

МЕТОДЫ ОПТИМИЗАЦИИ

Лекция 2: Введение в теорию сложности. Лагранжева теория двойственности

Панин Артем Александрович
Email: aaпанin1988@gmail.com

Лекции и материалы по курсу:

http://www.math.nsc.ru/LBRT/k5/Opt_MMf_2021.html

Новосибирский государственный университет

2021

Оптимизационная задача:

Оптимизационная задача P – четверка $(I_P, \mathfrak{F}_P, f_P, goal_P)$, где:

1. I_P – множество входов задачи.
2. \mathfrak{F}_P – функция, которая каждому входу $x \in I_P$ сопоставляет множество допустимых решений $\mathfrak{F}_P(x)$.
3. f_P – целевая функция задачи, а $f(x, y)$ – значение целевой функции для входа x на допустимом решении $y \in \mathfrak{F}_P(x)$.
4. Параметр $goal \in \{\min, \max\}$ определяет тип задачи: задача на минимум или задача на максимум.

(Когда понятно о какой задаче идет речь опускаем индекс P)

$\mathfrak{F}^*(x)$ – множество оптимальных решений задачи для входа x ,

$y^*(x)$ – оптимальное решение задачи,

$f^*(x)$ – оптимальное значение целевой функции задачи.

Сложность алгоритма

Сложность алгоритма – количество ресурса (обычно время), которое требуется для решения задачи. Для машины Тьюринга время измеряется количеством операций, выполняемых ею до момента перехода в финальное состояние. Сложность $T(|x|)$ представляется в виде $O(g(|x|))$, где $|x|$ – количество бит в представлении текущего экземпляра входа x (то есть существует константа c такая, что $|T(|x|)| \leq c |g(|x|)|$).

Эффективные алгоритмы

Если $T(|x|) = O(|x|^k)$, то алгоритм имеет полиномиальную сложность. Алгоритмы с полиномиальной сложностью используются в качестве формального аналога содержательного понятия эффективные алгоритмы. Все остальные алгоритмы имеют экспоненциальную сложность

Классы сложности

Класс сложности это множество задач, для каждой из которых существует алгоритм решения, временная сложность которого ограничена сверху некоторой функцией из заданного класса функций.

Первый уровень полиномиальной иерархии

Множество задач распознавания, решаемых за полиномиальное время детерминированными алгоритмами образуют класс P .

Множество задач распознавания, решаемых за полиномиальное время недетерминированными алгоритмами образуют класс NP .

Класс NP состоит из задач, для которых решение ("да" вход) может быть проверено (детерминированным алгоритмом) за полиномиальное время при наличии сертификата.

Класс NP

Пусть граф G – "да" вход задачи о гамильтоновом цикле. Тогда сам цикл выступает в качестве сертификата. Содержательно за недетерминированное время отгадываем цикл, а затем за детерминированное полиномиальное время проверяем, что это действительно гамильтонов цикл в графе G .

Итак, мы вводим класс NP и недетерминированные машины для того, чтобы получить понятийный аппарат, который позволяет нам сделать объектом исследования трудноразрешимые задачи.

Недетерминированные алгоритмы виртуальны по своей природе. У них нет реальных аналогов.

Второй уровень полиномиальной иерархии

Второй уровень полиномиальной иерархии определяется с помощью детерминированных и недетерминированных оракульных машин Тьюринга.

Задача распознавания L принадлежит классу Δ_2^P , если существует детерминированная оракульная машина Тьюринга, которая распознает за полиномиальное время задачу L , используя в качестве оракула некоторый язык из класса NP (NP -оракул). Класс Δ_2^P часто обозначают как P^{NP} . Аналогично определяются классы Σ_2^P и Π_2^P (соответственно, NP^{NP} и $co-NP^{NP}$).

Итерируя этот процесс получим полиномиальную иерархию.

Оптимизационные задачи: полиномиально разрешимые \rightarrow
 NPO -полные \rightarrow **неразрешимые**

ЛП \rightarrow ЦЛП \rightarrow ЦЛП \pm квадраты переменных

Рассмотрим задачу P с произвольными функциями f и φ_i :

$$f(x) \longrightarrow \min \quad (1)$$

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

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).

Пусть

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

Тогда

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

Задача P эквивалентна следующей:

$$g(x) \longrightarrow \min$$

Пусть

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

Рассмотрим задачу D :

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

(D) – задача двойственная к прямой (или исходной) задаче P .

$\lambda_1, \dots, \lambda_m$ – двойственные переменные,

x_1, \dots, x_n – прямые переменные.

Если $x \in Q$, $\lambda \geq 0$, то x – допустимое решение прямой задачи, а λ – допустимое решение двойственной задачи.

Лемма 1. (Слабая теорема двойственности).

$$\forall x \in Q \quad \forall \lambda \geq 0 \quad (h(\lambda) \leq f(x)).$$

Док-во: $h(\lambda) = \inf_{\tilde{x} \in R^n} L(\tilde{x}, \lambda) \leq L(x, \lambda) \leq \sup_{\tilde{\lambda} \geq 0} L(x, \tilde{\lambda}) = f(x). \blacksquare$

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

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

$$L(x^*, \lambda) \stackrel{(4)}{\leq} L(x^*, \lambda^*) \stackrel{(5)}{\leq} L(x, \lambda^*) \quad \forall x \in R^n, \forall \lambda \geq 0.$$

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

Док-во:

1) НЕОБХОДИМОСТЬ.

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

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

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

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

Из последнего следуют (4) и (5).

2) ДОСТАТОЧНОСТЬ.

Пусть $(\bar{x}, \bar{\lambda})$ – седловая точка функции Лагранжа. Тогда из (4) следует:

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

Предположим, что $\bar{x} \notin Q$. Т.е. $\exists i : \varphi_i(\bar{x}) > 0$. Тогда для достаточно большого $\lambda_i > 0$:

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

Противоречие. Следовательно, $\bar{x} \in Q$.

При $\lambda = 0$ имеем:

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

откуда:

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

Следовательно, $f(\bar{x}) = L(\bar{x}, \bar{\lambda})$. Из (5) следует, что:

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

По Лемме 2, \bar{x} и $\bar{\lambda}$ – оптимальные решения прямой и двойственной задачи, соответственно. ■

Следствие 1. Пусть $\bar{x} \in Q(P)$, $\bar{\lambda} \geq 0$. Следующие утверждения эквивалентны:

1. Пара $(\bar{x}, \bar{\lambda})$ — седловая точка функции Лагранжа.

2. $f(\bar{x}) = h(\bar{\lambda})$.

3. $\min_x \sup_{\lambda \geq 0} L(x, \lambda) = \max_{\lambda \geq 0} \inf_x L(x, \lambda)$.

Следствие 2. Пусть $x^*, \bar{x} \in Q$, $\lambda^*, \bar{\lambda} \geq 0$.

Если пары $(\bar{x}, \bar{\lambda})$ и (x^*, λ^*) — седловые точки функции Лагранжа, то пары (\bar{x}, λ^*) и $(x^*, \bar{\lambda})$ — также седловые точки функции Лагранжа, причем

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