Методы оптимизации. Теоремы.

Кирилл Захаров

2021 г.

Содержание

1	Линеиное программирование. Базисное решение, допустимое множество, оптимальное решение	1
2	Линейное программирование. Двойственная задача	1
3	Задача нелинейной безусловной оптимизации	2
4	Задача нелинейной условной оптимизации	3
5	Задача выпуклой оптимизации	4
6	Задача выпуклой квадратичной оптимизации	9

1 Линейное программирование. Базисное решение, допустимое множество, оптимальное решение

Теорема 1.1. Множество допустимых решений есть выпуклое множество.

Лемма 1.1. Базисные решения являются вершинами выпуклой многогранной области.

Теорема 1.2. Оптимальное решение является базисным решением. (Оптимальное решение лежит в углах выпуклой многогранной области).

2 Линейное программирование. Двойственная задача

Теорема 2.1 (Основное неравенство двойственности). Пусть заданы прямая задача $D: X \ f(X)$ и двойственная $\Omega: \Lambda \ \varphi(\Lambda)$. Тогда для любых допустимых планов прямой и двойственной задачи их целевые функции связаны неравенствами.

$$f(X) \to \min \Rightarrow f(X) \geqslant \varphi(\Lambda)$$

$$f(X) \to \max \Rightarrow f(X) \leqslant \varphi(\Lambda)$$
(1)

Теорема 2.2 (Критерий оптимальности Канторович). Если на допустимых планах прямой X и двойственной задачи Λ значения их целевых функций совпадают, то планы X и Λ являются оптимальными и наоборот.

Теорема 2.3. Для существования оптимального плана как прямой, так и двойственной задач \iff существование какого-либо допустимого плана для каждой из этих задач.

Теорема 2.4. Если прямая задача имеет оптимальное решение, то и двойственная имеет оптимальное решение.

Теорема 2.5. Если прямая задача не имеет решения из-за неограниченности целевой функции, то система ограничений двойственной задачи противоречива.

Теорема 2.6 (О дополняющей нежесткости). *Необходимым и достаточным условием того, что прямая и двойственная задачи имеют оптимальное решение, является выполнение условий дополняющей нежесткости.*

$$\lambda_j \left(\sum_{i=1}^N a_{ji} x_i - b_j \right) = 0$$

$$x_i \left(\sum_{j=1}^M a_{ji} \lambda_i - c_i \right) = 0$$
(2)

3 Задача нелинейной безусловной оптимизации

 $x \in O \subseteq \mathbb{R}^N$

Определение 3.1. $Y=(y_1,...,y_N)$ - точка локального минимума или максимума, если $\exists \ \varepsilon>0,$ такое что выполняется

$$f(Y) \leqslant f(Y + \delta X)$$
 или $f(Y) \geqslant f(Y + \delta X)$ (3)

для всех $\delta X = (\delta x_1, ..., \delta x_N) |0 < |\delta x_i| < \varepsilon$.

Определение 3.2. Y - точка строгого экстремума, если неравенства выполняются строго.

Определение 3.3. Y называется точкой глобального экстремума, если неравенства (3) выполняются во всей области.

$$\min f(x) = \max -f(X)$$

Определение 3.4. Функция, имеющая единственный экстремум называется унимодальной.

Лемма 3.1. Если область допустимых значений, определяемая системой ограничений равенств, содержит точку Y и ее окрестность, то M < N.

$$Y \subseteq D \land U_{\varepsilon}(Y) \subseteq D \Rightarrow M < N \tag{4}$$

Пемма 3.2. Пусть область допустимых значений, определяемая системой ограничений равенств задачи на условный экстремум, содержит хотя бы одну точку Y. Если набор градиентов $\operatorname{grad} \psi_j$ линейно независим u $\operatorname{rank} J = M < N$, то D вместе c каждой точкой X содержит некоторую непустую ее окрестность.

Теорема 3.1. Пусть задана функция f(x) и $x \in O = \mathbb{R}^1$. Если в точке Y функция f(x) имеет локальный экстремум, то $\frac{\partial f(Y)}{\partial x} = 0$.

