Математическое моделирование и методы оптимизации

Тремба Андрей Александрович
Москва, к.ф.-м.н., с.н.с.
Институт проблем управления РАН

21 июня 2019 г., Иннополис

Задача: познакомить с «оптимизацией»

10-часовой формат «краш»-курса по

- базовым понятиям,
- базовой теории,
- базовым методам,
- базовым инструментам,

... и немного математического моделирования.

Структура

1. Введение: моделирование и оптимизация
    - Формализация
    - Как решать
    - Когда остановиться
    - Типы задач

Структура

2. Теория: «геометрия», формулы и методы
   - Выпуклость
   - Градиент и градиентные методы
   - Безградиентные методы
   - Недифференцируемые функции

Структура (продолжение)

3. Типы задач: особенности и методы решения
   - Метод Ньютона
   - Промежуточные методы

Структура (продолжение)

4. Условная оптимизация
    - Простые методы
    - Функция Лагранжа и двойственность
    - CVX

Структура (продолжение)

5.
- Регуляризация
- Многокритериальная оптимизация
- Стохастические задачи/методы

Расписание

- **9:00 — 10:00** - введение
- **10:10 — 13:30** - теория и методы выпуклой оптимизации (основы)
- **14:10 — 16:45** - теория и методы («продвинутые»)
- **16:55 — 18:00** - задачи с ограничениями

Этапы в целом

1. Постановка задачи
2. Выбор способа решения
3. Нахождение решения
4. Анализ решения

Постановка: математическое моделирование

Предмет: физический процесс, структура, задача или объект.

Цель: добиться «результата».

Постановка: математическое моделирование

Средство: «математическое моделирование».

Итог: формальная формулировка задачи,

- математическое* описание объекта (модель),
- математическое описание цели.

Постановка: математическое моделирование

Предмет: физический процесс, структура, задача или объект.

Цель: добиться «результата».

Средство: «математическое моделирование».

Итог: формальная формулировка задачи,

- математическое* описание объекта (модель),
- математическое описание цели.

Типы моделей

- Статические

$$x \to M(x); F(\cdot) \Leftrightarrow f(x) = F(M(x))$$

- Динамические
  - Дискретные, например

$$x^{k+1} = Ax^k + Bu^k + Dw^k$$

  - Непрерывные, например

$$\dot{x}(t) = Ax(t) + Bu(t - \tau) + Dw(t))$$

Типы моделей

- Статические

$$x \to f(x)$$

- Динамические - оптимальное управление u
  Теория автоматического управления.
  - Дискретные, например
    $$x^{k+1} = Ax^k + Bu^k + Dw^k$$
  - Непрерывные, например
    $$\dot{x}(t) = Ax(t) + Bu(t-\tau) + Dw(t))$$

Типы моделей

- Статические - оптимизация по x

  Теория оптимизации.

$$x \to f(x)$$

- Динамические - оптимальное управление u

  Теория автоматического управления.

  - Дискретные, например

$$x^{k+1} = Ax^k + Bu^k + Dw^k$$

  - Непрерывные, например

$$\dot{x}(t) = Ax(t) + Bu(t-\tau) + Dw(t))$$

«Оптимизация» = «минимизация»*

$$\min_x f(x)$$

- Принятые обозначения, векторная запись (математическая база)
- Функция как чёрный ящик
- Как решать
- Когда остановиться
- Варианты постановок

Необходимые знания

- Линейная алгебра (векторы, матрицы)
- Математический анализ (пределы, множества, дифференцирование)
- Теория вероятности (моделирование, постановка задачи, методы)
- Численные методы

Книги для старта

- Поляк Б.Т. Введение в оптимизацию. М. Наука, 1983. (Переиздания 2014, 2019 гг.)
- Stephen Boyd and Lieven Vandenberghe Convex optimization. Cambridge University Press, 2004.
- А.В. Гасников Современные численные методы оптимизации. Метод универсального градиентного спуска. М. МФТИ, 2018.

