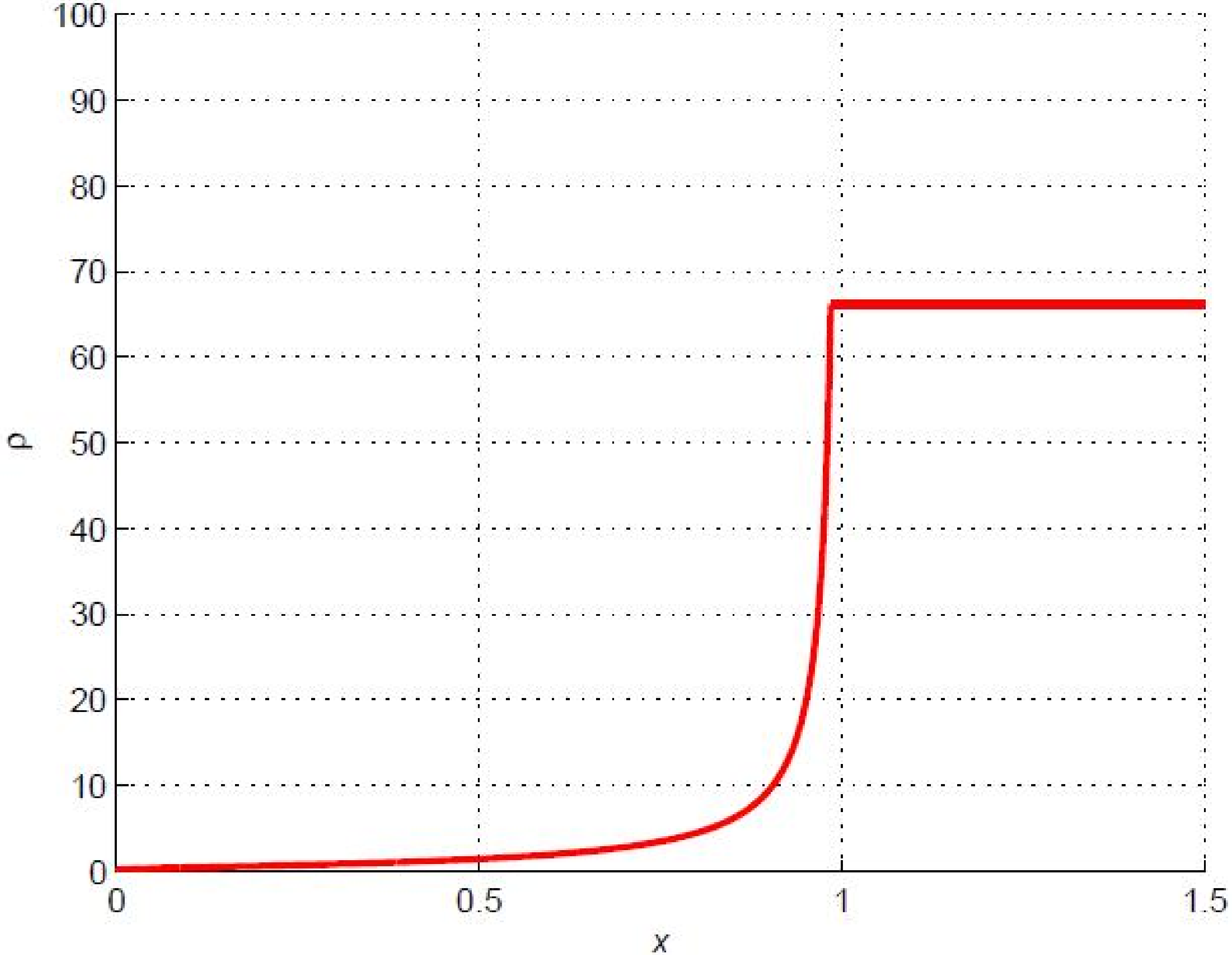
с функцией

$$\rho : \mathbb{R}_+ \to \mathbb{R}, \quad \rho(t) = \begin{cases} -1/\ln \hat{t}, \text{ если } t \geqslant \hat{t}, \\ -1/\ln t, \text{ если } t \in (0,\, \hat{t}), \\ 0, \text{ если } t = 0, \end{cases} \tag{7}$$

где $\hat{t} \in (0,\, 1)$ — параметр.

Эта функция стремится к нулю при $t \to 0+$ медленнее, чем t^ν при любом $\nu > 0$.

Локальная идентификация



Локальная идентификация

Предложение 1

Пусть функция $f : \mathbb{R}^n \to \mathbb{R}$ и отображение $G : \mathbb{R}^n \to \mathbb{R}^m$ дифференцируемы в точке $\bar{x} \in \mathbb{R}^n$. Пусть $\bar{x}$ — стационарная точка задачи (1), (2), причем функция $r : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}_+$ оценивает расстояние до множества $\{\bar{x}\} \times \mathcal{M}(\bar{x})$ вблизи $\bar{\mu} \in \mathcal{M}(\bar{x})$.

Тогда найдется окрестность U точки $(\bar{x}, \bar{\mu})$ такая, что для заданного формулами (6), (7) множества индексов справедливо

$$A(x, \mu) = A(\bar{x}) \quad \forall\, (x, \mu) \in U. \tag{8}$$

Доказательство

Аналогично случаю строгой дополнительности, но с использованием липшицевости G относительно $\bar{x}$ и соотношения $\lim_{t \to 0+} t^{\nu}/\rho(t) = 0$ для любого $\nu > 0$.