Теорема 3.2 (Необходимое условие экстремума 1-го порядка). Пусть задана функция f(X) и $X \in O = \mathbb{R}^N$. Пусть Y точка локального экстремума. Тогда $grad\ f(Y) = 0$.

Теорема 3.3 (Критерий Сильвестра).

- 1. Матрица A является положительно определенной \iff когда все ее угловые миноры больше 0;
- 2. Матрица A является отрицательно определенной \iff когда все ее угловые миноры образуют знакочередующийся ряд, начиная со знака «-»;
- 3. Матрица A является положительно полуопределенной $\iff A$ вырождена и все ее главные миноры $m_i(A) \geqslant 0$;
- 4. Матрица A является отрицательно полуопределенной \iff $m_i(A)=0$ или $sign\ m_i(A)=sign\ (-1)^i$.

Теорема 3.4 (Необходимое условие экстремума 2-го порядка). Пусть задана функция f(X) $X \in \mathbb{R}^N$. Пусть f(X) дважды дифференцируема в окрестности точки Y. Тогда если Y - точка локального минимума (максимума), то $H_f(Y)$ положительно полуопределенная (отрицательно полуопределенная).

Теорема 3.5 (Достаточное условие экстремума 2-го порядка). Пусть задана функция f(X) $X \in \mathbb{R}^N$. Пуста f(X) имеет стационарную точку, в которой вторые частные производные существуют и непрерывны. Если $H_f(Y)$ положительно определена (отрицательно определена), то Y точка минимума (максимума).

Теорема 3.6.

- 1. Пусть f(X) дифференцируема в точке $Y \in \mathbb{R}^N$. Тогда если $\delta X \in \mathbb{R}^N \mid \operatorname{grad} f(Y) \cdot \delta X < (>)0 \Rightarrow \delta X \in W_-(Y,f)(W_+(Y,f));$
- 2. Если $\delta X \in W_-(Y,f)(W_+(Y,f))$. Тогда $\operatorname{grad} f(Y) \cdot \delta X \leqslant (\geqslant) 0$.

4 Задача нелинейной условной оптимизации

Теорема 4.1 (Связь между $W_{+/-}(Y,f)$ и V(Y,f)). Если точка Y точка локального минимума (максимума), то $W_{-}(Y,f) \cap V(Y,f) = \emptyset$ ($W_{+}(Y,f) \cap V(Y,f) = \emptyset$).

Теорема 4.2 (Вейерштрасс). Пусть D - компакт u f(X) непрерывная функция определенная на D.

Тогда существует точка Y глобального минимума (максимума).

Теорема 4.3 (Необх. условие 1-го рода. Правило множителей Лагранжа). Пусть $Y \in D \subseteq \mathbb{R}^N$ - точка локального экстремума. Пусть $f(X), \psi_j(X)$ - непрерывно дифференцируемы и пусть в точке Y $J(Y) = \left\{ \frac{\partial \psi_j(Y)}{\partial x_i} \right\}$ имеет ранг равный M.

Тогда существуют неравные одновременно нулю вектор Λ' и $\lambda'_0 \mid$ точка $(\Lambda', \lambda'_0, Y)$ - стационарная точка функции Лагранжа, т.е. $\operatorname{grad} L(\Lambda', \lambda'_0, Y) = 0$

Теорема 4.4 (Необх. условие 2-го рода). Пусть $Y \in D \subseteq \mathbb{R}^N$ - точка локального минимума (максимума). Пусть $f(X), \psi_j(X)$ - дважды непрерывно дифференцируемы и пусть в точке Y $J(Y) = \left\{ \frac{\partial \psi_j(Y)}{\partial x_i} \right\}$ имеет ранг равный M.

Тогда в стационарной точке функции Лагранжа $(Y) \, \forall \, \delta X \neq 0 \, | \operatorname{grad} \psi_j(X) \cdot \delta X = 0$ выполняется неравенство $\delta X L_{XX}''(\Lambda',Y) \delta X^T \geqslant (\leqslant) 0$.