Обозначения
и основные соотношения

Многоликое равенство

$:=$ — «присваивание»

$\doteq$, $\triangleq$ — «по определению», «обозначает»
$(f(x) \doteq x^2)$

$=$ — «условие» (уравнение)

$\Leftrightarrow$ — «означает то же, что и»,
«эквивалентное выражение»

$\equiv$ — «эквивалентно», «тожественно равно»
(e.g. $0 \equiv x - 1 \cdot x$)

Многоликое равенство

$:=$ — «присваивание»

$\doteq$, $\triangleq$ — «по определению», «обозначает»

$(f(x) \doteq x^2)$

$=$ — «условие» (уравнение)

$\Leftrightarrow$ — «означает то же, что и»,

«эквивалентное выражение»

$\equiv$ — «эквивалентно», «тожественно равно»

(e.g. $0 \equiv x - 1 \cdot x$)

$=$ — «равняется»

Базовые символы

Умножение

$$(...)\cdot(...),\quad (...)\times(...)$$

Размерность матриц:

"число строк × число столбцов": $n \times m$

Скобки:

- группировка или блочное конструирование

$$(...),[...]$$

- определение множеств, выбор альтернативы

$$\{x : ...\}$$
$$\min\{a, b, c...\}$$

Числа (скаляры)

$$\alpha, \beta, \gamma, \dots$$
$$a, b, c, \dots, x, y, z$$
$$i, j, k, \dots, n, m$$

$$a \in \mathbb{R},\ a \in \mathbb{C}$$

(Скалярные) функции:

$$\log(x), e^x, f(x) = x^i, \sin(x), p(x) = \sum_{i=0}^{n} c_i x^i$$

Векторы

$$a, b, c, ..., x, y, z$$
$$b \in \mathbb{R}^n$$

Компоненты

$$b_i \in \mathbb{R}, \quad i = 1, ..., n$$

Вектор-столбец (по умолчанию):

$$x \in \mathbb{R}^{n \times 1} \Leftrightarrow x \in \mathbb{R}^n$$

Вектор-строка $c \in \mathbb{R}^{1 \times n}$

$$c = [c_1, c_2, ..., c_n]$$

Операции между векторами

Векторы $a, b, x, y \in \mathbb{R}^m$

(Покомпонентное) сложение

$$x + y = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_m + y_m \end{bmatrix} \in \mathbb{R}^m$$

(Стандартное) скалярное произведение

$$(x, y) = (y, x) = \sum_{i=1}^{m} x_i y_i \in \mathbb{R}$$

~~Векторное произведение (только в $\mathbb{R}^3$)*~~.

Операции между числами и векторами

$$a, b \in \mathbb{R}, \quad x, y \in \mathbb{R}^\ell$$

Покомпонентное умножение (определение):

$$a \cdot x = x \cdot a = \begin{bmatrix} a \cdot x_1 \\ a \cdot x_2 \\ \vdots \\ a \cdot x_\ell \end{bmatrix}$$

свойства дистрибутивности (проверьте!)

$$(a + b) \cdot x = ax + bx \in \mathbb{R}^\ell$$
$$a(x + y) = ax + ay \in \mathbb{R}^\ell$$

Матрицы

$$A, B, C, D, ..., X, Y, Z \in \mathbb{R}^{m\times n}$$

m строк, n столбцов (размерность?)

Обозначение компонент:

$$a_{i,j} \text{ or } A_{i,j}$$

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & a_{2,3} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \cdots & \vdots \\ a_{m-1,1} & a_{m-1,2} & a_{m-1,3} & \ddots & a_{m-1,n} \\ a_{m,1} & a_{m,2} & a_{m,3} & \cdots & a_{m,n} \end{bmatrix} \in \mathbb{R}^{m\times n}$$

Векторы как особые матрицы.

Operations between matrices

Matrices $A, B \in \mathbb{R}^{m\times n}, C \in \mathbb{R}^{m\times k}$

Componentwise addition:

$$A+B = \begin{bmatrix} a_{1,1}+b_{1,1} & \cdots & a_{1,n}+b_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1}+b_{m,1} & \cdots & a_{m,n}+b_{m,n} \end{bmatrix} \in \mathbb{R}^{m\times n}$$

Inner (dot) product (standard):

$$(A, B) = \sum_{i=1}^{m} \sum_{j=1}^{n} A_{i,j} B_{i,j} \in \mathbb{R}$$

Операции между матрицами

$A, B, X \in \mathbb{R}^{k \times m}$

Аналогично вектор-векторым операциям: ...

