Методы оптимизации. Теоремы.

Кирилл Захаров

2021 г.

Содержание

1 Линейное программирование **1**
1.1 Базисное решение, допустимое множество, оптимальное решение 1
1.2 Двойственная задача 1

2 Общая постановка задачи оптимизации **2**
2.1 Задача безусловной оптимизации 2
2.2 Задача условной оптимизации 3
2.3 Задача выпуклой оптимизации 4

1 Линейное программирование

1.1 Базисное решение, допустимое множество, оптимальное решение

Теорема 1.1.1. *Множество допустимых решений есть выпуклое множество.*

Лемма 1.1.1. *Базисные решения являются вершинами выпуклой многогранной области.*

Теорема 1.1.2. *Оптимальное решение является базисным решением. (Оптимальное решение лежит в углах выпуклой многогранной области).*

1.2 Двойственная задача

Теорема 1.2.1 (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.*

$$\begin{aligned} f(X) \to \min \Rightarrow f(X) \geqslant \varphi(\Lambda) \\ f(X) \to \max \Rightarrow f(X) \leqslant \varphi(\Lambda) \end{aligned} \tag{1}$$

Теорема 1.2.2 (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 1.2.3.

Теорема 1.2.4. *Если прямая задача имеет оптимальное решение, то и двойственная имеет оптимальное решение.*

Теорема 1.2.5. *Если прямая задача не имеет решения из-за неограниченности целевой функции, то система ограничений двойственной задачи противоречива.*

Теорема 1.2.6 (О дополняющей нежесткости). *Необходимым и достаточным условием того, что прямая и двойственная задачи имеют оптимальное решение, является выполнение условий дополняющей нежесткости.*

$$\lambda_j\left(\sum_{i=1}^{N} a_{ji}x_i - b_j\right) = 0 \tag{2}$$

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

2 Общая постановка задачи оптимизации

2.1 Задача безусловной оптимизации

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

Определение 2.1. $Y = (y_1, ..., y_N)$ - точка локального минимума или максимума, если $\exists\ \varepsilon > 0$, такое что выполняется

$$f(Y) \leqslant f(Y + \delta X) \text{ или } f(Y) \geqslant f(Y + \delta X) \tag{3}$$

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

Определение 2.2. Y - точка строгого экстремума, если неравенства выполняются строго.

Определение 2.3. Y называется точкой глобального экстремума, если неравенства (3) выполняются во всей области.

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

Определение 2.4. Функция, имеющая единственный экстремум называется унимодальной.

Лемма 2.1.1. *Если область допустимых значений, определяемая системой ограничений равенств, содержит точку Y и ее окрестность, то $M < N$.*

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

Лемма 2.1.2. *Пусть область допустимых значений, определяемая системой ограничений равенств задачи на условный экстремум, содержит хотя бы одну точку Y. Если набор градиентов $grad\,\psi_j$ линейно независим и $rank\, J = M < N$, то D вместе с каждой точкой X содержит некоторую непустую ее окрестность.*

Теорема 2.1.1. *Пусть задана функция $f(x)$ и $x \in O = \mathbb{R}^1$. Если в точке Y функция $f(x)$ имеет локальный экстремум, то $\dfrac{\partial f(Y)}{\partial x} = 0$.*

Теорема 2.1.2 (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 2.1.3 (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 2.1.4 (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 2.1.5 (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 2.1.6.

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

2.2 Задача условной оптимизации

Теорема 2.2.1 (Связь между $W_{+/-}(Y, f)$ и $V(Y, f)$)**.** *Если точка Y точка локального минимума (максимума), то $W_-(Y, f) \cap V(Y, f) = \varnothing$ ($W_+(Y, f) \cap V(Y, f) = \varnothing$).*

Теорема 2.2.2 (Вейерштрасс)**.** *Пусть D - компакт и $f(X)$ непрерывная функция определенная на D.*
Тогда существует точка Y глобального минимума (максимума).

2.3 Задача выпуклой оптимизации

Лемма 2.3.1. *Пересечение конечного или счетного числа выпуклых множеств есть выпуклое множество.*

Лемма 2.3.2. *Линейная комбинация* $\sum_{i=1}^{N} \alpha_i X_i$ *конечного числа выпуклых множеств* X_i *при любых* α_i *является выпуклым множеством.*

Лемма 2.3.3. *Если* $f_1(X), f_2(X), ..., f_M(X)$ *выпуклы (вогнуты) на выпуклом множестве* D*, то их линейная комбинация с неотрицательными коэффициентами* $f(X) = \sum\limits_{j=1}^{M} \alpha_j f_j(X)$ *будет выпуклой (вогнутой) функцией на* D*.*

Лемма 2.3.4. *Пусть* O *- выпуклое множество,* D *- произвольное множество.*
Пусть $g(X,Y) : O \in X \times D \in Y$*. Пусть* g *выпукла по* X *на* O *при* $\forall\, Y$ *и ограничена сверху по* Y *при* $\forall\, X$*.*
Тогда $f(X) = \sup\limits_{Y \in D} g(X,Y)$ *выпукла на* O*.*

Лемма 2.3.5 (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 2.3.1 (Условие выпуклости множества допустимых решений)**.** *Если* $\psi(X)$ *выпуклая (вогнутая) функция, то множество допустимых решений удовлетворяющее системе* $\psi(X) \leqslant b, x_i \geqslant 0$ $(\psi(X) \geqslant b, x_i \geqslant 0)$ *будет выпуклым.*

Теорема 2.3.2 (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* $grad\, f(Y) \cdot \delta X \geqslant (\leqslant)\, 0$*.* $(\delta X = X - Y)$

Теорема 2.3.3 (Достаточное условие экстремума)**.** *Если в задаче* $(*)$ *целевая функция задана на выпуклой области определения и дифференцируема в* $Y \in D$ *и если* $grad\, f(Y) \cdot \delta X \geqslant (\leqslant)\, 0$*, то* Y *точка* $\min$ (max).

Теорема 2.3.4 (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 2.3.5. *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: $(**)$

$$f(Y) = extr_D\, f(X)$$
$$D = \left\{ X \,\middle|\, X \in P; \psi_j(X) \leqslant 0, j = 1, ..., K; \psi_j(X) = 0, j = K+1, K+2, ..., M \right\} \subseteq \mathbb{R}^N \tag{5}$$

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