Теорема 4.5 (Достаточное условие экстремума). Пусть $Y \in D \subseteq \mathbb{R}^N$ - точка экстремума $u \ \psi_j(X) = 0$. Пусть $f(X), \psi_j(X)$ - дважды непрерывно дифференцируемы. Если существуют $\Lambda' = (\lambda_1, ..., \lambda_M) \neq 0 \land \lambda'_0 \neq 0 \ | \operatorname{grad} L(\Lambda', \lambda'_0, Y) = 0 \ u \ \operatorname{npu}$ этом $\delta X L''_{XX}(\Lambda', Y) \delta X^T > (<)0 \ \forall \ \delta X \neq 0 \ \partial \Lambda S \ \operatorname{somophix} \operatorname{grad} \psi_j(Y) \delta X = 0$. Тогда Y - точка локального минимима (максимума).

Теорема 4.6 (Достаточное условие экстремума в терминах матрицы Гессе функции Лагранжа). Пусть найдена стационарная точка функции Лагранжа.

Y - точка максимума, если начиная с углового минора порядка 2M+1 последующие N-M угловых миноров матрицы Гессе образуют знакочередующийся числовой ряд в котором знак первого члена совпадает со знаком $(-1)^{M+1}$.

Y - точка минимума, если начиная с углового минора порядка 2M+1 последующие N-M угловых миноров матрицы Гессе имеют знак $(-1)^M$.

5 Задача выпуклой оптимизации

Пемма 5.1. Пересечение конечного или счетного числа выпуклых множеств есть выпуклое множество.

Лемма 5.2. Линейная комбинация $\sum_{i=1}^{N} \alpha_i X_i$ конечного числа выпуклых множеств X_i при любых α_i является выпуклым множеством.

Лемма 5.3. Если $f_1(X), f_2(X), ..., f_M(X)$ выпуклы (вогнуты) на выпуклом множестве D, то их линейная комбинация c неотрицательными коэффициентами $f(X) = \sum_{i=1}^{M} \alpha_j f_j(X)$ будет выпуклой (вогнутой) функцией на D.

Лемма 5.4. Пусть O - выпуклое множество, D - произвольное множество.

 $\Pi ycmb\ g(X,Y):O\in X\times D\in Y.\ \Pi ycmb\ g$ выпукла по X на O при $\forall\ Y$ и ограничена сверху по Y при $\forall\ X.$

Тогда $f(X) = \sup_{Y \in D} g(X, Y)$ выпукла на O.

Лемма 5.5. Если функции $g_1(X), g_2(X), ..., g_M(X)$ выпуклы на выпуклом множестве $O \subset \mathbb{R}^N$ и $G(X) = (g_1(X), g_2(X), ..., g_M(X))$ - вектор-функция, образованная из них, q - монотонно неубывающая выпуклая функция, заданная на выпуклом множестве $D \subset \mathbb{R}^M$, и функция G(X) принимает значения из D, то функция f(X) = q(G(X)) выпукла на O.

Лемма 5.6. Если функция g выпукла на выпуклом множестве $O \subset \mathbb{R}^M$, A - матрица размера $M \times N$, $B \in \mathbb{R}^M$ - вектор и множество $D = \left\{ X \in \mathbb{R}^N : A \cdot X + B \in O \right\}$ непусто, то функция $f(X) = g(A \cdot X + B)$ выпукла на D.

Лемма 5.7 (Дифференциальный критерий выпуклости). Дважды непрерывно дифференцируемая функция f(X) выпукла (вогнута), если ее матрица Гессе является положительно полуопределенной (отрицательно полуопределенной). Если $H_f(Y)$ положительно (отрицательно) определена, то f(X) строго выпукла (вогнута).

Выпуклая задача оптимизации: (*)

$$f(Y)=extr_D\ f(X)$$

$$D=\left\{X\ \middle| X\in P, \psi_j(X)\leqslant (\geqslant,=)0, j=1,...,M\right\}\subseteq \mathbb{R}^N$$
 D - выпуклое множество, $f(x)$ - выпукла на D

Теорема 5.1 (Условие выпуклости множества допустимых решений). Если $\psi(X)$ выпуклая (вогнутая) функция, то множество допустимых решений удовлетворяющее системе $\psi(X) \leq b, x_i \geq 0$ ($\psi(X) \geq b, x_i \geq 0$) будет выпуклым.

