Методы оптимизации. Семинар 10. Условия оптимальности.

Александр Катруца

Московский физико-технический институт,
Факультет Управления и Прикладной Математики

31 октября 2016 г.

Напоминание

- Конус возможных направлений
- Касательный конус
- Острый экстремум

Мотивация

Вопрос 0

Когда существует решение оптимизационной задачи?

Вопрос 1

Как проверить, что точка является решением оптимизационной задачи?

Вопрос 2

Из каких условий можно найти решение оптимизационной задачи?

Существование решения

Теорема Вейерштрасса

Пусть $X \subset R^n$ компактное множество и пусть $f(x)$ непрерывная функция на X. Тогда точка глобального минимума функции $f(x)$ на X существует.

Эта теорема гарантирует, что решение подавляющего большинства разумных задач существует.

Условия оптимальности

Определение

Условием оптимальности будем называть некоторое выражение, выполнимость которого даёт необходимое и (или) достаточное условие экстремума.

Классы задач:

- Общая задача минимизации
- Задача безусловной минимизации
- Задача минимизации с ограничениями типа равенств
- Задача минимизации с ограничениями типа равенств и неравенств

Общая задача минимизации

Задача

$$f(x) \to \min_{x \in X}$$

Критерий оптимальности

Пусть $f(x)$ определена на множестве $X \subset \mathbb{R}^n$. Тогда

1. если x^* точка минимума $f(x)$ на X, то $\partial_X f(x^*) \neq \emptyset$ и $0 \in \partial_X f(x^*)$
2. если для некоторой точки $x^* \in X$ существует субдифференциал $\partial_X f(x^*)$ и $0 \in \partial_X f(x^*)$, то x^* — точка минимума $f(x)$ на X.

Какие недостатки у приведённого критерия?

Примеры

- $\mathbf{x}^\top\mathbf{x} + \alpha\|\mathbf{x} - \mathbf{c}\|_2 \to \min\limits_{\mathbf{x}\in\mathbb{R}^n}, \ \alpha > 0$
- $\mathbf{x}^\top\mathbf{x} + \alpha\|\mathbf{c}^\top\mathbf{x} - b\|_2 \to \min\limits_{\mathbf{x}\in\mathbb{R}^n}, \ \alpha > 0$
- Ограничение на допустимое множество

$$(x+2)^2 + |y+3| \to \min_{(x,y)\in\mathbb{R}^2}$$

$$\text{s.t. } 8 + 2x - y \leq 0$$

Задача безусловной минимизации

Задача: $f(x) \to \min\limits_{x \in \mathbb{R}^n}$.

Критерий оптимальности для выпуклых функций

Пусть $f(x)$ выпуклая функция на $\mathbb{R}^n$. Тогда точка x^* решение задачи безусловной минимизации $\Leftrightarrow 0 \in \partial f(x^*)$.

Следствие

Если $f(x)$ выпукла и дифференцируема на $\mathbb{R}^n$. Тогда точка x^* решение задачи безусловной минимизации $\Leftrightarrow$ $\nabla f(x^*) = 0$.

Достаточное условие для невыпуклых функций

Пусть f дважды дифференцируема на $\mathbb{R}^n$ и x^* такая что $\nabla f(x^*) = 0$. Тогда если $\nabla^2 f(x^*) \succ 0$, то x^* точка строгого локального минимума $f(x)$ на $\mathbb{R}^n$.

Примеры

- $x_1 e^{x_1} - (1 + e^{x_2}) \cos x_2 \to \min$
- Функция Розенброка:
  $$(1 - x_1)^2 + \alpha \sum_{i=2}^{n} (x_i - x_{i-1})^2 \to \min, \ \alpha > 0$$
- $x_1^2 + x_2^2 - x_1 x_2 + e^{x_1 + x_2} \to \min$

Задача

$$f(x) \to \min_{x \in \mathbb{R}^n}$$
$$\text{s.t. } g_i(x) = 0, \ i = 1, \ldots, m$$

Лагранжиан

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

Критерий оптимальности

Пусть $f(x)$ и $g_i(x)$ дважды дифференцируемы в точке x^* и непрерывно дифференцируемы в некоторой окрестности x^*. Пусть также $\nabla_x L(x^*, \boldsymbol{\lambda}) = 0$. Тогда если $\mathbf{h}^\top \nabla^2 L(x^*, \boldsymbol{\lambda})\mathbf{h} > 0$, где $\mathbf{h} \in T(\mathbf{x}^* | G)$ — касательный конус, то x^* — точка локального минимума.

