Методы оптимизации
Семинар 4. Условия оптимальности. Теорема Каруша-Куна-Таккера

Лобанов Александр Владимирович

Московский физико-технический институт
Факультет инноваций и высоких технологий

lobanov.av@mipt.ru

22 сентября 2022 г.

Напоминание

Теорема Вейерштрасса

Пусть $S \subset \mathbb{R}^n$ – компактное множество, а $f(x)$ – непрерывная функция на S. Тогда точка глобального минимума функции $f(x)$ на S существует.

Напоминание

Теорема Вейерштрасса

Пусть $S \subset \mathbb{R}^n$ – компактное множество, а $f(x)$ – непрерывная функция на S. Тогда точка глобального минимума функции $f(x)$ на S существует.

Множетели Лагранжа

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

$$\text{s.t. } h_i(x) = 0, \quad i = 1, ..., p$$

Основная идея метода множителей Лагранжа заключается в переходе от условной оптимизации к безусловной за счет увеличения размерности задачи:

$$L(x, \mu) = f(x) + \sum_{i=1}^{p} \mu_i h_i(x) \to \min_{x \in \mathbb{R}^n, \mu \in \mathbb{R}^p}$$

Безусловная оптимизация

Пусть $f(x) : \mathbb{R}^n \to \mathbb{R}$ – дважды дифференцируемая функция.

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

Безусловная оптимизация

Пусть $f(x) : \mathbb{R}^n \to \mathbb{R}$ – дважды дифференцируемая функция.

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

Необходимое условие

Если x^* – локальный минимум $f(x)$, тогда

$$\nabla f(x^*) = 0$$

Безусловная оптимизация

Пусть $f(x) : \mathbb{R}^n \to \mathbb{R}$ – дважды дифференцируемая функция.

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

Необходимое условие

Если x^* – локальный минимум $f(x)$, тогда

$$\nabla f(x^*) = 0$$

Достаточное условие

Если $f(x)$ в некоторой точке x^* удовлетворяет следующим условиям:

$$H_f(x^*) = \nabla^2 f(x^*) \succ (\prec) 0,$$

тогда (при выполнении необходимого условия) x^* является локальным минимум (максимумом) функции $f(x)$.

Пусть $f(x) : \mathbb{R}^n \to \mathbb{R}$ – дважды дифференцируемая функция.

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

Необходимое условие

Если x^* – локальный минимум $f(x)$, тогда

$$\nabla f(x^*) = 0$$

Достаточное условие

Если $f(x)$ в некоторой точке x^* удовлетворяет следующим условиям:

$$H_f(x^*) = \nabla^2 f(x^*) \succ (\prec) 0,$$

тогда (при выполнении необходимого условия) x^* является локальным минимум (максимумом) функции $f(x)$.

(Peano surface $f(x, y) = (2x^2 - y)(y - x^2)$).

Оптимизации с ограничениями типа равенств (интуитивно)

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

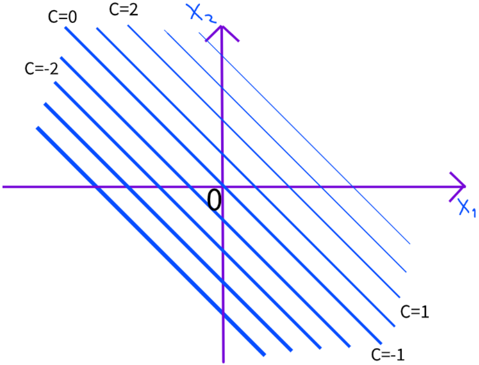
$$\text{s.t. } h(x) = 0$$

Попробуем проиллюстрировать подход к решению данной задачи а простом примере с $f(x) = x_1 + x_2$ и $h(x) = x_1^2 + x_2^2 - 2$.

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

$$\text{s.t. } h(x) = 0$$

Попробуем проиллюстрировать подход к решению данной задачи а простом примере с $f(x) = x_1 + x_2$ и $h(x) = x_1^2 + x_2^2 - 2$.

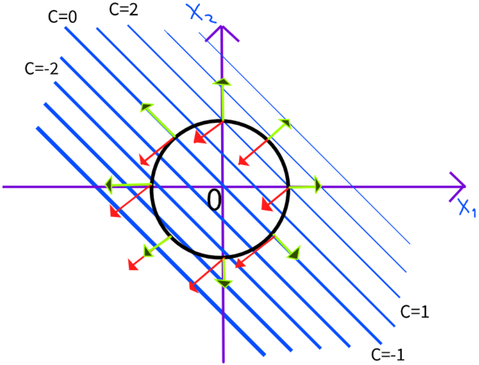


Оптимизации с ограничениями типа равенств (интуитивно)

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

$$\text{s.t. } h(x) = 0$$

Попробуем проиллюстрировать подход к решению данной задачи а простом примере с $f(x) = x_1 + x_2$ и $h(x) = x_1^2 + x_2^2 - 2$.



Оптимизации с ограничениями типа равенств (интуитивно)

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

$$
\text{s.t. } h(x) = 0
$$

Попробуем проиллюстрировать подход к решению данной задачи а простом примере с $f(x) = x_1 + x_2$ и $h(x) = x_1^2 + x_2^2 - 2$.

Чтобы двигаться от точки X_R по допустимому множеству в сторону убывания функции, нужно гарантировать два условия:

$$
\langle \delta x, \nabla h(x_R) \rangle = 0
$$

$$
\langle \delta x, -\nabla f(x_R) \rangle > 0
$$

Предположим, что в процессе движения мы подошли к точке, где $\boxed{-\nabla f(x) = \mu \nabla h(x)}$. Тогда мы подошли к точке допустимого множества, отодвинувшись от которой не получится сократить нашу функцию. Это локальный минимум в условной задаче.