Умножение покомпонентно

$$a \cdot B = B \cdot a = ... \in \mathbb{R}^{k \times m}$$

и дистрибутивно (проверяется):

$$(a + b) \cdot X = aX + bX \in \mathbb{R}^{k \times m}$$
$$a(A + B) = aA + aB \in \mathbb{R}^{k \times m}$$

деление на число:

$$C/a = ... \in \mathbb{R}^{k \times m}$$

Связь скаляров, векторов и матриц

Матрицы с единичной размерностью $\mathbb{R}^{k\times 1}$ or $\mathbb{R}^{1\times n}$ могут использоваться как векторы.

Матрицы размерности $\mathbb{R}^{1\times 1}$ и векторы размерностью $\mathbb{R}^{1}$ считаются скалярами.

Matlab, Octave, Scilab
Numpy*

Особая операция: матричное умножение

$A, B \in \mathbb{R}^{k\times m}, C \in \mathbb{R}^{m\times \ell}, R \in \mathbb{R}^{\ell\times n}$

Определение

$$A\cdot C = \begin{bmatrix} \sum_{j=1}^m a_{1,j}c_{j,1} & \sum_{j=1}^m a_{1,j}c_{j,2} & \cdots & \sum_{j=1}^m a_{1,j}c_{j,\ell} \\ \vdots & \vdots & \vdots & \vdots \\ \sum_{j=1}^m a_{k,j}c_{j,1} & \sum_{j=1}^m a_{k,j}c_{j,2} & \cdots & \sum_{j=1}^m a_{k,j}c_{j,\ell} \end{bmatrix} \in \mathbb{R}^{k\times}$$

Оно дистрибутивно (проверяется):

$$(A+B)\cdot C = AC + BC \in \mathbb{R}^{k\times \ell}$$

и ассоциативно:

$$A\cdot C\cdot R = (A\cdot C)\cdot R = A\cdot (C\cdot R) \in \mathbb{R}^{k\times n}$$

Как правило, $AB \neq BA$

Особые операции над матрицами

Транспонирование

$$(\cdot)^\top \text{ or } (\cdot)^T$$

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & a_{2,3} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{k,1} & a_{k,2} & a_{k,3} & \cdots & a_{k,n} \end{bmatrix} \in \mathbb{R}^{k\times n}$$

$$(A^T)_{i,j} = A_{j,i}, \quad i = 1, \ldots, n, \ j = 1, \ldots, k$$

$$A^T = \begin{bmatrix} a_{1,1} & a_{2,1} & \cdots & a_{k,1} \\ a_{1,2} & a_{2,2} & \cdots & a_{k,2} \\ a_{1,3} & a_{2,3} & \cdots & a_{k,3} \\ \vdots & \vdots & \vdots & \vdots \\ a_{1,n} & a_{2,n} & \cdots & a_{k,n} \end{bmatrix} \in \mathbb{R}^{n\times k}$$

Специальные матрицы

Квадратные матрицы

$$A \in \mathbb{R}^{m \times m}$$

Симметричные матрицы:

$$A \in \mathbb{R}^{m \times m}, \quad A_{i,j} = A_{j,i}$$

Положительно (полу) определённые матрицы:

$$\det M_i(A) > 0 \;\; (\det M_i(A) \geq 0), \;\; i = 1, \ldots, m$$

$$A \in \mathbb{S}_+ \Leftrightarrow A \succeq 0 \Leftrightarrow A \geq 0$$

$$A \in \mathbb{S}_{++} \Leftrightarrow A \succ 0 \Leftrightarrow A > 0$$

Единичная матрица:

$I \in \mathbb{R}^{k \times k} : I_{i,j} = 1$ если $i = j$, иначе 0, $\quad i, j = 1, \ldots$

