Методы оптимизации
Оптимизация с ограничениями неравенства. Выпуклое программирование

Н. В. Артамонов

МГИМО МИД России

6 октября 2023 г.

Содержание

1. Постановка задачи. Выпуклая оптимизация
2. Функция Лагранжа
3. Функция Лагранжа в форме Куна-Таккера
4. Пример: Consumption–Leisure choice

Задача максимизации

$$\max f(\boldsymbol{x})$$
$$s.t. \begin{cases} g_1(\boldsymbol{x}) \leq C_1 \\ \vdots \\ g_k(\boldsymbol{x}) \leq C_k \end{cases} \tag{1}$$

Важно!

Согласованность знаков неравенства в ограничениях с максимизацией целевой функции.

Задача минимизации

$$\min f(\boldsymbol{x})$$
$$s.t. \begin{cases} g_1(\boldsymbol{x}) \geq C_1 \\ \vdots \\ g_k(\boldsymbol{x}) \geq C_k \end{cases} \tag{2}$$

Важно!

Как и раньше согласованность знаков неравенства в ограничениях с минимизацией целевой функции.

Задача минимизации

$$\min f(\boldsymbol{x})$$
$$s.t. \begin{cases} g_1(\boldsymbol{x}) \geq C_1 \\ \vdots \\ g_k(\boldsymbol{x}) \geq C_k \end{cases} \tag{2}$$

Важно!

Как и раньше согласованность знаков неравенства в ограничениях с минимизацией целевой функции.

Очевидно, что задача (2) сводится к задаче (1) с целевой функцией $(-f)$ и с ограничениями $-g_j(\boldsymbol{x}) \leq -C_j$.

Задача выпуклой оптимизации

Для глобального экстремума потребуется выпуклость функций в задаче оптимизации.

Определение

Экстремальная задача (1) *называется задачей выпуклого программирования или выпуклой оптимизации, если*

- f *вогнута*
- $g_1, \ldots, g_k$ *выпуклы*

Определение

Экстремальная задача (2) *называется задачей выпуклого программирования или выпуклой оптимизации, если*

- f *выпукла*
- $g_1, \ldots, g_k$ *вогнуты*

Матричная запись ограничений

Обозначим

$$\boldsymbol{g}(\boldsymbol{x}) = \begin{pmatrix} g_1(\boldsymbol{x}) \\ \vdots \\ g_k(\boldsymbol{x}) \end{pmatrix} \qquad \boldsymbol{c} = \begin{pmatrix} C_1 \\ \vdots \\ C_k \end{pmatrix}$$

Перепишем задачи (1), (2) в виде

$$\begin{aligned} \max f(\boldsymbol{x}) \\ s.t.\ \boldsymbol{g}(\boldsymbol{x}) \le \boldsymbol{c} \end{aligned} \qquad \begin{aligned} \min f(\boldsymbol{x}) \\ s.t.\ \boldsymbol{g}(\boldsymbol{x}) \ge \boldsymbol{c} \end{aligned}$$

1 Постановка задачи. Выпуклая оптимизация

2 Функция Лагранжа

3 Функция Лагранжа в форме Куна-Таккера

4 Пример: Consumption–Leisure choice

Функция Лагранжа

Функцией Лагранжа для задачи (1) и задачи (2) имеет вид

$$\mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) - \sum_{j=1}^{k} \lambda_j (g_j(\boldsymbol{x}) - C_j) = f(\boldsymbol{x}) - \boldsymbol{\lambda}^\top (\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{c})$$

Дополнительные переменные $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_k)^\top$ называются множителями Лагранжа (Lagrange multipliers).

Условие Слейтера

Определение

Будем говорить, что для ограничений задачи (1) *выполнено* *условие Слейтера*, *если существует* $\boldsymbol{x}_0 \in \mathbb{R}^n$ *такой, что*

$$\boldsymbol{g}(\boldsymbol{x}_0) < \boldsymbol{c}$$

Будем говорить, что для ограничений задачи (2) *выполнено* *условие Слейтера*, *если существует* $\boldsymbol{x}_0 \in \mathbb{R}^n$ *такой, что*

$$\boldsymbol{g}(\boldsymbol{x}_0) > \boldsymbol{c}$$

Замечание

Условие Слейтера относится только к системе ограничений и не касается целевой функции

Основной результат для (1)

Теорема

Пусть задача (1) *является задачей выпуклого программирования и для ограничений выполнено условие Слейтера.*