Локальная идентификация

Если функция r оценивает расстояние до $\{\bar{x}\} \times \mathcal{M}(\bar{x})$ вблизи любого $\bar{\mu} \in \mathcal{M}(\bar{x})$, и в точке $\bar{x}$ выполнено условие регулярности Мангасариана–Фромовица, то существует $\delta > 0$ такое, что (8) имеет место при

$$U = \{(x, \mu) \in \mathbb{R}^n \times \mathbb{R}^m \mid \operatorname{dist}((x, \mu), \{\bar{x}\} \times \mathcal{M}(\bar{x})) \leqslant \delta\}.$$

Это следует из компактности $\mathcal{M}(\bar{x})$.

Если $\mathcal{M}(\bar{x}) = \{\bar{\mu}\}$ (т.е. выполнено строгое условие Мангасариана–Фромовица), то можно идентифицировать и множество $A_+(\bar{x})$, а значит и $A_0(\bar{x}) = A(\bar{x}) \setminus A_+(\bar{x})$, и сразу положить $\mu_i = 0$, $i \in A_0(\bar{x})$.

Локальная идентификация

Положим

$$A_+(x,\,\mu) = \{i = 1,\,\ldots,\,m \mid \mu_i \geqslant \rho(r(x,\,\mu))\},\quad x \in \mathbb{R}^n,\,\mu \in \mathbb{R}^m. \tag{9}$$

Предложение 2

В предположениях предложения 1, если стационарной точке $\bar{x}$ задачи (1), (2) отвечает единственный множитель Лагранжа $\bar{\mu}$, то найдется окрестность U точки $(\bar{x},\,\bar{\mu})$ такая, что для заданного формулами (7), (9) множества индексов $A_+(x,\,\mu)$ справедливо

$$A_+(x,\,\mu) = A_+(\bar{x}) \quad \forall\,(x,\,\mu) \in U.$$

Локальная идентификация

Вместо функции ρ из (7) в предложениях 1 и 2 можно использовать

$$\rho : \mathbb{R}_+ \to \mathbb{R}, \quad \rho(t) = t^\theta,$$

где $\theta \in (0, \nu)$. При этом также выполняется ключевое свойство $\lim_{t \to 0+} t^\nu / \rho(t) = 0$.

Такой способ предполагает знание ν в (5).

Оценки расстояния

Обычно оценивающая расстояние до множества $\{\bar{x}\} \times \mathcal{M}(\bar{x})$ функция r определяется через невязку системы Каруша–Куна–Таккера (3), (4):

$$r(x,\, \mu) = \|\Psi(x,\, \mu)\|, \quad x \in \mathbb{R}^n,\, \mu \in \mathbb{R}^m, \tag{10}$$

где отображение $\Psi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^s$ непрерывно и таково, что множество решений системы (3), (4) совпадает с множеством решений уравнения

$$\Psi(x,\, \mu) = 0.$$

Оценки расстояния

Согласно теореме 4.4.2, для r из (10) оценка (5) имеет место при $\nu = 1$, если Ψ локально липшицево и дифференцируемо в точке $(\bar{x}, \bar{\mu})$ по любому направлению, и

$$\{(\xi, \zeta) \in \mathbb{R}^n \times \mathbb{R}^m \mid \Psi'((\bar{x}, \bar{\mu}); (\xi, \zeta)) = 0\} = \{(0, 0)\}. \qquad (11)$$

Подразумевает, что $\mathcal{M}(\bar{x}) = \{\bar{\mu}\}$.

Если $\psi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ — функция дополнительности, то можно взять $s = n + m$ и

$$\Psi(x, \mu) = \left(\frac{\partial L}{\partial x}(x, \mu), \psi(\mu, -G(x))\right), \quad x \in \mathbb{R}^n, \mu \in \mathbb{R}^m. \quad (12)$$

Как и для комплементарных задач в разд. 4.4.4, для случаев, когда ψ — функция естественной невязки или функция Фишера–Бурмайстера, известны критерии выполнения (11).

Оценки расстояния

Более тонкий результат:

Теорема 1

Пусть функция $f : \mathbb{R}^n \to \mathbb{R}$ и отображение $G : \mathbb{R}^n \to \mathbb{R}^m$ дважды дифференцируемы в точке $\bar{x} \in \mathbb{R}^n$. Пусть $\bar{x}$ — стационарная точка задачи (1), (2), а $\bar{\mu} \in \mathbb{R}^m$ — отвечающий ей множитель Лагранжа, причем выполнено достаточное условие второго порядка оптимальности

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

где $C(\bar{x}) = \{\xi \in \mathbb{R}^n \mid \langle g_i'(\bar{x}), \xi\rangle \leqslant 0,\ i \in A(\bar{x}),\ \langle f'(\bar{x}), \xi\rangle \leqslant 0\}$ — критический конус задачи в точке $\bar{x}$.

Оценки расстояния

Теорема 1 (завершение)

Тогда для функции $r : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}_+$, вводимой согласно (10), (12), где ψ — функция естественной невязки или функция Фишера–Бурмайстера, имеет место оценка (5) при $\nu = 1$ для некоторого $c > 0$ и некоторой окрестности U точки $(\bar{x}, \bar{\mu})$.

Здесь не подразумевается единственность множителя Лагранжа.

Оценки расстояния

Оценки расстояния другого типа: для вещественно-аналитических f и G оценка (5) имеет место при некотором (неизвестном!) $\nu > 0$ для r из (10) при $s = n + 2m + 1$ и

$$\Psi(x, \mu) = \left(\frac{\partial L}{\partial x}(x, \mu), \min\{\mu, 0\}, \max\{G(x), 0\}, \langle \mu, G(x) \rangle \right)$$

без всяких условий регулярности.