ЛЕКЦИЯ № 1

от 10.02.14

Лектор:

http://www.math.nsc.ru/LBRT/k5/mo.html

1. Понятие экстремальной задачи

2. Элементы алгоритмической теории экстремальных задач

3. Классификация задач

Лагранжева теория двойственности

1. Определения

2. Теорема о седловой точке

3. Линейное программирование

4. Теория двойственности линейного программирования

ЛИТЕРАТУРА

[1] *Алексеева Е. В., Кутненко О. А., Плясунов А. В. Численные методы оптимизации. Новосибирск: НГУ, 2008.*

[2] Болтянский В. Г. Математические методы оптимального управления. М.:Наука, 1969.

[3] Васильев Ф. П. Методы оптимизации. М.: Факториал Пресс, 2002.

[4] *Глебов Н. И., Кочетов Ю. А., Плясунов А. В. Методы оптимизации. Новосибирск: НГУ, 2000.*

[5] Иоффе А. Д., Тихомиров В. М. Теория экстремальных задач. М.: Наука, 1974.

[6] *Ларин Р. М., Плясунов А. В., Пяткин А. В. Методы оптимизации. Примеры и задачи. Новосибирск: НГУ, 2003, 2009.*

[7] Мину М. Математическое программирование. М.: Наука, 1990.

ЛИТЕРАТУРА

[8] Моисеев Н. Н., Иванилов Ю. П., Столярова Е. М. Методы оптимизации. М.: Наука, 1978.

[9] Понтрягин Л. С. и др. Математическая теория оптимальных процессов. М.: Наука, 1976.

[10] Сухарев А. Г., Тимохов А. В., Федоров В.В. Курс методов оптимизации. М.: Физматлит, 2005.

[11] Схрейвер А. Теория линейного и целочисленного программирования. М.: Мир, 1991.

[12] Ху Т. Целочисленное программирование и потоки в сетях. М.: Мир, 1974.

[13] Юдин Д.Б., Гольштейн Е.Г. Линейное программирование. Теория, методы и приложения. М.:Наука, 1969.

ЭКСТРЕМАЛЬНАЯ (ОПТИМИЗАЦИОННАЯ) ЗАДАЧА (P)

Найти:

$$\min f(x) \tag{1}$$

при условии, что

$$\varphi_i(x) \le 0, i = \overline{1, m}, \tag{2}$$

$$x \in S \subseteq R^n \text{ или } Z^n \text{ или } B^n. \tag{3}$$

$x = (x_1, \ldots, x_n)$– вектор переменных;

f – целевая функция задачи;

$\varphi_i(x) \le 0, i = \overline{1, m}$, $x \in S$ – ограничения задачи.

Методы оптимизации $\equiv$ Теория оптимизации $\equiv$ Теория экстремальных задач $\equiv$ Математическое программирование

1. Теоретическое исследование вопросов существования оптимальных решений экстремальных задач.

2. Необходимые и/или достаточные условия экстремума.

3. Разработка численных методов решения.

4. Исследование сложности задач (классы PO, FPTAS, PTAS, APX, poly-APX, exp-APX, NPO).

Источник экстремальных задач: – экономика, техника и т.д. (см. слайды).

Цели лекционного курса:

– Изучение ряда базовых алгоритмов, которые используются для решения **конечномерных задач оптимизации**.

– Получение теоретических и концептуальных представлений, достаточных для понимания, оценки этих алгоритмов и, если необходимо, создания новых.

ОСНОВНЫЕ ОПРЕДЕЛЕНИЯ

Вектор $\boldsymbol{x}$ – допустимое решение задачи $\boldsymbol{P}$, если выполняются ограничения (2),(3).

$\boldsymbol{Q(P) = \{x \in R^n | \varphi_i(x) \leq 0, i = \overline{1, m}, x \in S\}}$ – множество допустимых решений задачи $\boldsymbol{P}$.

Оптимальное решение (глобальный минимум): любое допустимое решение задачи, на котором достигается минимум целевой функции $\boldsymbol{f}$ на множестве $\boldsymbol{Q(P)}$.

1. $\boldsymbol{g(x) = 0 \equiv g(x) \leq 0, -g(x) \leq 0}$.

$\boldsymbol{g(x) \leq 0 \equiv g(x) + y = 0}$, где $\boldsymbol{y \geq 0}$.