Возможные варианты

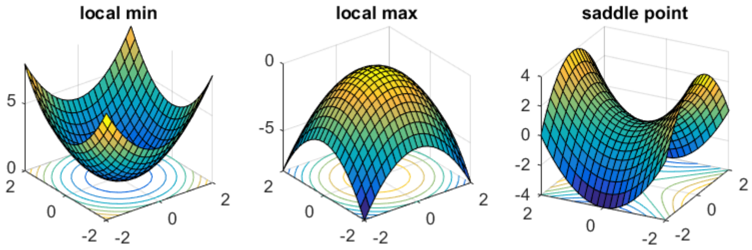

Рисунок взят из блога
`http://www.offconvex.org/2016/03/22/saddlepoints/`

Примеры

- $\sum\limits_{i=1}^{n} \alpha_i x_i^4 \to \mathrm{extr}_{\mathbf{x} \in G},\ G = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{c}^\top \mathbf{x} = 1\},$
  $\alpha_i > 0,\ c_i > 0$
- $x_1 + 4x_2 + 9x_3 \to \mathrm{extr}_{\mathbf{x} \in G},\ G = \left\{ \frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} = 1 \right\}$
- Примеры из задачника по матану на метод множителей Лагранжа

Задача

$$\min_{x \in \mathbb{R}^n} f(x)$$

$$\text{s.t. } g_i(x) = 0, \; i = 1, \ldots, m$$

$$h_j(x) \leq 0, \; j = 1, \ldots, p$$

Лагранжиан

$$L(x, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{j=1}^{p} \mu_j h_j(x)$$

Условия оптимальности

Необходимое условие

Пусть x^* решение задачи математического программирования, и функции f, h_j, g_i дифференцирумы. Тогда найдутся такие $\boldsymbol{\mu}^*$ и $\boldsymbol{\lambda}^*$, что выполнены следующие условия:

- $g_i(x^*) = 0$
- $h_j(x^*) \leq 0$
- $\mu_j^* \geq 0$
- $\mu_j^* h_j(x^*) = 0$
- $\nabla_x L(x^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) = 0$

Если задача выпуклая, то это же условие является достаточным.

Условия оптимальности (cont'd)

Если задача невыпуклая, то

Достаточное условие первого порядка

Если для стационарной точки $(x^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ число активных неравенств $|J|$ такое что $n = m + |J|$ и $\mu_j > 0$, $j \in J$, то эта точка является точкой минимума.

Достаточное условие второго порядка

Если в задаче математического программирования число активных ограничений меньше размерности задачи, то точка x^* яляется решением задачи, если выполнены условия

$$\mathbf{z}^\mathsf{T} \nabla^2_{xx} L(x^*) \mathbf{z} > 0$$

для

- $\mathbf{z} \neq 0$ и $\nabla g_i^\mathsf{T}(x^*)\mathbf{z} = 0$
- при $j \in J$ и $\mu_j > 0$, $\nabla h_j^\mathsf{T}(x^*)\mathbf{z} = 0$
- при $j \in J$ и $\mu_j = 0$, $\nabla h_j^\mathsf{T}(x^*)\mathbf{z} \leq 0$

Примеры

- Пример 1

$$\begin{aligned} &\operatorname{extr}(x_1 - 3)(x_2 - 2)\\ \text{s.t. } &x_1 + 2x_2 = 4\\ &x_1^2 + x_2^2 \le 5\\ &x_1 \ge 0,\ x_2 \ge 0 \end{aligned}$$

- Пример 2

$$\operatorname{extr} \sum_{i=1}^{n} \frac{c_i}{x_i}$$

$$\text{s.t. } \sum_{i=1}^{n} a_i x_i \le b$$

$$x_i > 0,\ b > 0,\ c_i > 0,\ a_i > 0$$

- Пример 3

$$\begin{aligned} &\operatorname{extr}(x_1 x_3 - 2x_2)\\ \text{s.t. } &2x_1 - x_2 - 3x_3 \le 10\\ &3x_1 + 2x_2 + x_3 = 6\\ &x_2 \ge 0 \end{aligned}$$

Резюме

- Существование решения оптимизационной задачи
- Условия оптимальности для
  - общей задачи оптимизации
  - задачи безусловной оптимизации
  - задачи оптимизации с ограничениями типа равенств
  - задачи оптимизации с ограничениями типа равенств и неравенств