Теорема 5.2 (Необходимо условие экстремума). Если в задаче (*) целевая функция задана на выпуклой области определения и дифференцируема в $Y \in D$ и если Y - точка локального минимума (максимума), то grad $f(Y) \cdot \delta X \geqslant (\leqslant) 0$. ($\delta X = X - Y$)

Лемма 5.8. Если точка локального экстремума Y является внутренней точкой D, то $\operatorname{grad} f(Y) = 0$.

Лемма 5.9. Пусть $D = \{ X \mid X \in \mathbb{R}^N, a_j \leqslant x_j \leqslant b_j, j = 1, ..., N \}, -\infty \leqslant a_j \leqslant b_j \leqslant \infty.$ Тогда

$$\frac{\partial f(Y)}{\partial x_j} = \begin{cases}
\geqslant (\leqslant)0, & y_j = a_j \neq -\infty \\
0, & a_j \leqslant y_j \leqslant b_j \\
\leqslant (\geqslant)0, & y_j = b_j \neq +\infty
\end{cases}$$
(5)

Лемма 5.10. Пусть $D = \left\{ X \mid X \in \mathbb{R}^N, x_j \geqslant 0, j = 1, ..., S \right\}.$

Тогда в точке локального минимума (максимума)

$$npu\ j=1,...,S: \frac{\partial f(Y)}{\partial x_j}\geqslant (\leqslant)0\ ecnu\ y_j=0\ u\ \frac{\partial f(Y)}{\partial x_j}=0\ ecnu\ y_j>0;$$
 $npu\ j=S+1,...,S+N: \frac{\partial f(Y)}{\partial x_i}=0\ \forall\ y_j.$

Теорема 5.3 (Достаточное условие экстремума). Если в задаче (*) целевая функция задана на выпуклой области определения и дифференцируема в $Y \in D$ и если $\operatorname{grad} f(Y) \cdot \delta X \geqslant (\leqslant) 0$, то Y точка min (max).

Лемма 5.11. Пусть f(X) - выпуклая (вогнутая) функция, определенная на $D \subseteq \mathbb{R}^N$ и дифференцируемая во внутренней точке $Y \in D$. Если Y - стационарная точка функции f(X), $grad\ f(Y) = 0$, то Y - точка экстремума f(X) на D.

Теорема 5.4 (Единственность точки экстремума задачи выпуклой оптимизации). Если выпуклая функция f(X) определенная на D имеет точку локального минимума (максимума), то эта точка является точкой глобального минимума (максимума).

Теорема 5.5. Пусть f(X) выпуклая функция определенная на D. Пусть f(X) достигает глобального минимума (максимума) на E.

Тогда E выпуклое множество. (E - множество точек глобального минимума (максимума) функции f(X))

Общая (неклассическая) постановка задачи оптимизации: (**)

$$f(Y) = extr_D f(X)$$

$$D = \left\{ X \middle| X \in P; \psi_j(X) \le 0, j = 1, ..., K; \psi_j(X) = 0, j = K + 1, K + 2, ..., M \right\} \subseteq \mathbb{R}^N$$
(6)

Функция Лагранжа:
$$L(\Lambda, \lambda_0, X) = \lambda_0 f(X) + \sum_{j=1}^M \lambda_j \psi_j(X)$$

Теорема 5.6 (Необходимое условие экстремума в форме принципа Лагранжа). *Пусть есть задача* (**). *Пусть выполняются следующие условия:*

- 1. Р выпуклое множество;
- 2. $f(X), \psi_1(X), ..., \psi_K(X)$ дифференцируемы в точке $Y \in D$;
- 3. $\psi_{K+1}(X),...,\psi_{K+M}(X)$ непрерывно дифференцируемы в окрестности $U_{\varepsilon}(Y)$.

Если Y - точка локального минимума (максимума) задачи (**) и при этом $\exists \Lambda' = (\lambda'_1,...,\lambda'_M) \neq 0 \land \lambda'_0 \neq 0 \mid \forall X \in P \ u \ \forall \delta X = X - Y \ выполняются условия Куна-Таккера.$

$$\sum_{i} \frac{\partial L(\Lambda', \lambda'_{0}, Y)}{\partial x_{i}} \delta x_{i} \geqslant (\leqslant) 0 \tag{7}$$

$$\lambda'_{i}\psi_{j}(Y) = 0, j = 1, ..., K \tag{8}$$

$$\lambda_{i}^{'} \geqslant (\leqslant)0, j = 1, \dots, K \tag{9}$$

 $\lambda_{K+1}^{'},...,\lambda_{M}^{'}$ могут иметь любой знак.