2. $\boldsymbol{\max_{x \in Q} g(x) \equiv \min_{x \in Q} -g(x)}$

ОСНОВНЫЕ ОПРЕДЕЛЕНИЯ

Задача оптимизации решена, если
– либо найдено её оптимальное решение,
– либо найден конечный инфинум целевой функции на множестве $\boldsymbol{Q(P)}$, в случае, когда оптимального решения не существует,
– либо доказано, что целевая функция неограничена снизу на множестве допустимых решений,
– либо установлено, что множество допустимых решений задачи $\boldsymbol{P}$ пусто.

КЛАССИФИКАЦИЯ ЗАДАЧ

В зависимости от природы множества $\boldsymbol{S}$ задачи оптимизации классифицируются как:

– **дискретные (комбинаторные)** — $\boldsymbol{S}$ конечно или счетно,

– **целочисленные** — $\boldsymbol{x} \in \boldsymbol{S} \subseteq \boldsymbol{Z}^n$,

– **булевы** — $\boldsymbol{x} \in \boldsymbol{S} \subseteq \boldsymbol{B}^n$,

– **вещественные (непрерывные)** — $\boldsymbol{x} \in \boldsymbol{S} \subseteq \boldsymbol{R}^n$,

– **бесконечномерные** — $\boldsymbol{S}$ подмножество гильбертова пространства.

КЛАССИФИКАЦИЯ ЗАДАЧ

Если $\boldsymbol{S} = \boldsymbol{R}^n$ или $\boldsymbol{Z}^n$ или $\boldsymbol{B}^n$, $(\boldsymbol{m} = \mathbf{0})$, то задача $\boldsymbol{P}$ – **задача безусловной оптимизации**. В противном случае говорят о **задаче условной оптимизации**.

Подробности в пособии (изучить самостоятельно) **Алексеева Е. В., Кутненко О. А., Плясунов А. В. Численные методы оптимизации. Новосибирск: НГУ, 2008.**

ОСНОВЫ АЛГОРИТМИЧЕСКОЙ ТЕОРИИ ЭКСТРЕМАЛЬНЫХ ЗАДАЧ

Вычислимость $\equiv$ Тип вычислительных устройств $\equiv$ Программы $\equiv$ Конечные наборы инструкций

Пример невычислимого вещественного числа

Идея: использовать 10 проблему Гильберта $\equiv$ разрешимы или нет диофантовы уравнения в целых числах?

$\boldsymbol{d = \sum_{n \in D} 4^{-n}}$, где $\boldsymbol{D}$ – гёделевские номера разрешимых диофантовых уравнений.

Иерархия сложностных классов экстремальных задач.

ВЫЧИСЛИМОСТЬ ВЕЩЕСТВЕННЫХ ФУНКЦИЙ

Алгебраическая вычислимость:

Данные представляются (аргумент x) точно и любая операция (+, -, *, /) выполняется точно за один шаг при вычислении $\boldsymbol{f(x)}$

Битовая вычислимость:

По заданному "хорошему"рациональному приближению $\boldsymbol{x}$ можно получить "хорошее"рациональное приближение $\boldsymbol{f(x)}$

ВЫЧИСЛИМОСТЬ ВЕЩЕСТВЕННЫХ ФУНКЦИЙ: ПРОБЛЕМЫ

Алгебраическая вычислимость:

Функции константы вычислимы (сравнить с примером)

Невычислимы трансцендентные функции. Например: e^x и $\sqrt{x}$.

Битовая вычислимость:

Все вычислимые функции должны быть непрерывны

БИТОВАЯ ВЫЧИСЛИМОСТЬ: УТОЧНЕНИЕ

Функция $\varphi(n) \in \{\frac{m}{2^k} \mid m \in Z, k \in N\}$ такая, что

$$\|\varphi(l) - x\| < 2^{-l}$$

называется представлением числа $x \in R$

Число x вычислимо тогда и только тогда, когда вычислимо его представление

Пиксель – шар радиуса $\varepsilon > 0$

Пиксель вычислим, если вычислимы координаты его центра и радиус. Например, $B(\frac{m}{2^k}, 2^{-l})$

Образ – набор пикселей заданного размера (разрешение)