Тогда $\hat{\mathbf{x}}$ *является решением экстремальной задачи тогда и только тогда, когда он удовлетворяет системе*

$$\begin{cases} \mathcal{L}'_{x_i} = 0 & i = 1, \ldots, n \\ \lambda_j \mathcal{L}'_{\lambda_j} = 0 & j = 1, \ldots, k \\ \lambda_j \geq 0, \mathcal{L}'_{\lambda_j} \geq 0 & j = 1, \ldots, k \end{cases} \tag{3}$$

Кроме того, $\hat{\mathbf{x}}$ *– глобальный максимум в задаче* (1)*.*

Замечание

В отличие от задачи оптимизации с ограничениями равенства здесь важен знак λ_j

Необходимые условия минимума

Следствие (Необходимые условия минимума)

Чтобы точка $\mathbf{x}$ *была локальным минимумом в задаче* (2) *необходимо, чтобы для некоторых чисел* $\hat{\lambda}_1, ..., \hat{\lambda}_k$ *выполнялась система*

$$\begin{cases} \mathcal{L}'_{x_i} = 0 & i = 1, \ldots, n \\ \lambda_j \mathcal{L}'_{\lambda_j} = 0 & j = 1, \ldots, k \\ \lambda_j \geq 0, \mathcal{L}'_{\lambda_j} \leq 0 & j = 1, \ldots, k \end{cases} \tag{4}$$

Кроме того, $\hat{\mathbf{x}}$ *– глобальный минимум в задаче* (11).

Определение

Условие $\lambda_j \mathcal{L}'_{\lambda_j} = 0$ *называются* *условием дополняющей нежесткости*

Интерпретация множителей Лагранжа

Пусть $\hat{x}$ – решение задачи (1) или задачи (2). Рассмотрим это решение как функцию от $C_1, \dots, C_k$:

$$\hat{x} = \hat{x}(C_1, \dots, C_k) \qquad \hat{f} = f(\hat{x}) = \hat{f}(C_1, \dots, C_k)$$

Теорема

Оптимальное решение есть гладкая функция от $C_1, \dots, C_k$ и выполнены равенства

$$\frac{\partial \hat{f}}{\partial C_j} = \hat{\lambda}_j \qquad j = 1, \dots, k$$

1 Постановка задачи. Выпуклая оптимизация

2 Функция Лагранжа

3 Функция Лагранжа в форме Куна-Таккера

4 Пример: Consumption–Leisure choice

Задача максимизации

Рассмотрим задачу максимизации с ограничениями неравенства в следующей форме

$$
\begin{gathered}
\max\ f(\boldsymbol{x}) \\
s.t. \begin{cases} g_j(\boldsymbol{x}) \leq C_j & j = 1, \ldots, k \\ x_i \geq 0 & i = 1, \ldots, n \end{cases}
\end{gathered}
\tag{max}
$$

Особенности

- Всего $n + k$ ограничений
  - k «нетривиальных» ограничений
  - n ограничений неотрицательности переменных
- согласованность знаков неравенств в ограничениях с максимизацией целевой функции

Задача минимизации

Аналогично для задачи минимизации:

$$\min\ f(\mathbf{x})$$
$$s.t. \begin{cases} g_j(\mathbf{x}) \geq C_j & j = 1, \ldots, k \\ x_i \geq 0 & i = 1, \ldots, n \end{cases} \tag{min}$$

Особенности

- Всего $n + k$ ограничений
  - k «нетривиальных» ограничений
  - n ограничений неотрицательности переменных
- согласованность знаков неравенств в ограничениях с минимизацией целевой функции

Задача выпуклой оптимизации

Для глобального экстремума потребуется выпуклость функций в задаче оптимизации.

Определение

Экстремальная задача (max) *называется* *задачей выпуклого программирования* *или* *выпуклой оптимизации, если*

- f *вогнута*
- $g_1, \ldots, g_k$ *выпуклы*

Определение

Экстремальная задача (min) *называется* *задачей выпуклого программирования* *или* *выпуклой оптимизации, если*

- f *выпукла*
- $g_1, \ldots, g_k$ *вогнуты*

Матричная запись ограничений

Обозначим

$$\boldsymbol{g}(\boldsymbol{x}) = \begin{pmatrix} g_1(\boldsymbol{x}) \\ \vdots \\ g_k(\boldsymbol{x}) \end{pmatrix} \qquad \boldsymbol{c} = \begin{pmatrix} C_1 \\ \vdots \\ C_k \end{pmatrix}$$