Теорема 5.7 (Достаточное условие экстремума). Пусть есть задача (**). Пусть выполняются следующие условия:

- 1. Р выпуклое множество;
- 2. $f(X), \psi_1(X), ..., \psi_K(X)$ дифференцируемы в точке $Y \in D$;
- 3. $f(X), \psi_1(X), ..., \psi_K(X)$ выпуклы на P;
- 4. $\psi_{K+1}(X),...,\psi_{M}(X)$ линейны.

Если существуют такие $\Lambda' \neq 0 \land \lambda_0' \neq 0$, что $\forall X \in P$ выполняются условия Куна-Таккера.

Тогда Y - точка минимума (максимума).

Лемма 5.12. Пусть Y - точка минимума (максимума). Пусть Y - внутренняя точка P.

Тогда
$$\frac{\partial L(\Lambda', \lambda'_0, Y)}{\partial x_i} = 0, i = 1, ..., N.$$

Если $P = \left\{ X \in P \middle| a_i \leqslant x_i \leqslant b_i, i = 1, ..., N \right\}.$

Tог ∂a

$$\frac{\partial L}{\partial x_i} = \begin{cases}
\geqslant (\leqslant)0, & y_i = a_i \neq -\infty \\
0, & a_i < y_j < b_i \\
\leqslant (\geqslant)0, & y_i = b_i \neq +\infty
\end{cases}$$
(10)

Если
$$P = \left\{ X \in P \mid x_i \geqslant 0, i = 1, ..., S, 0 \leqslant S \leqslant N \right\}.$$

Тогда

$$\frac{\partial L(\Lambda', \lambda'_0, Y)}{\partial x_i} \geqslant (\leqslant)0; y_i \frac{\partial L(\Lambda', \lambda'_0, Y)}{\partial x_i} = 0, \qquad i = 1, ..., S$$
(11)

$$\frac{\partial L(\Lambda', \lambda'_0, Y)}{\partial x_i} = 0, \qquad i = S + 1, S + 2, ..., N$$
(12)

Условие регулярности: линейная независимость набора градиентов ограничений в D.

Теорема 5.8 (Необходимое условие экстремума Куна-Таккера). *Пусть есть задача* (**). *Пусть выполняются следующие условия:*

- 1. Р выпуклое множество;
- 2. $f(X), \psi_1(X), ..., \psi_K(X)$ дифференцируемы в точке $Y \in D$;
- 3. $\psi_1(X), ..., \psi_K(X)$ выпуклы на P;
- 4. $\psi_{K+1}(X),...,\psi_{M}(X)$ линейны.

И выполнено одно из условий:

а) ограничения равенства отсутствуют, т.е. K = M и система $\psi_j(X) < 0, j = 1,...,M$ имеет решение на P;

- b) P полиэ ∂p и $\psi_i(X), j = 1, ..., K$ линейны;
- c) P полиэдр $u \psi_{S+1}(X), ..., \psi_K$ линейны u система ограничений $\psi_j(X) \leqslant 0, j = 1, ..., S$ имеет хотя бы одно допустимое решение.

Если Y - точка локального минимума (максимума) Тогда существуют такие $\Lambda' \neq 0 \land \lambda_0' \neq 0$, что $\forall X \in P$ выполняются условия Куна-Таккера.

Теорема 5.9 (Необходимые и достаточные условия экстремума Куна-Таккера в дифф. форме). Пусть есть задача (**). Пусть выполняются следующие условия:

- 1. P выпуклое множество;
- 2. $f(X), \psi_1(X), ..., \psi_K(X)$ дифференцируемы в точке $Y \in D$;
- 3. $\psi_1(X), ..., \psi_K(X)$ выпуклы на P;
- 4. $\psi_{K+1}(X),...,\psi_{M}(X)$ линейны;
- 5. f(x) выпукла.

U выполняется одно из a)-c).