Соответственно, образ вычислим, если он состоит из вычислимых пикселей

БИТОВАЯ ВЫЧИСЛИМОСТЬ: УТОЧНЕНИЕ

Ограниченное множество $\boldsymbol{S} \subseteq \boldsymbol{R}^n$ вычислимо, если вычислима следующая функция

$\boldsymbol{f(d,k) = 1}$, если $\boldsymbol{B(d, 2^{-k}) \bigcap S \neq \emptyset}$

$\boldsymbol{f(d,k) = 0}$, если $\boldsymbol{B(d, 2 * 2^{-k}) \bigcap S = \emptyset}$

$\boldsymbol{f(d,k) = 0}$ или $\boldsymbol{1}$, иначе

$\boldsymbol{d \in \{\frac{m}{2^k} \mid m \in Z, k \in N\}}$

Ограниченная функция определённая на ограниченном множестве вычислима, если вычислим её график

Лагранжева теория двойственности

Рассмотрим задачу $\boldsymbol{P}$ с произвольными функциями $\boldsymbol{f}$ и $\boldsymbol{\varphi_i}$:

$$\boldsymbol{f(x) \longrightarrow \min} \tag{1}$$

$$\boldsymbol{\varphi_i(x) \le 0, i = \overline{1, m}.} \tag{2}$$

Определение 1. Функцию

$$\boldsymbol{L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i \varphi_i(x),}$$

определенную при всех $\boldsymbol{x}$ и $\boldsymbol{\lambda}$, назовем функцией Лагранжа для задачи (1), (2).

Лагранжева теория двойственности

Определение 2. Пара $(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ называется седловой точкой функции Лагранжа, если

$$\boldsymbol{L(x^*, \lambda) \overset{(3)}{\leq} L(x^*, \lambda^*) \overset{(4)}{\leq} L(x, \lambda^*)\ \forall x \in R^n, \forall \lambda \geq 0.}$$

Пусть

$$\boldsymbol{g(x) = \sup_{\lambda \geq 0} L(x, \lambda)}$$

Тогда

$$\boldsymbol{g(x) = \begin{cases} f(x),\ x \in Q, \\ +\infty,\ x \notin Q. \end{cases} \Longrightarrow}$$

Лагранжева теория двойственности

Задача $\boldsymbol{P}$ эквивалентна следующей:

$$\boldsymbol{g(x) \longrightarrow \min}$$

Пусть

$$\boldsymbol{h(\lambda) = \inf_{x \in R^n} L(x, \lambda).}$$

Рассмотрим задачу $(\boldsymbol{D})$:

$$\boldsymbol{h(\lambda) \longrightarrow \max_{\lambda \geq 0}.}$$

$(\boldsymbol{D})$ – задача двойственная к прямой (или исходной) задаче $\boldsymbol{P}$. $\boldsymbol{\lambda_1, \ldots, \lambda_m}$ – двойственные переменные, а $\boldsymbol{x_1, \ldots, x_n}$ — прямые переменные.

Лагранжева теория двойственности

Если $\boldsymbol{x} \in \boldsymbol{Q}, \boldsymbol{\lambda} \geq \mathbf{0}$, то $\boldsymbol{x}$ — допустимое решение прямой задачи, а $\boldsymbol{\lambda}$ — допустимое решение двойственной задачи.

Лемма 1. (Слабая теорема двойственности).

$$\forall \boldsymbol{x} \in \boldsymbol{Q} \ \forall \boldsymbol{\lambda} \geq 0 \ (\boldsymbol{h}(\boldsymbol{\lambda}) \leq \boldsymbol{f}(\boldsymbol{x})).$$

Лемма 2. Если $\overline{\boldsymbol{x}} \in \boldsymbol{Q}$ и $\overline{\boldsymbol{\lambda}} \geq 0$ и $\boldsymbol{f}(\overline{\boldsymbol{x}}) = \boldsymbol{h}(\overline{\boldsymbol{\lambda}})$, то $\overline{\boldsymbol{x}}$ и $\overline{\boldsymbol{\lambda}}$ — оптимальные решения задачи $\boldsymbol{P}$ и $\boldsymbol{D}$, соответственно.

Лагранжева теория двойственности

Теорема 1. Вектора $\overline{x}, \overline{\lambda}$ — оптимальные решения прямой и двойственной задачи и $f(\overline{x}) = h(\overline{\lambda})$ тогда и только тогда, когда пара $(\overline{x}, \overline{\lambda})$ — седловая точка функции Лагранжа. При этом $L(\overline{x}, \overline{\lambda}) = f(\overline{x}) = h(\overline{\lambda})$.

