ЛЕКЦИЯ №7.
Методы нелинейного программирования

Задача нелинейного программирования

Задача математического программирования в общем формулируется так:
Найти $X \in \mathbb{R}^n$, которое удовлетворяет

$$\underset{X \in G}{extr}\left\{ f(X) \right\},$$
$$G = \left\{ X : f_i(X) \begin{Bmatrix} = \\ \le \end{Bmatrix} 0, \quad i = \begin{Bmatrix} \overline{1,k} \\ \overline{k+1,m} \end{Bmatrix}, \quad X \ge 0 \right\}. \tag{7.1}$$

Если хотя бы одна из функций $f_i\left(X\right)$, $i = \overline{1,m}$ нелинейная, то задача (7.1) называется задачей нелинейного программирования (ЗНП).

Строгие равенства $f_i = 0$ можно использовать для сокращения вектора *X*. Следовательно, ЗНП можно сформулировать так. Найти *X*, которое удовлетворяет

$$\underset{X \in G}{extr}\left\{ f(X) \right\},$$
$$G = \left\{ X : f_i(X) \le 0, \quad i = \overline{1,m}, \quad X \ge 0 \right\}. \tag{7.2}$$

Напротив, продлевая вектор *X* балансовыми переменными, можно исключить неравенства из ЗНП (7.1). Тогда ЗНП формулируется так:

$$\underset{X \in G}{extr}\left\{ f(X) \right\},$$
$$G = \left\{ X : f_i(X) = 0, \quad i = \overline{1,m}, \quad X \ge 0 \right\}. \tag{7.3}$$

Задачи (7.1)-(7.3) предусматривают отыскание глобального экстремума, то есть такого x_0, для которого $f\left(X_0\right) \begin{Bmatrix} \le \\ \ge \end{Bmatrix} f\left(X\right), \quad \forall X \in G$. Из курса математического анализа известно, что непрерывная функция $f\left(X\right)$, заданая в замкнутой области *G*, достигает в ней глобальных экстремумов.

Рассмотрим классификацию задач НЛП.

Классические задачи нелинейной оптимизации

Эти задачи предполагают нахождении экстремума нелинейной функции $f\left(X\right)$, $X \in \mathbb{R}^n$ при нелинейных ограничениях-ровностей

$$f_i\left(X\right)=0,\quad i=\overline{1,m},\quad m\le n.$$

Их еще называют *задачами отыскания условного экстремума*.

Если $m=0$, то имеем классическую задачу отыскания безусловного экстремума функции $f\left(X\right)$, $X\in\mathbb{R}^n$.

Для классических задач оптимизации существенным является требование гладкости (существование и непрерывность у функций $f_i\left(X\right)$, $i=\overline{1,m}$ производных частей по крайней мере ко 2-ом порядке включительно).

Классические задачи оптимизации могут быть решены классическими методами с использованием аппарата дифференциального исчисления. Однако трудности вычислительного характера, которые возникают при этом настолько значительные, что для решения практических задач этого типа необходимо применять другие методы.

Замечание. В зависимости от количества экстремумов на интервале поиска целевая функция может быть *унимодальной* (имеет один экстремум) или полимодальной (имеет несколько экстремумов).

На практике большинство функций – полимодальные. Поиск глобального экстремума полимодальной функции крайне сложная задача, при решении которой возникает проблема слойности: сходимость метода оптимизации ухудшается при приближении к точному решению оптимизационной задачи из-за того, что он блуждает по локальным минимумам.

Задачи квадратичного программирования

В этих задачах нужно минимизировать квадратичную функцию при линейных ограничениях и при условии, что $f\left(X\right)$ является или выпуклой, или вогнутой функцией.

Задачи квадратичного программирования (ЗКП) можно отнести к классу задач выпуклого программирования. Но их выделяют в отдельный класс как через специфику целевой функции, так и через специфику условий-ограничений.