Точка Y локального минимума (максимума) существует $\iff \exists \Lambda' \neq 0 \land \lambda'_0 \neq 0$, такие что $\forall X \in P$ выполняются условия Kуна-Таккера.

Задача: (* * *)

$$f(Y) = \min f(X)$$

$$D = \left\{ X \mid X \in P, \psi_j(X) \le 0, j = 1, ..., K; \psi_j(X) = 0, j = K + 1, ..., M \right\} \subseteq \mathbb{R}^N$$

Теорема 5.10. Пусть есть задача (* * *). Пусть выполняются следующие условия:

- 1. Р выпуклое множество;
- 2. $f(X), \psi_1(X), ..., \psi_K(X)$ выпуклы на P;
- 3. $\psi_{K+1}(X),...,\psi_{M}(X)$ линейны;
- 4. D непусто.

Тогда существуют такие $\Lambda^{'} \neq 0 \wedge \lambda_{0}^{'} \neq 0$, что $\forall X \in P$ выполняются неравенства

$$\lambda_{0}'f^{*} \leqslant \lambda_{0}'f(X) + \sum_{j=1}^{M} \lambda_{j}\psi_{j}(X) = L(\Lambda', \lambda_{0}', X)$$
 (13)

$$\lambda_{j}^{'} \geqslant 0, j = 1, ..., K \tag{14}$$

Теорема 5.11 (Достаточное условие существования вектора Куна-Таккера). Пусть есть задача (***). Пусть выполняются следующие условия:

- 1. Р выпуклое множество;
- 2. $f(X), \psi_1(X), ..., \psi_K(X)$ выпуклы на P;
- 3. $\psi_{K+1}(X),...,\psi_{M}(X)$ линейны.

M выполняется одно из a)-c). Тогда вектор Kуна-Таккера существует.

Условие регулярности задачи (* * *): существование вектора Куна-Таккера.

Теорема 5.12 (Выпуклость двойственной задачи). В двойственной задаче Q - выпукло $u \varphi$ вогнута (выпукла вверх) на Q.

Теорема 5.13 (Основное неравенство двойственности для задачи выпуклой оптимизации). $\forall X \in D$ прямой задачи $u \ \forall \Lambda \in \mathcal{L}$ двойственной задачи справедливо неравенство $f(X) \geqslant \varphi(\Lambda)$.

Теорема 5.14 (Теорема двойственности). Если прямая задача имеет решение и оно конечно и выполнено условие регулярности (th 11).

Тогда множество решений двойственной задачи непусто и совпадает со множеством векторов Куна-Таккера прямой задачи. И целевые функции прямой и двойственной задач совпадают.

Теорема 5.15 (Связь между решением прямой и двойственной задачи). Если для прямой задачи (* * *) выполнено условие регулярности и допустимое множество двойственной задачи непусто, то двойственная задача имеет решение. Если допустимое множество пусто, то минимум прямой задачи это $(-\infty)$.

Теорема 5.16 (Теорема Куна-Таккера в форме двойственности). Если выполнено условие теоремы 11 для прямой задачи, то точка Y есть решение прямой задачи \iff существует вектор Kуна-Таккера, такой что $f(Y) = \varphi(\Lambda)$.

Теорема 5.17 (Теорема Куна-Таккера в терминах седловой точки). Если выполнено условие теоремы 11 для прямой задачи, то точка Y есть решение прямой задач \iff существует вектор Куна-Таккера, такой что (Y,Λ') - седловая точка функции Лагранжа.

Теорема 5.18 (Об условиях одновременного достижения экстремума прямой и двойственной задачи). Если выполнено условие теоремы 11, то точки Y и Λ' есть решения прямой и двойственной задач \iff выполнено соотношение двойственности.

Или точки Y и Λ' есть решения прямой и двойственной задач $\iff (Y,\Lambda')$ - седловая точка функции Лагранжа.

6 Задача выпуклой квадратичной оптимизации

Теорема 6.1. Для того, чтобы существовал вектор Куна-Таккера $\Lambda' \in \mathcal{L}$, удовлетворяющий условиям $\begin{cases} CX + S + \Lambda' = 0 \\ \lambda_j'(A_jY - b_j) = 0, j = 1, ..., K \end{cases} \iff Y \in D$ - точкой минимума.