Методы оптимизации
Семинар 5. Введение в двойственность

Лобанов Александр Владимирович

Московский физико-технический институт
Факультет инноваций и высоких технологий

lobanov.av@mipt.ru

29 сентября 2022 г.

Таблица успеваемости группы: Б05-027

№	Студент	Посещаемость															Д/З				К/Р		Текущая оценка за курс МО
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	1	2	3	4	1	2	
1	Аджима Никита	+	+	+	н												⦶						10
2	Алексеев Максим	+	+	+	+												-2						8
3	Белый Антон	н	+	+	+												⦶						10
4	Вандакуров Артем	+	+	+	+												⦶						10
5	Грачев Кирилл	+	+	+	н												⦶						10
6	Егоров Гордей	+	+	+	+												⦶	⦶					10
7	Мустафин Артём	н	+	+	+												⦶						10
8	Мухаметгалин Артур	н	н	+	+												-2						8
9	Русскин Николай	н	+	+	н												⦶						10
10	Рябухин Никита	н	н	+	+												-2						8
11	Турлыбеков Олжас	+	н	+	+												-2						8
12	Хоружий Тимофей	+	н	н	н												⦶						10
13	Хрол Ариана	+	+	н	+												⦶						10
14	Челышкин Артём	+	+	н	+												⦶						10
15	Яфаров Владимир	+	н	+	н												-2						8

Обозначения: «+» – присутствовал на семинаре, «н» — отсутствовал на семинаре, «⦶» — сдано Д/З

Дата обновления: 27 сентября 2022 г.

Задача выпуклой оптимизации

Определение

Задачи математического программирования следующего типа называются задачами выпуклой оптимизации:

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

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

Задача выпуклой оптимизации

Определение

Задачи математического программирования следующего типа называются задачами выпуклой оптимизации:

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

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

$$Ax = b,$$

где все функции $f(x), g_1(x), ..., g_m(x)$ – выпуклые функции и ограничение типа равенств – аффинно.

Что будем делать сегодня?

Что будем делать сегодня?

Прямая задача

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

Что будем делать сегодня?

Прямая задача

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

Двойственная задача

$$g(y) \to \max_{y \in \Omega}$$

Что будем делать сегодня?

Прямая задача

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

Двойственная задача

$$g(y) \to \max_{y \in \Omega}$$

Мы построим $g(y)$, сохраняющую равномерную оценку:

$$g(y) \leq f(x) \quad \forall x \in S, \; \forall y \in \Omega$$

А как следствие:

$$\max_{y \in \Omega} g(y) \leq \min_{x \in S} f(x)$$

Рассмотрим один возможный способ построения $g(y)$ в случае, когда мы имеем общую задачу математичекого программирования с функциональными ограничениями.

Основные обозначения

Задача математического программирования

Основные обозначения

Задача математического программирования

$$f_0(x) \to \min_{x \in \mathbb{R}^n}$$

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

Основные обозначения

Задача математического программирования и Лагранжиан

$$f_0(x) \to \min_{x \in \mathbb{R}^n}$$

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

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

Основные обозначения

Задача математического программирования и Лагранжиан

$$f_0(x) \to \min_{x \in \mathbb{R}^n}$$

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

$$L(x, \lambda, \mu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \mu_i h_i(x) = f_0(x) + \lambda^T f(x) + \mu^T h(x)$$

Основные обозначения

Задача математического программирования и Лагранжиан

$$f_0(x) \to \min_{x \in \mathbb{R}^n}$$

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

$$L(x, \lambda, \mu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \mu_i h_i(x) = f_0(x) + \lambda^T f(x) + \mu^T h(x)$$

$$\mathbb{D} = \text{dom } f_0(x) \cap \bigcap_{i=1}^{m} \text{dom } f_i(x) \cap \bigcap_{j=1}^{p} \text{dom } h_j(x)$$

Основные обозначения

Задача математического программирования и Лагранжиан

$$f_0(x) \to \min_{x \in \mathbb{R}^n}$$

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

$$L(x, \lambda, \mu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \mu_i h_i(x) = f_0(x) + \lambda^T f(x) + \mu^T h(x)$$

Двойственная функция

Определим двойственную функцию Лагранжа $g : \mathbb{R}^m_+ \times \mathbb{R}^p \to \mathbb{R}$ как минимальное значение Лагранжиана по x: для всех $\lambda \in \mathbb{R}^m_+, \mu \in \mathbb{R}^p$

$$g(\lambda, \mu) = \inf_{x \in \mathbb{D}} L(x, \lambda, \mu) = \inf_{x \in \mathbb{D}} \left(f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \mu_i h_i(x) \right)$$

Нижняя граница (оценка снизу)

Нижняя оценка

Двойственная функция дает нижние оценки оптимального значения исходной задачи p^*. Для всех $\lambda \geq 0, \mu$:

$$g(\lambda, \mu) \leq p^*$$

Нижняя граница (оценка снизу)

