Новосибирский государственный университет

Факультет информационных технологий

Курс: «Методы оптимизации», 2015 г.

Лектор: Иваненко Дмитрий Сергеевич

ЛЕКЦИЯ № 1

1. Введение в методы оптимизации

2. Понятие экстремальной задачи

3. Классификация задач

4. Лагранжева теория двойственности

ЛИТЕРАТУРА

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

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

[3] Васильев Ф. П. Численные методы решения экстремальных задач. М.: Наука, 1980.

[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.*

ЛИТЕРАТУРА

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

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

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

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

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

Введение в теорию эстремальных задач

Методы (теория) оптимизации $\equiv$

$\equiv$ Теория экстремальных задач $\equiv$

$\equiv$ Математическое программирование

Актуальность дисциплины связана с появлением ЭВМ и феноменом **переборных** задач. Такие задачи можно решить с помощью перебора всех возможных вариантов, число которых растет **экспоненциально** в зависимости от размеров задачи. Для большинства таких задач неизвестно **эффективных** алгоритмов решения, т. е. они являются **NP-полными** или **NP-трудными**.

Вопрос

$$\mathbf{P} = \mathbf{NP}?$$

до сих пор открыт.

Предмет теории:

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.

Введение в теорию эстремальных задач

Приложения:

- в технике: управление размерами и оптимизация структур, оптимальное планирование сложных технических или информационных систем, сетей трубопроводов, компьютеров

- в автоматике: оптимальное управление системами, управление призводством, робототехника

- в транспорте: логистика, управление транспортными потоками, составление расписаний

- а также: в экономике, промышленности, медицине и др.

Введение в теорию эстремальных задач

Приложения (в прикладной математике):

- в численном анализе: аппроксимация, решение линейных и нелинейных задач

- в исследовании операций и в комбинаторной оптимизации: алгоритмы для оптимизации технико-экономических систем, задач оптимального управления, задач на графах, задач целочисленного и булевого программированияи др.

- теория игр: невозможна без понятия седловой точки, методов решения экстремальных задач

Цели лекционного курса:

– Углублённое изучение теории линейного программирования

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

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

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.

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 инфимум 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 $\boldsymbol{S}$:

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

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

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

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

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

КЛАССИФИКАЦИЯ ЗАДАЧ

Если $(\boldsymbol{m} = \mathbf{0})$, т. е.

$$\boldsymbol{S} = \boldsymbol{R}^{\boldsymbol{n}} \text{ или } \boldsymbol{S} = \boldsymbol{Z}^{\boldsymbol{n}} \text{ или } \boldsymbol{S} = \boldsymbol{B}^{\boldsymbol{n}},$$

то задача $\boldsymbol{P}$ – **задача безусловной оптимизации**. В противном случае говорят о **задаче условной оптимизации**.

КЛАССИФИКАЦИЯ ЗАДАЧ

В зависимости от свойств функций $\boldsymbol{f}$ и $\boldsymbol{\varphi_i(x)}$ возникает более тонкое деление экстремальных задач на классы, например:

непрерывное математическое программирование: $\boldsymbol{f}$, $\boldsymbol{\varphi_i(x)}$ – непрерывные, произвольные, нелинейные, $\boldsymbol{S}$ – связное, компактное подмножество $\boldsymbol{R^n}$

КЛАССИФИКАЦИЯ ЗАДАЧ

выпуклое программирование: $\boldsymbol{f}$, $\boldsymbol{\varphi_i(x)}$ произвольные, выпуклые, $\boldsymbol{S \subset R^n}$ – выпуклое множество

линейное программирование: $\boldsymbol{f}$, $\boldsymbol{\varphi_i(x)}$ – линейные, $\boldsymbol{S = \{x \in R^n | Ax \le b\}}$

Подробнее (изучить самостоятельно):

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

Лагранжева теория двойственности

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

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

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

Определение 1. Функцию

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

определенную при всех x и λ, назовем функцией Лагранжа для задачи (1), (2).

Лагранжева теория двойственности

Пусть

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

Тогда

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

Следовательно, радача $\boldsymbol{P}$ эквивалентна задаче безусловной оптимизации:

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

Лагранжева теория двойственности

Пусть

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

Рассмотрим задачу (D):

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

Задача D называется **двойственной** к исходной (или прямой) задаче P. Переменные $\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_m$ называются **двойственными переменными**

Лемма 1. (Слабая теорема двойственности).