Перепишем задачи (max), (min) в виде

$$\max f(\boldsymbol{x}) \qquad\qquad \min f(\boldsymbol{x})$$

$$s.t. \begin{cases} \boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{c} \\ \boldsymbol{x} \geq 0 \end{cases} \qquad\qquad s.t. \begin{cases} \boldsymbol{g}(\boldsymbol{x}) \geq \boldsymbol{c} \\ \boldsymbol{x} \geq 0 \end{cases}$$

Функция Лагранжа в форме Куна-Таккера

Определим для задач (max), (min) функцию Лагранжа в форме Куна – Таккера, включив в нее только нетривиальные ограничения

$$\mathcal{Z}(\boldsymbol{x}, \boldsymbol{\mu}) = f(\boldsymbol{x}) - \sum_{j=1}^{k} \mu_j (g_j(\boldsymbol{x}) - C_j) = f(\boldsymbol{x}) - \boldsymbol{\mu}^\top (\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{c})$$

с множителями Лагранжа

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_k \end{pmatrix}$$

Условие Слейтера

Определение

Будем говорить, что для ограничений задачи (max) *выполнено* *условие Слейтера*, *если существует* $\boldsymbol{x}_0 \in \mathbb{R}^n$ *такой, что*

$$\boldsymbol{g}(\boldsymbol{x}_0) < \boldsymbol{c} \qquad \boldsymbol{x}_0 > 0$$

Будем говорить, что для ограничений задачи (min) *выполнено* *условие Слейтера*, *если существует* $\boldsymbol{x}_0 \in \mathbb{R}^n$ *такой, что*

$$\boldsymbol{g}(\boldsymbol{x}_0) > \boldsymbol{c} \qquad \boldsymbol{x}_0 > 0$$

Замечание

Условие Слейтера относится только к системе ограничений и не касается целевой функции

Основной результат для (max)

Теорема (Kuhn – Tucker, 1951)

Пусть задача (max) *является задачей выпуклого программирования и для ограничений выполнено условие Слейтера.*

Тогда $\hat{x}$ *является решением экстремальной задачи тогда и только тогда, когда он удовлетворяет системе*

$$\begin{cases} x_i \mathcal{Z}'_{x_i} = 0 & i = 1, \ldots, n \\ \mu_j \mathcal{Z}'_{\mu_j} = 0 & j = 1, \ldots, k \\ x_i \geq 0, \ \mathcal{Z}'_{x_i} \leq 0 & i = 1, \ldots, n \\ \mu_j \geq 0, \ \mathcal{Z}'_{\mu_j} \geq 0 & j = 1, \ldots, k \end{cases} \tag{5}$$

Кроме того, $\hat{x}$ *– глобальный максимум в задаче* (max).

Основной результат для (min)

Теорема (Kuhn – Tucker, 1951)

Пусть задача (min) *является задачей выпуклого программирования и для ограничений выполнено условие Слейтера.*

Тогда $\hat{x}$ *является решением экстремальной задачи тогда и только тогда, когда он удовлетворяет системе*

$$\begin{cases} x_i \mathcal{Z}'_{x_i} = 0 & i = 1, \dots, n \\ \mu_j \mathcal{Z}'_{\mu_j} = 0 & j = 1, \dots, k \\ x_i \geq 0, \ \mathcal{Z}'_{x_i} \geq 0 & i = 1, \dots, n \\ \mu_j \geq 0, \ \mathcal{Z}'_{\mu_j} \leq 0 & j = 1, \dots, k \end{cases} \tag{6}$$

Кроме того, $\hat{x}$ *– глобальный минимум в задаче* (min).

Некоторые замечания

Системы (5), (6) имеют $n + k$ уравнений и содержат $2n + 2k$ неравенств для отбора корней.

Алгоритм решения задач (max), (min)

- Проверяем выпуклость/вогнуть целевой функции и ограничений
- Проверяем условие Слейтера для ограничений
- Решаем систему (5) или (6)

Интерпретация множителей Лагранжа

Множители Лагранжа μ_j имеют простую экономическую интерпретацию как предельные значения.

Рассмотрим оптимальное решение задачи (max) и задачи (min) и оптимальное значение целевой функции как функцию от $C_1, \ldots, C_k$:

$$\hat{\boldsymbol{x}} = \hat{\boldsymbol{x}}(C_1, \ldots, C_k) \qquad \hat{f} = f(\hat{\boldsymbol{x}}) = \hat{f}(C_1, \ldots, C_k).$$

Интерпретация множителей Лагранжа

Множители Лагранжа μ_j имеют простую экономическую интерпретацию как предельные значения.