Оптимизации с ограничениями типа равенств

Теперь определим функцию Лагранжа

$$L(x, \mu) = f(x) + \mu h(x).$$

Тогда точка x^* является локальным минимумом задачи, описанной выше, тогда и только тогда, когда:

Необходимые условия

$$\nabla_x L(x^*, \mu^*) = 0$$
$$\nabla_\mu L(x^*, \mu^*) = 0$$

Достаточные условия

$$\langle y, \nabla^2_{xx} L(x^*, \mu^*) y \rangle > 0,$$
$$\forall y \neq 0 \in \mathbb{R}^n : \nabla h(x^*)^T y = 0$$

Оптимизации с ограничениями типа равенств (в общем случае)

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

$$\text{s.t. } h_i(x) = 0, \ i = 1, ..., p$$

Решение

$$L(x, \mu) = f(x) + \sum_{i=1}^{p} \mu_i h_i(x) = f(x) + \mu^T h(x).$$

Пусть $f(x)$ и $h_i(x)$ дважды дифференцируемы в точке x^* и непрерывно дифференцируемы в некоторой окрестности x^*. Тогда условия локального минимума для $x \in \mathbb{R}^n$, $\mu \in \mathbb{R}^p$ записываются как

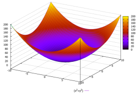
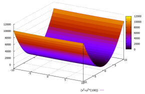
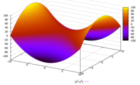
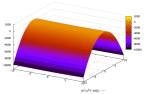
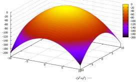
$$\nabla_x L(x^*, \mu^*) = 0$$

$$\nabla_\mu L(x^*, \mu^*) = 0$$

$$\langle y, \nabla^2_{xx} L(x^*, \mu^*) y \rangle > 0,$$

$$\forall y \neq 0 \in \mathbb{R}^n : \nabla h_i(x^*)^T y = 0$$

Зависимость критической точки от поведения Гессиана

$y^t H_f y$	Определенность H	x^*	
> 0	Положительно опр.	Минимум	
≥ 0	Положительно полуопр.	Складка	
$\neq 0$	Не определен	Седловая точка	
≤ 0	Отрицательно полуопр.	Складка	
< 0	Отрицательно опр.	Максимум	

Оптимизация с ограничениями типа неравенств

Задача условной оптимизации

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

$$\text{s.t. } g(x) \leq 0$$

Задача условной оптимизации

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

$$\text{s.t. } g(x) \leq 0$$

Случай 1

Рассмотрим пример, где $f(x) = x_1^2 + x_2^2$, $g(x) = x_1^2 + x_2^2 - 1$

Оптимизация с ограничениями типа неравенств

Задача условной оптимизации

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

$$\text{s.t. } g(x) \leq 0$$

Случай 1

Рассмотрим пример, где $f(x) = x_1^2 + x_2^2$, $g(x) = x_1^2 + x_2^2 - 1$

Случай 2

Рассмотрим пример, где $f(x) = (x_1 - 1)^2 + (x_2 + 1)^2$, $g(x) = x_1^2 + x_2^2 - 1$

Оптимизация с ограничениями типа неравенств

Задача условной оптимизации

Таким образом, при решении следующей задачи:

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

$$\text{s.t. } g(x) \leq 0$$

возможны два случая:

$g(x) \leq 0$ неактивно. $g(x^*) < 0$

$$g(x^*) < 0$$
$$\nabla f(x^*) = 0$$
$$\nabla^2 f(x^*) > 0$$

$g(x) \leq 0$ активно. $g(x^*) = 0$

$$g(x^*) = 0$$
$$-\nabla f(x^*) = \lambda \nabla g(x^*), \lambda > 0$$
$$\langle y, \nabla^2_{xx} L(x^*, \lambda^*) y \rangle > 0,$$
$$\forall y \neq 0 \in \mathbb{R}^n : \nabla g(x^*)^T y = 0$$

Оптимизация с ограничениями типа неравенств

Объединяя два возможных случая, мы можем записать общие условия задачи:

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

$$\text{s.t. } g(x) \leq 0$$

Определим функцию Лагранжа:

$$L(x, \lambda) = f(x) + \lambda g(x)$$

тогда точка x^* является локальным минимумом тогда и только тогда, когда:

- $\nabla_x L(x^*, \lambda^*) = 0$
- $\lambda^* \geq 0$
- $\lambda^* g(x^*) = 0$
- $g(x^*) \leq 0$
- $\langle y, \nabla^2_{xx} L(x^*, \lambda^*) y \rangle > 0, \forall y \neq 0 \in \mathbb{R}^n : \nabla g(x^*)^T y \leq 0$

Общая формулировка задачи математического программирования

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

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

$$h_i(x) = 0, \ i = 1, ..., p$$

Функция Лагранжа

Решение включает в себя построение функции Лагранжа:

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

Условия Каруша-Куна-Такера

Необходимые условия

Пусть x^* – решение задачи математического программирования, и функции f, g_i, h_i дифференцируемы. Тогда найдутся такие λ^*, μ^*, что выполнены следующие условия:

$$\nabla_x L(x^*, \lambda^*, \mu^*) = 0$$

$$\nabla_\mu L(x^*, \lambda^*, \mu^*) = 0$$

$$\lambda_i^* \geq 0, \;\; i = 1, ..., m$$

$$\lambda_i^* g_i(x^*) = 0, \;\; i = 1, ..., m$$

$$g_i(x^*) \leq 0, \;\; i = 1, ..., m$$