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

Доказательство. Пусть $\boldsymbol{\lambda} \geq \mathbf{0}$. По определению функции $\boldsymbol{h}(\boldsymbol{\lambda})$:

$$\boldsymbol{h}(\boldsymbol{\lambda}) \leq \boldsymbol{L}(\boldsymbol{x}, \boldsymbol{\lambda}) = \boldsymbol{f}(\boldsymbol{x}) + \sum_{i=1}^{m} \boldsymbol{\lambda}_i \boldsymbol{\varphi}_i(\boldsymbol{x}).$$

$\forall \boldsymbol{x} \in \boldsymbol{Q}$ справедливо $\boldsymbol{\lambda}_i \boldsymbol{\varphi}_i(\boldsymbol{x}) \leq \mathbf{0}$, следовательно:

$$\boldsymbol{h}(\boldsymbol{\lambda}) \leq \boldsymbol{f}(\boldsymbol{x}). \qquad \blacksquare$$

Лагранжева теория двойственности

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

Определение 2. Пара (x^*, λ^*) называется седловой точкой функции Лагранжа, если $\forall x \in R^n, \forall \lambda \geq 0$

$$L(x^*, \lambda) \overset{(3)}{\leq} L(x^*, \lambda^*) \overset{(4)}{\leq} L(x, \lambda^*).$$

Лагранжева теория двойственности

Теорема 1. **Векторы $\overline{x}, \overline{\lambda}$ — 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 $f(\overline{x}) = L(\overline{x}, \overline{\lambda}) = h(\overline{\lambda})$.**

Доказательство. *Необходимость.* Пусть $\overline{x}, \overline{\lambda}$ — оптимальные решения прямой и двойственной задачи.

$$h(\overline{\lambda}) \stackrel{def}{=} \inf_{x \in R^n} L(x, \overline{\lambda}) \leq L(\overline{x}, \overline{\lambda})$$

$$L(\overline{x}, \overline{\lambda}) \leq \sup_{\lambda \geq 0} L(\overline{x}, \lambda) \stackrel{def}{=} g(\overline{x}).$$

Лагранжева теория двойственности

В силу оптимальности, $\overline{x} \in Q$ и $f(\overline{x}) = g(\overline{x})$. Из равенства $f(\overline{x}) = h(\overline{\lambda})$ следует:

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

Тогда:

$$L(\overline{x}, \lambda) \leq \sup_{\lambda \geq 0} L(\overline{x}, \lambda) =$$

$$= L(\overline{x}, \overline{\lambda}) =$$

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

Лагранжева теория двойственности

Достаточность. Пусть $(\bar{x}, \bar{\lambda})$ — седловая точка функции Лагранжа. Тогда из (3) следует:

$$\sum_{i=1}^{m} (\lambda_i - \bar{\lambda}_i)\varphi_i(\bar{x}) \le 0 \quad \forall \lambda \ge 0.$$

Предположим, что $\exists\ i : 1 \le i \le m, \varphi_i(\bar{x}) > 0$. Тогда для достаточно большого $\lambda_i > 0$:

$$\sum_{i=1}^{m} (\lambda_i - \bar{\lambda}_i)\varphi_i(\bar{x}) > 0.$$

Получено противоречие, следовательно, $\bar{x} \in Q$.

При $\boldsymbol{\lambda} = \mathbf{0}$ имеем:

$$\sum_{i=1}^{m} \overline{\lambda}_i \varphi_i(\overline{x}) \geq 0,$$

откуда:

$$\overline{\lambda}_i \varphi_i(\overline{x}) = 0 \quad \forall i = \overline{1,\, m}.$$

Следовательно, $f(\overline{x}) = L(\overline{x}, \overline{\lambda})$.
Из (4) следует, что:

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

По Лемме 2, $\overline{x}$ и $\overline{\lambda}$ — оптимальные решения задачи прямой и двойственной задач, соответственно. ■

Лагранжева теория двойственности

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

Доказательство. Покажем эквивалентность первого и третьего утверждений. По теореме 1, пара $(\overline{x}, \overline{\lambda})$ — седловая точка функции Лагранжа тогда и только тогда, когда $\overline{x}$ и $\overline{\lambda}$ — оптимальные решения прямой и двойственной задач.

Лагранжева теория двойственности

Другими словами,

$$\min_{x \in Q} f(x) = f(\overline{x}) = L(\overline{x}, \overline{\lambda}) = h(\overline{\lambda}) = \max_{\lambda \geq 0} h(\lambda).$$

Поскольку

$$\min_{x \in Q} f(x) = \min_{x \in R^n} g(x),$$

то

$$\min_{x} \sup_{\lambda \geq 0} L(x, \lambda) = L(\overline{x}, \overline{\lambda}) = \max_{\lambda \geq 0} \inf_{x} L(x, \lambda).$$

Упражнение. Завершить доказательство Следствия 1.

Лагранжева теория двойственности

Следствие 2. Пусть $x^*, \overline{x} \in Q$, $\lambda^*, \overline{\lambda} \geq 0$. Если пары $(\overline{x}, \overline{\lambda})$ и (x^*, λ^*) — седловые точки функции Лагранжа, то пары $(\overline{x}, \lambda^*)$ и $(x^*, \overline{\lambda})$ — также седловые точки функции Лагранжа, причем

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

Упражнение. Доказать Следствие 2.