Задачи выпуклого (вогнутого) программирования

ЗНП (7.2), в которых целевая функция $f\left(X\right)$ является выпуклой вниз (вверх), а допустимое множество − выпуклым, относят к классу задач выпуклого программирования (ЗВП). Методы решения этих задач являются наиболее разработанными в нелинейном программировании.

Примером таких методов может быть группа градиентных методов.

Теорема Куна-Таккера

Рассмотрим задачу математического программирования.

Найти $X=\left(x_1,x_2,...,x_n\right)^T$, которая обеспечивает

$$z = \max_{X} \left\{ f\left(X\right) \right\},$$
$$G = \left\{ X : f_i(X) \geq 0, \quad X \geq 0, \quad i = \overline{1,m} \right\}, \tag{7.4}$$

и для которой складываем функцию Лагранжа $\Lambda = (\lambda_1, ..., \lambda_m)^T$:

$$L(X,\Lambda) = f(X) + \sum_{i=1}^{m} \lambda_i f_i(X). \tag{7.5}$$

Если выполняется условие регулярности $\exists X \in G: f_i(X) > O, \ \forall i = \overline{1,m}$, то имеет место такая теорема.

Теорема. Точка $X_0 \geq 0$ тогда и только тогда есть решением задачи (7.4), если существует вектор $\Lambda_0 \geq 0$, такой, при котором для всех $X \in G$ и $\Lambda \geq 0$ выполняется условие:

$$L(X,\Lambda_0) \leq L(X_0,\Lambda_0) \leq L(X_0,\Lambda). \tag{7.6}$$

Точка $\left(X_0,\Lambda_0\right)$ называется точкой седла функции $L(X,\Lambda)$.

Если $f\left(X\right)$ и $f_i\left(X\right)$ дифференцированные функции, то условия (7.6) эквивалентны так называемым локальным условиям Куна-Таккера :

$$\left.\left(\frac{dF}{dx_j}\right)\right|_{(X_0,\Lambda_0)} \leq 0, \quad x_j^{(0)} \cdot \left.\left(\frac{dF}{dx_j}\right)\right|_{(X_0,\Lambda_0)} = 0, \quad x_j^{(0)} \geq 0, \quad j = \overline{1,n}, \tag{7.7}$$

$$\left.\left(\frac{dF}{d\lambda_i}\right)\right|_{(X_0,\Lambda_0)} \geq 0, \quad \lambda_i^{(0)} \cdot \left.\left(\frac{dF}{d\lambda_i}\right)\right|_{(X_0,\Lambda_0)} = 0, \quad \lambda_i^{(0)} \geq 0, \quad i = \overline{1,m}. \tag{7.8}$$

Квадратичное программирование

Нелинейная задача с линейными ограничениями и квадратичной целевой функцией называется задачей квадратичного программирования:

$$f(x) = \max_{x \in X} \left\{ \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij} x_i x_j + \sum_{j=1}^{n} c_j x_j \right\},$$
$$\sum_{j=1}^{n} a_{ij} x_j = b_i, \quad i = \overline{1,m}, \tag{7.7}$$
$$x_j \geq 0, \quad j = \overline{1,n},$$

где $X = (x_1, ..., x_n)^T$ и $d_{ij} = d_{ji}$, $\left(i,j\right) = \overline{1,n}$.

Обозначим $A=\left(a_{ij}\right)_{mn}$, $B=(b_1,...,b_m)^T$, $C=(c_1,...,c_n)$, $D=(d_{ij})_n$. Тогда задача (7.7) примет вид

$$\begin{gathered} f(X)=\max_X\left\{X^TDX+CX\right\}, \\ AX=B, \\ X\geq 0. \end{gathered} \tag{7.8}$$