Матричные функции

$$A \in \mathbb{R}^{m \times m}$$

Собственные векторы, собственные значения:

$$Ax_i = \lambda_i x_i, \;\; \lambda_i \in \mathbb{C}, \; x_i \in \mathbb{C}^m, \; i = 1, \ldots, m$$

След:

$$tr A = \sum_{i=1,\ldots,m} a_{i,i} = \sum_{i=1,\ldots,m} \lambda_i$$

Детерминант:

$$\det A = \prod_{i=1}^{m} \lambda_i \in \mathbb{R}$$

Скалярное произведение

- (Комплексно сопряжённая) симметрия

$$(y, x) = (x, y)$$

- Линейность по первому* аргументу

$$\begin{aligned} (ax, y) &= a(x, y) \\ (x + z, y) &= (x, y) + (z, y) \end{aligned}$$

- Положительная полуопределённость

$$\begin{aligned} (x, x) &\geq 0 \\ (x, x) = 0 &\Leftrightarrow x = 0 \end{aligned}$$

Задаёт норму: $\|x\| \doteq \sqrt{(x, x)}$

Стандартное скалярное произведение

$x, y \in \mathbb{R}^n$:

$$(x, y) = \sum_{i=1,n} x_i y_i = x^T y$$

Для конечномерных объектов

$$(S, R) = \sum_{i,j,k} S_{i,j,k} R_{i,j,k}$$

Существуют и иные скалярные произведения, например, взвешенные.

Многомерная геометрия

$x, c, d \in \mathbb{R}^n, \gamma, b \in \mathbb{R}$

Гиперплоскость:

$$(c, x) = b$$

Прямая:

$$x + \gamma d, \quad \gamma \in \mathbb{R}$$

Тела

Полупространство

$$(a, x) \leq b$$

Полиэдральное множество, многогранник

$$(a_i, x) \leq b_i,\ i = 1, ..., m$$

Параллелотоп («брус»)

$$|x_i| \leq c_i$$

Шар, заданный нормой

$$\|x\| \leq r$$

Эллипсоид

$$\{x : (x - c)^T Q(x - c) \leq r^2\},\ Q \succ 0$$

Ещё немного обозначений

- Производная (скалярная, векторых функций) $f'(\cdot)$
- (Цепное) правило вычисления сложных производных
$$\Big(f\big(g(x)\big)\Big)' = f'\big(g(x)\big)g'(x)$$
- Частная производная
$$f'_x(x, y, z, \ldots),\ f'_y(x, y, z, \ldots),\ \ldots\ x, y, z \in \mathbb{R}$$
или
$$\frac{\partial f(x_1, x_2, \ldots)}{\partial x_2},\ \frac{\partial}{\partial x_2} f(x_1, x_2, \ldots),\ \frac{\partial}{\partial x_2} f(x),\ \ x \in \mathbb{R}^n$$

Скалярные функции скалярных переменных

$$f : \mathbb{R} \to \mathbb{R}$$

Разложение в ряд Тейлора:

$$f(x + \Delta) = \sum_{k=0}^{\infty} f^{(k)}(x) \frac{\Delta^k}{k!}$$

$$f(y) = f(x) + f'(x)(y - x) + f''(x)\frac{(y-x)^2}{2} + \\ + \sum_{k=3}^{\infty} f^{(k)}(x)\frac{(y-x)^k}{k!}$$

Пределы и множества

Скалярный предел:

$$\{x^k\} : x^k \to x^+$$

Граница:

$$\partial Q \quad (\text{or } bnd\, Q)$$

Замкнутое множество:

$$\partial Q \subseteq Q$$

Пересечение $A \cap B$, объединение $X \cup Y$,
Сумма Минковского
$A \oplus B \doteq \{a + b : a \in A, b \in B\}$

Ключевые элементы

- Обозначения
- Скаляры, векторы, матрицы и операции над ними
- (Стандартное) скалярное произведение
- (Скалярные) производные

Как решать оптимизационные задачи

Ключевой слайд: как решать задачи

1. Записать задачу

$$\min_{x} f(x)$$

Ключевой слайд: как решать задачи

1. Записать задачу

$$\min_x f(x)$$

2. Получить ответ
   - MATLAB:
     `fminunc(f)`, `fminsearch(f)`, `fmincon(f)`
   - Python (Scipy):
     `scipy.optimize(f)`
   - CVX:
     `prob.solve()`, `minimize(·)`