Следствие 1. Пусть $\overline{x} \in Q(P)$, $\overline{\lambda} \geq 0$. Следующие утверждения эквивалентны:

1. Пара $(\overline{x}, \overline{\lambda})$ — седловая точка функции Лагранжа.

2. $f(\overline{x}) = h(\overline{\lambda})$.

3. $\min\limits_{x} \sup\limits_{\lambda \geq 0} L(x, \lambda) = \max\limits_{\lambda \geq 0} \inf\limits_{x} L(x, \lambda)$.

Лагранжева теория двойственности

Следствие 2. **Пусть** $\boldsymbol{x}^*, \overline{\boldsymbol{x}} \in \boldsymbol{Q}$, $\boldsymbol{\lambda}^*, \overline{\boldsymbol{\lambda}} \geq \mathbf{0}$.

Если пары $(\overline{\boldsymbol{x}}, \overline{\boldsymbol{\lambda}})$ и $(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ — седловые точки функции Лагранжа, то пары $(\overline{\boldsymbol{x}}, \boldsymbol{\lambda}^*)$ и $(\boldsymbol{x}^*, \overline{\boldsymbol{\lambda}})$ — также седловые точки функции Лагранжа, причем

$$\boldsymbol{L}(\boldsymbol{x}^*, \overline{\boldsymbol{\lambda}}) = \boldsymbol{L}(\overline{\boldsymbol{x}}, \boldsymbol{\lambda}^*) = \boldsymbol{L}(\overline{\boldsymbol{x}}, \overline{\boldsymbol{\lambda}}) = \boldsymbol{L}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*).$$

Линейное программирование (ЛП)

Задача линейного программирования (ЛП) в канонической форме:

$$w(x) = (c, x) \longrightarrow \min \tag{5}$$

$$Ax = b, \tag{6}$$

$$x \geq 0, \tag{7}$$

где $c = (c_j), x = (x_j) \in R^n$, $A = (a_{ij}) - (m \times n)$ матрица, $b = (b_i) \in R^m$, $m \leq n$, $rang(A) = m$.

$$Ax = b \equiv (a_i, x) = b_i, i = \overline{1, m},$$

$$Ax = b \equiv \sum_{j=1}^{n} A_j x_j = b$$

ЛП: двойственная задача

Задача $(\mathbf{5}) - (\mathbf{7})$ эквивалентна задаче

$$\boldsymbol{(c, x) \longrightarrow \min}$$

$$\begin{aligned} \boldsymbol{(a_i, x) - b_i \le 0,} &\qquad \boldsymbol{\lambda_i^1 \ge 0} \\ \boldsymbol{-(a_i, x) + b_i \le 0,} &\qquad \boldsymbol{\lambda_i^2 \ge 0} \\ \boldsymbol{-x_j \le 0.} &\qquad \boldsymbol{\mu_j \ge 0} \end{aligned}$$

Ее функция Лагранжа:

$$\boldsymbol{L(x, \lambda^1, \lambda^2, \mu) = (c, x) + (\lambda^1, Ax - b) + (\lambda^2, -Ax + b) + (\mu, -x) =}$$

$$\boldsymbol{= \Big(c + (\lambda^1 - \lambda^2)A - \mu, x\Big) - (\lambda^1 - \lambda^2, b).}$$

Следовательно, целевая функция двойственной задачи имеет вид:

ЛП: двойственная задача

$$h(\boldsymbol{\lambda}^1, \boldsymbol{\lambda}^2, \boldsymbol{\mu}) = \inf_{\boldsymbol{x}} \boldsymbol{L}(\boldsymbol{x}, \boldsymbol{\lambda}^1, \boldsymbol{\lambda}^2, \boldsymbol{\mu}) =$$

$$= \begin{cases} -(\boldsymbol{b}, \boldsymbol{\lambda}^1 - \boldsymbol{\lambda}^2), & \text{если } \boldsymbol{c} + (\boldsymbol{\lambda}^1 - \boldsymbol{\lambda}^2)\boldsymbol{A} - \boldsymbol{\mu} = \boldsymbol{0}, \\ -\infty, & \text{иначе.} \end{cases} \Longrightarrow$$