Нижняя оценка

Двойственная функция дает нижние оценки оптимального значения исходной задачи p^*. Для всех $\lambda \geq 0, \mu$:

$$g(\lambda, \mu) \leq p^*$$

Попробуем доказать

Предположим, что некоторые $\hat{x}$ – точки из допустимого множества для исходной задачи, т.е. $f_i(\hat{x}) \leq 0$ и $h_j(\hat{x}) = 0,\ \lambda \geq 0$. Тогда мы имеем:

$$L(\hat{x}, \lambda, \mu) = f_0(\hat{x}) + \lambda^T f(\hat{x}) + \mu^T h(\hat{x}) \leq f_0(\hat{x})$$

Следовательно

$$g(\lambda, \mu) = \inf_{x \in \mathbb{D}} L(x, \lambda, \mu) \leq L(\hat{x}, \lambda, \mu) \leq f_0(\hat{x})$$

Нижняя граница (оценка снизу)

Нижняя оценка

Двойственная функция дает нижние оценки оптимального значения исходной задачи p^*. Для всех $\lambda \geq 0, \mu$:

$$g(\lambda, \mu) \leq p^*$$

Двойственная задача оптимизации

Какова наилучшая нижняя оценка, которую можно получить из двойственной функции Лагранжа? Именно этот порос приводи к следующей задачи оптимизации:

$$g(\lambda, \mu) \to \max_{\lambda \in \mathbb{R}^m, \mu \in \mathbb{R}^p}$$

$$\text{s.t. } \lambda \geq 0$$

Оптимальными множителями Лагранжа называют (λ^*, μ^*), если они оптимальны для поставленной задачи.

Функция	$\boxed{\text{Прямая задача}}$	$\boxed{\text{Двойственная задача}}$
	$f_0(x)$	$g(\lambda, \mu) = \inf\limits_{x \in \mathbf{dom} f_0} L(x, \lambda, \mu)$
Переменные	$x \in \mathbb{R}^n$	$\lambda \in \mathbb{R}^m, \mu \in \mathbb{R}^p$
Ограничения	s.t. $f_i(x) \leq 0, \; i = 1, ..., m$ $h_i(x) = 0, \; i = 1, ..., p$	$\lambda_i \geq 0, \forall i \in \overline{1, m}$
Задача	$f_0(x) \to \min\limits_{x \in \mathbb{R}^n}$ s.t. $f_i(x) \leq 0, \; i = 1, ..., m$ $h_i(x) = 0, \; i = 1, ..., p$	$g(\lambda,\mu) \to \max\limits_{\lambda \in \mathbb{R}^m, \mu \in \mathbb{R}^p}$ s.t. $\lambda \geq 0$
Решение	$x^*, \quad p^* = f_0(x^*)$	$\lambda^*, \mu^*, \quad d^* = g(\lambda^*, \mu^*)$

Сильная и слабая двойственность

Двойственный зазор

Оптимальные значения целевой функции в прямой и двойственной задаче связаны соотношением:

$$d^* \leq p^*$$

При этом разница между ними называется двойственным зазором.

Двойственный зазор

Оптимальные значения целевой функции в прямой и двойственной задаче связаны соотношением:

$$d^* \leq p^*$$

При этом разница между ними называется двойственным зазором.

Слабая двойственность

Если $d^* < p^*$, то свойство называют слабой двойственностью.
Если $d^* = p^*$, то – сильной двойственностью.

Сильная и слабая двойственность

Двойственный зазор

Оптимальные значения целевой функции в прямой и двойственной задаче связаны соотношением:

$$d^* \leq p^*$$

При этом разница между ними называется двойственным зазором.

Слабая двойственность

Если $d^* < p^*$, то свойство называют слабой двойственностью.
Если $d^* = p^*$, то – сильной двойственностью.

Условия Слейтора

Если задача выпуклая и существует x, лежащий внутри допустимой области, т.е. ограничения типа неравенств выполнены как строгие неравенства, то выполнено свойство сильной двойственности

Полезные свойства двойственности

Построение нижней границы решения прямой задачи

Бывает, что решить прямую задачу может быть сложно. С помощью двойственной задачи можно получить некоторую нижнюю оценку, подставив любое $y \in \Omega$ в $g(y)$.

Проверка разрешимости задачи и достижимости решения

Из неравенства $\max_{y \in \Omega} g(y) \le \min_{x \in S} f_0(x)$ следует, что если $\min_{x \in S} f_0(x) = -\infty$, то $\Omega = \varnothing$ и наоборот.

Упрощение процесса решения задачи

Иногда легче решить двойную задачу, чем основную. Например, в случае, если выполняется свойство строгой двойственности: $g(y^*) = f(x^*)$.

Критерий субоптимальности

$f_0(x) - f^* \le f_0(x) - g(y) = \varepsilon$.

Способы использования: критерий остановки в итерационном процессе, теоретическая оценка сходимости алгоритма, проверка оптимальности данной точки.