Запишем функцию Лагранжа

$$L(X,\Lambda)=X^TDX+CX+\Lambda^T(B-AX),$$

где $\Lambda=(\lambda_1,...,\lambda_m)$, и проверим условия теоремы Куна-Таккера. Они выполняются, если квадратичная форма X^TDX определенна отрицательно. Из ограничений задачи (7.8) и локальных условий Куна-Таккера получаем систему соотношений

$$\begin{cases} B-AX=0, \\ X\,2D+C-\Lambda A\leq 0. \end{cases} \tag{7.9}$$

Решение $\left(X_0,\Lambda_0\right)$ системы (7.9) содержит решение X_0 задачи (7.8) и решений Λ_0 двойной задачи. При решении системы (7.9) следует учитывать одно из условий дополняющей нежёсткости $\left(X2D+C-\Lambda A\right)X=0$, нелинейность которой создает значительные трудности. Однако существуют алгоритмы, которые успешно преодолевают эти трудности (например, алгоритм Вольфа).

Графический метод решения ЗНП

При $n=2$ область G задачи (7.2) в плоскости x_1Ox_2 может быть построена как сечение областей $f_i\left(x_1,x_2\right)\leq 0$, $i=\overline{1,m}$, а линия $f_i\left(x_1,x_2\right)=z_0$ при изменении параметра z_0 – это линия уровня, которая “входит” в область G в точке минимума (находится ли в области, оставаясь точкой), а “выходит” из области в точке максимума функции $f\left(x_1,x_2\right)$.

Контрольный пример 7.1

Задачу квадратичного программирования решить графическим методом:

$$f(X)=\max_{x\in X}\left\{-2x_1^2-x_2^2+20x_1+16x_2\right\},$$

$$\begin{cases} x_1+2x_2\leq 16, \\ 5x_1+2x_2\leq 40, \\ x_j\geq 0,\ j=\overline{1,2}. \end{cases}$$

Решение

Построим область ограничений *G* для данной задачи квадратичного программирования (рис 7.1).

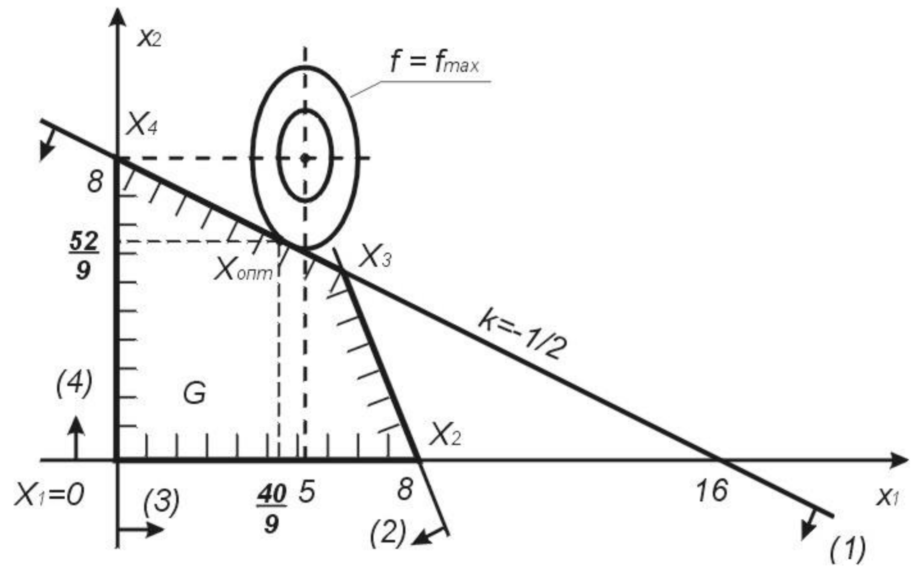

Рисунок 7.1 – Графический метод решения задачи квадратичного программирования

Выпишем уравнение линии уровня

$$-2x_1^2 - x_2^2 + 20x_1 + 16x_2 = f \in \text{const},$$

и упростим его, выделив полные квадраты. Получим:

$$2\left(x_1 - 5\right)^2 + \left(x_2 - 8\right)^2 = 114 - f = c^2,$$

или

$$\frac{(x_1-5)^2}{\left(c/\sqrt{2}\right)^2} + \frac{(x_2-8)^2}{(c)^2} = 1.$$

Последнее уравнение описывает эллипс с центром в точке (5,8) и полуосями $c/\sqrt{2}$ и *c*.