Ключевой слайд: как решать задачи (не совсем)

1. Записать задачу

$$\min_x f(x)$$

2. Получить ответ
   - MATLAB:
     `fminunc(f)`, `fminsearch(f)`, `fmincon(f)`
   - Python (Scipy):
     `scipy.optimize(f)`
   - CVX:
     `prob.solve()`, `minimize(·)`

Примеры (a)

$$f(x) = \frac{1}{x}$$

$$\min f(x) = ?$$

Примеры (б)

$$f(x) = \frac{1}{x}, \quad x \geq 0$$

$$\min f(x) = ?$$

Примеры (в)

$$f(x) = \frac{1}{x}, \quad a \geq x \geq 0$$

$$\min f(x) = ?$$

Примеры (г)

$$f(x) = \frac{1}{x}, \quad |x| > a$$

$$\min f(x) =?$$

Примеры (д)

$$f(x) = \frac{1}{x}, \quad |x| \geq a$$

$$\min f(x) = ?$$

Если что-то пошло не так...

Нужен метод:

«Путь, способ, приём теоретического исследования или практического осуществления чего-нибудь».

Цель (в контексте оптимизации)

$$f^* = \min_x f(x) \quad \left(f^* = \inf_x f(x)\right)$$

$$x^* = \arg\min f(x)$$

Алгоритм

Последовательность:

$$x^0 \to x^1 \to \ldots \to x^k \to \ldots$$

$$\{x^k\}$$

Цель:

$$\{x^k\} \to x^*$$

$$f(x^k) \to f(x^*) = f^*$$

Алгоритм

Последовательность:

$$x^0 \to x^1 \to \ldots \to x^k \to \ldots$$

$$\{x^k\}$$

Цель:

$$\{x^k\} \to x^*$$

$$f(x^k) \to f(x^*) = f^*$$

Вариант: «релаксационная последовательность»

$$f(x^{k+1}) \le f(x^k)$$

Критерии остановки

Есть $\{x^k\}, \quad k = 0, 1, \dots$

- $$\|x^k - x^*\| \to 0$$
- $$f(x^k) - f^* \to 0$$
- $$f(x^{k-1}) - f(x^k) \to 0$$
- $$\|x^k - x^{k-1}\| \to 0$$

Функция как чёрный ящик

$$\widehat{x} = \text{Алгоритм}(\ldots)$$

Функция как чёрный ящик

$$\widehat{x} = \text{Алгоритм}(...) \approx x^*$$

$$x^{k+1} = \text{Очередной\_шаг}(\{x^k\}, ...)$$

Функция как чёрный ящик

$$\widehat{x} = \text{Алгоритм}(\dots)$$

$$x^{k+1} = \text{Очередной\_шаг}(\{x^k\}, \dots)$$

- $f(x^k)$
- $f'(x^k), \nabla f(x^k)$
- $\nabla^2 f(x^k)$
- Мета-информация о функции и ограничениях.

Типы задач оптимизации

- Выпуклые и невыпуклые
- Гладкие и негладкие
- С ограничениями и без ограничений (условные/безусловные)
- Линейное и нелинейное программирование
- Детерминированные и стохастические
- Непрерывные и дискретные
- Точные и робастные
- и т.д.

Выводы по первой части

1. Подготовка к оптимизации
2. Векторы и обозначения
3. Как «решить» любою задачу оптимизации
4. Как не решить задачу оптимизации
5. «Чёрный ящик» и критерии остановки алгоритма

Далее

«Минимизация»,

«непрерывная»,

«конечномерная»,

методами

- нулевого порядка,
- первого порядка,
- второго порядка,

Далее

«Минимизация»,

«непрерывная»,

«конечномерная»,

методами

- нулевого порядка,
- первого порядка,
- второго порядка,

в основном - «выпуклые» задачи.

Что будет интересовать

- Тип задачи
- Фундаментальные идеи и подходы
- Выбор подходящего метода решения
- Эффективность методов
- Преобразование задач