ЛЕКЦИЯ № 1

от 07.09.16

Лектор: Панин Артём (Александрович)

http://www.math.nsc.ru/LBRT/k5 Opt-FIT-2016.html (вместо - нижнее подчеркивание)

1. Понятие экстремальной задачи

2. Элементы алгоритмической теории экстремальных задач

3. Классификация задач

Лагранжева теория двойственности

1. Определения

2. Теорема о седловой точке

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

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

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

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

f – целевая функция задачи;

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

Источник экстремальных задач: – экономика, техника и д.р.

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.**

ЛАГРАНЖЕВА ТЕОРИЯ ДВОЙСТВЕННОСТИ

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

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

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

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

Пусть

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

Тогда

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

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).

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

Док-во: $\boldsymbol{h}(\boldsymbol{\lambda}) = \inf\limits_{\tilde{x} \in R^n} \boldsymbol{L}(\tilde{\boldsymbol{x}}, \boldsymbol{\lambda}) \leq \boldsymbol{L}(\boldsymbol{x}, \boldsymbol{\lambda}) \leq \sup\limits_{\tilde{\lambda} \geq 0} \boldsymbol{L}(\boldsymbol{x}, \tilde{\boldsymbol{\lambda}}) = \boldsymbol{f}(\boldsymbol{x})$.

■

Лемма 2. Если $\overline{\boldsymbol{x}} \in \boldsymbol{Q}$ и $\overline{\boldsymbol{\lambda}} \geq \mathbf{0}$ 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

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

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

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

Док-во:

Лагранжева теория двойственности

1) НЕОБХОДИМОСТЬ.

Пусть $\overline{\boldsymbol{x}}, \overline{\boldsymbol{\lambda}}$ – оптимальные решения прямой и двойственной задачи. Тогда:

$$\boldsymbol{f(\overline{x}) = \sup_{\lambda \geq 0} L(\overline{x}, \lambda) \geq L(\overline{x}, \overline{\lambda}) \geq \inf_{x \in R^n} L(x, \overline{\lambda}) = h(\overline{\lambda}).}$$

Но $\boldsymbol{f(\overline{x}) = h(\overline{\lambda})}$. Возьмем произвольные $\boldsymbol{x \in R^n}$, $\boldsymbol{\lambda \geq 0}$. Получается, что

$$\boldsymbol{L(\overline{x}, \lambda) \leq \sup_{\lambda \geq 0} L(\overline{x}, \lambda) = L(\overline{x}, \overline{\lambda}) = \inf_{x \in R^n} L(x, \overline{\lambda}) \leq L(x, \overline{\lambda}).}$$

Из последнего следуют (3) и (4).

Лагранжева теория двойственности

2) ДОСТАТОЧНОСТЬ.

Пусть $(\overline{\boldsymbol{x}}, \overline{\boldsymbol{\lambda}})$ – седловая точка функции Лагранжа. Тогда из (3) следует:

$$\sum_{i=1}^{m}(\boldsymbol{\lambda}_i - \overline{\boldsymbol{\lambda}}_i)\boldsymbol{\varphi}_i(\boldsymbol{x}) \leq \mathbf{0}, \forall \boldsymbol{\lambda} \geq \mathbf{0}.$$

Предположим, что $\overline{\boldsymbol{x}} \notin \boldsymbol{Q}$. Т.е. $\exists \boldsymbol{i} : \boldsymbol{\varphi}_i(\overline{\boldsymbol{x}}) > \mathbf{0}$. Тогда для достаточно большого $\boldsymbol{\lambda}_i > \mathbf{0}$:

$$\sum_{i=1}^{m}(\boldsymbol{\lambda}_i - \overline{\boldsymbol{\lambda}}_i)\boldsymbol{\varphi}_i(\boldsymbol{x}) > \mathbf{0}.$$

Противоречие. Следовательно, $\overline{\boldsymbol{x}} \in \boldsymbol{Q}$.

Лагранжева теория двойственности

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

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

откуда:

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

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

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

По Лемме 2, $\overline{\boldsymbol{x}}$ и $\overline{\boldsymbol{\lambda}}$ – оптимальные решения прямой и двойственной задачи, 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. $\min\limits_{x} \sup\limits_{\lambda \geq 0} L(x, \lambda) = \max\limits_{\lambda \geq 0} \inf\limits_{x} L(x, \lambda)$.

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

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