При росте f полуоси $c/\sqrt{2}$ и *c* уменьшаются (эллипс “взимается”). Точкой максимума функции $f\left(x\right)$ в области *G* будет точка, в которой эллипс, “взимаясь”, покинет область *G*. Это будет точка касания эллипса и прямой $x_1 + 2x_2 = 16$, угловой коэффициент которой $k = -1/2$.

Найдем координаты этой точки из условия, что она принадлежит прямой и $\frac{dx_2}{dx_1}=k=-\frac{1}{2}$ для уравнения линии уровня. Продифференцируем это уравнение по переменной x_1:

$$20+16\frac{dx_2}{dx_1}-4x_1-2x_2\frac{dx_2}{dx_1}=0.$$

$$20+16\left(-\frac{1}{2}\right)-4x_1-2x_2\left(-\frac{1}{2}\right)=0.$$

Окончательно имеем систему уравнений:

$$\begin{cases} x_1+2x_2=16, \\ 4x_1-2x_2=12, \end{cases}$$

решение которой и будет решением данной задачи квадратичного программирования:

$$X_{\text{опт}}=\left(\frac{40}{9},\ \frac{52}{9}\right),$$

$$f_{\max}=108\frac{4}{9}.$$

Проверим выполнение условий теоремы Куна-Таккера, т.е. покажем, что существует $\lambda_0\geq 0$, при котором в точке оптимума выполняются условия теоремы для функции $F(X,\Lambda)$:

$$F(X,\Lambda)=-2x_1^2-x_2^2+20x_1+16x_2+\lambda_1\cdot(16-x_1-2x_2)+\lambda_2\cdot(40-5x_1-2x_2).$$

Находим:

$$\frac{dF}{dx_1}=-4x_1+20-\lambda_1-5\lambda_2,\ \frac{dF}{dx_2}=-2x_2+16-2\lambda_1-2\lambda_2,$$

$$\frac{dF}{d\lambda_1}=16-x_1-2x_2,\ \frac{dF}{d\lambda_2}=40-5x_1-2x_2.$$

λ_2 приобретает нулевое значение, поскольку при подстановке $x_1=\frac{40}{9}$ и $x_2=\frac{52}{9}$ в выражение $\frac{dF}{d\lambda_2}$, имеем значение больше нуля, а за условием (7.8) $\lambda_2^0\cdot\left(\frac{dF}{d\lambda_2}\right)\Bigg|_{(X_0,\Lambda_0)}=0$.

В соответствии с условиями (7.7) производные $\left.\left(\dfrac{dF}{dx_j}\right)\right|_{(X_0,\Lambda_0)}$, $j=\overline{1,2}$ приобретают нулевые значения, поскольку координаты вектора X_0 отличаются от нуля.

Находим $\lambda_1=\dfrac{20}{9}$. Следовательно, в точке $(X_0,\Lambda_0)=\left(\dfrac{40}{9},\dfrac{52}{9},\dfrac{20}{9},0\right)$ выполняются условия Куна-Таккера, и она действительно является точкой экстремума.

Контрольный пример 7.2

Задачу квадратичного программирования на минимум решить аналитическим и графическим методом:

$$f(X)=\max_{x\in X}\left\{5x_1-5x_2-2x_1^2-2x_2^2\right\},$$

$$\begin{cases}2x_1+3x_2\le 10,\\ x_1-x_2\le 13,\\ x_j\ge 0,\quad j=\overline{1,2}.\end{cases}$$

Решение

Графический метод.

Построим область ограничений *G* для данной задачи квадратичного программирования (рис 7.2).

