Вводная лекция

Мальковский Н. В.

Санкт-Петербургский Академический Университет

Что такое математическая оптимизация?

Общая задача минимизации: найти для функции $f : \mathcal{D} \to \mathbb{R}$ такую точку x^*, что

$$f(x) \geq f(x^*) \quad \forall x \in \mathcal{D}$$

или по заданному ϵ такую точку x_ϵ, что

$$f(x_\epsilon) \leq f(x^*) + \epsilon$$

или

$$||x_\epsilon - x^*|| \leq \epsilon.$$

Что изучает математическая оптимизация?

- Общие закономерности задач математической оптимизации
- Аналитические методы решения задач оптимизации
- Численные методы решения задач оптимизации и их свойства (алгоритмическая сложность)

О чем курс?

Основное направление курса – выпуклая оптимизация, основные темы:

- Методы безусловной оптимизации (градиентный/субградиентный спуск, случайный поиск, метод Ньютона) и их анализ
- Теория двойственности
- Методы условной оптимизации (симплекс-метод, проективный градиентный спуск, метод эллипсоидов, метод внутренней точки) и их анализ
- Некоторые стандартные обобщения: стохастические и покоординатные варианты методов

Основные параметры задачи

- Свойства множества $\mathcal{D}$ (дискретность, связность, выпуклость)
- Свойства функции f (непрерывность, дифференцируемость, выпуклость)
- Доступные к измерению величины
  - Значения f и её производных
  - Вспомогательные функции: проекция на $\mathcal{D}$, опорная гиперплоскость к $\mathcal{D}$ и т. п.

Виды алгоритмов оптимизации

- Точные: алгоритм за конечное число арифметических действий получает значение x^*
- Приближенные: для некоторого $\gamma \geq 0$ и любого $\epsilon > \gamma$ алгоритм за конечное число операций получает x_ϵ такое, что

$$||x_\epsilon - x^*|| \leq \epsilon \quad \text{или} \quad f(x_\epsilon) - f(x^*) \leq \epsilon$$

- Асимптотические: алгоритм строит последовательность $x_k \in \mathcal{D}$ такую, что

$$x_k \to x^* \quad \text{или} \quad f(x_k) \to f(x^*)$$

Комбинаторная оптимизация

- Простая функция f
- Дискретное множество $\mathcal{D}$
- Большинство задач не разрешимы полиномиально
- ~~Используются в основном точные алгоритмы~~
- Основные темы: оптимизация на графах, матроиды, субмодулярные функции и т. д.

Выпуклая оптимизация

- Выпуклая функция f
- Выпуклое множество $\mathcal{D}$ (важно: выпуклость исключает дискретность)
- Используются в основном асимптотические методы, иногда точные (симплекс-метод)
- Все методы имеют сходимость по ϵ не хуже $\mathcal{O}(1/\sqrt{\epsilon})$, количество вычислений на одну итерацию полиномиально относительно размера пространства
- Основные темы: линейное программирование, двойственность, градиентный спуск и прочее

Замечания

Следующие методы также распространены для различных задач оптимизации:

- Метод отжига/генетические методы: в основном используются для задач, где неприменимы другие методы. Не имеют гарантий результата.
- Метод ветвей и границ: также используется, когда больше нечего применить, однако имеет гарантию сходимости, хоть и может при этом очень долго искать ответ.
- Динамическое программирование: изначально задумано для оптимизации процессов, однако в последствии нашло применение в гораздо большем классе задач.

Простая задача оптимизации

Пусть $\mathcal{D}$ – открытое множество, $f : \mathcal{D} \to \mathbb{R}$ – дифференцируемая функция. Задача

$$\text{минимизировать} \quad f(x), \quad x \in \mathcal{D}$$

исторически является одной из первых задач оптимизация общего вида. Решение этой задачи обычно получается из условий стационарности (теорема Ферма)

$$\nabla f(x) = 0_{\mathcal{D}}.$$

Задача оптимизации с ограничениями

Пусть $\mathcal{D}$ – открытое множество, $f : \mathcal{D} \to \mathbb{R}$, $g : \mathcal{D} \to \mathbb{R}^m$ – дифференцируемые функции. Задача

$$\begin{array}{ll} \text{минимизировать} & f(x), \quad x \in \mathcal{D} \\ \text{при условии} & g(x) = 0_m \end{array}$$

впервые была изучена Лагранжем, позднее в 20 веке обобщена до задач с ограничением в виде неравенств. На основе этих результатов выстроена современная теория двойственности в задачах оптимизации.

Две статистические задачи

Метод максимального правдоподобия:

$$\text{максимизировать } P\{y \mid x\},$$

где $P\{y \mid x\}$ – вероятность y при условии x, x – подлежащие оценки параметры системы, x – наблюдаемые параметры системы.

Минимизация функционала среднего риска:

$$\text{минимизировать} \quad f(x),$$

где $F(x) = E_\omega F(x, \omega)$.

Основная сложность таких задач заключается в том, что ни функция f, ни F заранее не известны, есть только возможность измерить $F(x, \omega)$ со случайной реализацией ω.

Задачи трекинга

Дана последовательность функций $f_k : \mathcal{D} \to \mathbb{R}$, известно, что точки минимума x_k^*, x_{k+1}^* функций f_k и f_{k+1} отличаются не сильно, задача трекинга заключается в построении такой последовательности x_k, что для фиксированного $\epsilon > 0$ существует $K : \ \forall k \geq K$ выполняется

$$||x_k - x_k^*|| \leq \epsilon.$$

Пример: отслеживание положения движущегося объекта в пространстве по видеосъемке.

Седловые задачи

Оптимизация худшего случая:

$$\text{минимизировать } \max_y f(x, y)$$

Теория игр: два игрока ходят одновременно, первый может выбирать из множества X, второй – их множества Y. Пусть $f : X \times Y \to \mathbb{R}$, $f(x, y)$ – плата первого игрока второму при условии, что он выбрал x, а его соперник – y. Задача первого игрока:

$$\text{минимизировать } \max_y f(x, y).$$

Задача второго игрока:

$$\text{максимизировать } \min_x f(x, y)$$

Вариационные задачи

Пусть $f, g : \mathcal{D} \times \mathcal{U} \to \mathbb{R}$. Вариационные задачи обычно заключаются в нахождении минимума функционала на некотором множестве кривых и может быть сформулирована например следующим образом

$$\begin{array}{ll} \text{минизировать} & \int_a^b g(x(t), u(t))dt \\ \text{при условии} & \frac{d}{dt}x(t) = f(x(t), u(t)). \end{array}$$

Часто к этой задаче добавляются краевые условия вида $x(a) = x_0$, $x(b) = x_1$. Основной сутью этой задачи является выбор оптимальной траектории для некоторой динамической системы.

Рекомендованная литература

- Нестеров. Ю. Е. *Методы выпуклой оптимизации* , 2010 (перевод с англ. его же книги 1998 года *Introductory lectures on convex programming*)
- Поляк Б. Т. *Введение в оптимизацию*, 1986
- Boyd S., Vandenberghe L. *Convex Optimization*, 2004
- Luenberger D., Yinyu Y. *Linear and nonlinear optimization*, 2008