Рассмотрим оптимальное решение задачи (max) и задачи (min) и оптимальное значение целевой функции как функцию от $C_1, \ldots, C_k$:

$$\hat{x} = \hat{x}(C_1, \ldots, C_k) \qquad \hat{f} = f(\hat{x}) = \hat{f}(C_1, \ldots, C_k).$$

Теорема

$\hat{x} = \hat{x}(C_1, \ldots, C_k)$ *есть гладкая функция и выполнены равенства*

$$\mu_j = \frac{\partial \hat{f}}{\partial C_j} \qquad j = 1, \ldots, k$$

1 Постановка задачи. Выпуклая оптимизация

2 Функция Лагранжа

3 Функция Лагранжа в форме Куна-Таккера

4 Пример: Consumption–Leisure choice

Постановка задачи

Экономический агент имеет два «товара»: «отдых» l (leisure, в часах) и потребление x.

Пусть w – почасовая оплата и P – цена потребления.

Агент располагает общим временем H, которое он может тратить на работу и на отдых, и также имеет фиксированный доход M (non-labor income).

Функция полезности экономического агента $U(x, l)$.

Постановка задачи

Рассмотрим задачу оптимизации

$$\max \; U(x, l)$$

$$s.t. \begin{cases} Px + wl \le wH + M \\ 0 \le l \le H \\ x \ge 0 \end{cases}$$

Вопрос: какая интерпретация у такой постановки?

Функция Лагранжа в форме К.-Т.

Перепишем в виде задачи (max)

$$\max\ U(x,l)$$
$$s.t. \begin{cases} Px + wl \le wH + M \\ l \le H \\ x, l \ge 0 \end{cases}$$

Функция Лагранжа в форме К.-Т.

Перепишем в виде задачи (max)

$$\max\ U(x,l)$$
$$s.t.\begin{cases} Px+wl\le wH+M \\ l\le H \\ x,l\ge 0\end{cases}$$

Функция Лагранжа имеет вид

$$\mathcal{Z}(x,l,\mu_1,\mu_2)=U(x,l)-\mu_1(Px+wl-wH-M)-\mu_2(l-H)$$

и

$$\mathcal{Z}'_x=U'_x-\mu_1P \qquad \mathcal{Z}'_l=U'_l-\mu_1w-\mu_2$$

Consumption–Leisure choice как задача выпуклой оптимизации

Ограничения задаются линейными функциями $\Rightarrow$ выпуклыми функциями.

Если функция полезности U вогнута, то задача Consumption–Leisure choice есть задача выпуклой оптимизации.

Consumption–Leisure choice как задача выпуклой оптимизации

Ограничения задаются линейными функциями $\Rightarrow$ выпуклыми функциями.

Если функция полезности U вогнута, то задача Consumption–Leisure choice есть задача выпуклой оптимизации.

Вывод

Если функция полезности U вогнута, то решение системы уравнений-неравенств (5) есть решение задачи оптимизации.

Система К.-Т. для Consumption–Leisure choice

Система (5) имеет вид

$$
\begin{cases}
x_i \mathcal{Z}'_{x_i} = 0 \\
\mu_j \mathcal{Z}'_{\mu_j} = 0 \\
\mathcal{Z}'_{x_i}, \mathcal{Z}'_{\mu_j} \le 0 \\
x_i, \mu_j \ge 0
\end{cases}
\Rightarrow
\begin{cases}
x(U'_x - \mu_1 P) = 0 \\
l(U'_l - \mu_1 w - \mu_2) = 0 \\
-\mu_1(Px + wl - wH - M) = 0 \\
-\mu_2(l - H) = 0 \\
U'_x - \mu_1 P \le 0 \\
U'_l - \mu_1 w - \mu_2 \le 0 \\
wH + M - Px - wl \ge 0 \\
H - l \ge 0 \\
x, l, \mu_1, \mu_2 \ge 0
\end{cases}
\tag{7}
$$

Общие свойства решения сситемы (7)

Если $\mu_1 = 0$, то

$$U'_x - \mu_1 P \leq 0 \Rightarrow U'_x \leq 0$$

Общие свойства решения сситемы (7)

Если $\mu_1 = 0$, то

$$U'_x - \mu_1 P \le 0 \Rightarrow U'_x \le 0$$

Но! Стандартные условия на функцию полезности: $U'_x, U'_l > 0$.

Противоречие!

Общие свойства решения ссистемы (7)

Если $\mu_1 = 0$, то

$$U'_x - \mu_1 P \leq 0 \Rightarrow U'_x \leq 0$$

Но! Стандартные условия на функцию полезности: $U'_x, U'_l > 0$.