Двойственная задача

$$h(\boldsymbol{\lambda}^1, \boldsymbol{\lambda}^2, \boldsymbol{\mu}) \longrightarrow \sup_{\boldsymbol{\lambda}^1 \geq 0, \boldsymbol{\lambda}^2 \geq 0, \boldsymbol{\mu} \geq 0},$$

эквивалентна задаче ЛП:

$$-(\boldsymbol{b}, \boldsymbol{\lambda}^1 - \boldsymbol{\lambda}^2) \longrightarrow \max$$

$$\boldsymbol{c} + (\boldsymbol{\lambda}^1 - \boldsymbol{\lambda}^2)\boldsymbol{A} - \boldsymbol{\mu} = \boldsymbol{0} \equiv \boldsymbol{c} + (\boldsymbol{\lambda}^1 - \boldsymbol{\lambda}^2)\boldsymbol{A} \geq \boldsymbol{0}.$$

Умножим ограничения на $-\mathbf{1}$, обозначим $\boldsymbol{y} = -(\boldsymbol{\lambda}^1 - \boldsymbol{\lambda}^2)$

ЛП: двойственная задача

Получим

$$(\boldsymbol{b}, \boldsymbol{y}) \longrightarrow \max \tag{8}$$

$$\boldsymbol{y}\boldsymbol{A} \le \boldsymbol{c}. \tag{9}$$

Замечание 1. Для задач (5)-(7) и (8)-(9) выполняются все утверждения: л. 1, л. 2, теор. 1, следствия 1 — 2.

Теорема 2. Задача двойственная к задаче (8)-(9) совпадает с исходной задачей (5)-(7).

Доказательство. Задача (8)-(9) эквивалентна задаче

$$-(\boldsymbol{b}, \boldsymbol{y}) \longrightarrow \min$$

$$\boldsymbol{y}\boldsymbol{A} \le \boldsymbol{c}.$$

ЛП: двойственная задача

$$\boldsymbol{yA \le c} \equiv \text{системе неравенств } \boldsymbol{(y, A_j) - c_j \le 0}, \qquad \boldsymbol{x_j \ge 0}$$

(сопоставили каждому ограничению двойственную переменную (множитель Лагранжа))

Функция Лагранжа:

$$\boldsymbol{L(y, x) = -(b, y) + (x, yA - c) =}$$

$$\boldsymbol{-(b, y) + (Ax, y) - (x, c) = (Ax - b, y) - (c, x).}$$

Целевая функция двойственной задачи

$$\boldsymbol{h(x) = \inf_y L(y, x) =}$$

$$\begin{cases} \boldsymbol{-(c, x)}, & \text{если } \boldsymbol{Ax = b}, \\ \boldsymbol{-\infty}, & \text{иначе.} \end{cases} \implies$$

ЛП: двойственная задача

Задача

$$\max_{x \geq 0} h(x)$$

эквивалентна задаче ЛП:

$$-(c, x) \longrightarrow \max$$

$$Ax = b,$$

$$x \geq 0$$

или

$$\min (c, x)$$

$$Ax = b,$$

$$x \geq 0.$$

■

ЛП: двойственная задача

Прямая задача		Двойственная задача
$\boldsymbol{w(x) = \sum_{j=1}^{n} c_j x_j \to \min}$		$\boldsymbol{z(y) = \sum_{i=1}^{m} b_i y_i \to \max}$
$\boldsymbol{a_i x \geq b_i}$	$\boldsymbol{i \in I_1}$	$\boldsymbol{y_i \geq 0}$
$\boldsymbol{a_i x = b_i}$	$\boldsymbol{i \in I_2}$	$\boldsymbol{y_i -}$ своб.
$\boldsymbol{x_j \geq 0}$	$\boldsymbol{j \in J_1}$	$\boldsymbol{y A_j \leq c_j}$
$\boldsymbol{x_j -}$ своб.	$\boldsymbol{j \in J_2}$	$\boldsymbol{y A_j = c_j.}$

Упражнение.

Техника получения этой схемы: либо повторить выкладки, приведшие к задаче (8), (9), либо воспользоваться сводимостью общей задачи ЛП к задаче ЛП в канонической форме и применить готовый рецепт (задача (8), (9)).