Противоречие!

Следовательно всегда $\mu_1 > 0$ и на оптимальном решении

$$Px + wl = wH + M$$

т.е. ограничение будет равенством.

Как решать систему (7)?

Так как $x, l \geq 0$, то нужно последовательно рассматривать случаи:

1. $x, l > 0$;
2. $x > 0, l = 0$;

Как решать систему (7)?

Так как $x, l \geq 0$, то нужно последовательно рассматривать случаи:

1. $x, l > 0$;
2. $x > 0, l = 0$;

Важно!

Случай $x = 0$ не реализуется

Случай $x = 0$

В самом деле, если $x = 0$, то

$$Px + wl - wH - M = w(l - H) - M < -M < 0 \quad (\text{т.к. } l \leq H)$$

Случай $x = 0$

В самом деле, если $x = 0$, то

$$Px + wl - wH - M = w(l - H) - M < -M < 0 \quad (\text{т.к. } l \leq H)$$

Из третьего уравнения (7)

$$\mu_1(Px + wl - wH - M) = 0 \Rightarrow \mu_1 = 0$$

Случай $x = 0$

В самом деле, если $x = 0$, то

$$Px + wl - wH - M = w(l - H) - M < -M < 0 \quad (\text{т.к. } l \leq H)$$

Из третьего уравнения (7)

$$\mu_1(Px + wl - wH - M) = 0 \Rightarrow \mu_1 = 0$$

Тогда

$$U'_x - \mu_1 P \leq 0 \Rightarrow U'_x \leq 0$$

Но! Стандартные условия на функцию полезности: $U'_x, U'_l > 0$.

Противоречие!

Случай $l = 0, x > 0$

Так как $Px + wl = wH + M$, то $\hat{x} = (M + wH)/P$.

Случай $l = 0, x > 0$

Так как $Px + wl = wH + M$, то $\hat{x} = (M + wH)/P$.

Из первого уравнения системы (7):

$$U'_x = \mu_1 P \Rightarrow \hat{\mu}_1 = U'_x(\hat{x}, 0)/P.$$

Случай $l = 0, x > 0$

Так как $Px + wl = wH + M$, то $\hat{x} = (M + wH)/P$.

Из первого уравнения системы (7):

$$U'_x = \mu_1 P \Rightarrow \hat{\mu}_1 = U'_x(\hat{x}, 0)/P.$$

Далее $\mu_2(l - H) = 0 \Rightarrow \hat{\mu}_2 = 0$.

Случай $l = 0, x > 0$

Так как $Px + wl = wH + M$, то $\hat{x} = (M + wH)/P$.

Из первого уравнения системы (7):

$$U'_x = \mu_1 P \Rightarrow \hat{\mu}_1 = U'_x(\hat{x}, 0)/P.$$

Далее $\mu_2(l - H) = 0 \Rightarrow \hat{\mu}_2 = 0$.

Осталось неравенство $U'_l - \mu_1 w - \mu_2 \leq 0$.

Случай $l = 0, x > 0$

Так как $Px + wl = wH + M$, то $\hat{x} = (M + wH)/P$.

Из первого уравнения системы (7):

$$U'_x = \mu_1 P \Rightarrow \hat{\mu}_1 = U'_x(\hat{x}, 0)/P.$$

Далее $\mu_2(l - H) = 0 \Rightarrow \hat{\mu}_2 = 0$.

Осталось неравенство $U'_l - \mu_1 w - \mu_2 \leq 0$.

Вывод

$(\hat{x}, 0)$ – решение задачи Consumption-Leisure Choice $\iff$

$$PU'_l(\hat{x}, 0) \leq wU'_x(\hat{x}, 0).$$

Случай $x, l > 0$

Получим систему

$$\begin{cases} U'_x - \mu_1 P = 0 \\ U'_l - \mu_1 w - \mu_2 = 0 \\ Px + wl = wH + M \\ \mu_2(l - H) = 0 \\ \mu_1, \mu_2, x, l \geq 0 \\ l \leq H \end{cases}$$

Случай $x, l > 0$

Получим систему

$$\begin{cases} U'_x - \mu_1 P = 0 \\ U'_l - \mu_1 w - \mu_2 = 0 \\ Px + wl = wH + M \\ \mu_2(l - H) = 0 \\ \mu_1, \mu_2, x, l \geq 0 \\ l \leq H \end{cases}$$

Два случая

1. $\hat{l} = H$ и $\hat{x} = M/P$;
2. $l < H$ и $\mu_2 = 0$.