Выпишем уравнение линии уровня:

$$-2x_1^{\ 2}-2x_2^{\ 2}+5x_1-5x_2=f\in\text{const},$$

или

$$(x_1-\frac{5}{4})^2+(x_2+\frac{5}{4})^2=\frac{100-16f}{32}.$$

Последнее уравнение описывает круг с центром в точке $\left(\dfrac{5}{4},-\dfrac{5}{4}\right)$ и радиусом $\dfrac{100-16f}{32}$.

При уменьшении f радиус $\dfrac{100-16f}{32}$ увеличивается (круг “растягивается”). Точкой минимума функции $f(x)$ в области *G* будет точка, в которой круг, “растягиваясь”, коснется области *G*. Это будет точка касания круга и прямой $x_2=0$; угловой коэффициент которой $k=0$. Найдем координаты этой точки из условия, что она принадлежит прямой $x_2=0$ и

$\dfrac{dx_2}{dx_1} = k = 0$ для уравнения линии уровня. Продифференцируем это уравнение по переменной x_1:

$$5 - 5\frac{dx_2}{dx_1} - 4x_1 - 4x_2\frac{dx_2}{dx_1} = 0,$$

$$5 - 5 \cdot 0 - 4x_1 - 2x_2 \cdot 0 = 0.$$

Окончательно имеем систему уравнений:

$$\begin{cases} x_1 = \dfrac{5}{4}, \\ x_2 = 0, \end{cases}$$

решение которой и будет решением данной задачи квадратичного программирования:

$$X_{\text{опт}} = \left(\frac{5}{4},\ 0\right),$$
$$f_{\max} = 3\frac{1}{8}.$$

Аналитический метод

Составим вспомогательную систему уравнений.

$$\begin{cases} 2x_1 + 3x_2 + x_3 = 10, \\ x_1 - x_2 + x_4 = 13, \\ -4x_1 - 2\lambda_1 - \lambda_2 + y_1 = -5, \\ -4x_2 - 3\lambda_1 + \lambda_2 + y_2 = 5, \\ -\lambda_1 + y_3 = 0, \\ -\lambda_2 + y_4 = 0, \\ x_j, y_j \geq 0, \quad x_j \cdot y_j = 0, \quad j = \overline{1,4}. \end{cases}$$

где λ_1 и λ_2 – множители Лагранжа.

Запишем эту систему уравнений в виде:

$$
\begin{cases}
2x_1 + 3x_2 + x_3 = 10, \\
x_1 - x_2 + x_4 = 13, \\
4x_1 - y_1 + 2y_3 + y_4 = 5, \\
-4x_2 + y_2 - 3y_3 + y_4 = 5, \\
x_j, y_j \geq 0, \quad x_j \cdot y_j = 0, \quad j = \overline{1,4}.
\end{cases}
$$

Опорное решение последней системы содержит решение данной задачи. Найдем это решение методом искусственного базиса.

Составим вспомогательную ЗЛП:

$$\varphi = \max\{z_1\},$$

$$
\begin{cases}
2x_1 + 3x_2 + x_3 = 10, \\
x_1 - x_2 + x_4 = 13, \\
4x_1 - y_1 + 2y_3 + y_4 + z_1 = 5, \\
-4x_2 + y_2 - 3y_3 + y_4 = 5, \\
x_j, y_j \geq 0, \quad z_1 \geq 0, \quad x_j \cdot y_j = 0, \quad i = 1, \quad j = \overline{1,4}.
\end{cases}
$$

Решим последнюю задачу симплексным методом (таблица 7.1).

Таблица 7.1

i	$Б$	$C_б$	U_0	0	0	0	0	0	0	0	0	1
				U_1	U_2	U_3	U_4	U_5	U_6	U_7	U_8	U_9
1	U_3	0	10	2	3	1	0	0	0	0	0	0
2	U_4	0	13	1	-1	0	1	0	0	0	0	0
3	U_9	1	5	4	0	0	0	-1	0	2	1	1
4	U_6	0	5	0	-4	0	0	0	1	-3	1	0
m+1	$\varphi_j - C_j$		4	0	0	0	0	-1	0	2	1	0
1	U_3	0	15/2	0	3	1	0	1/2	0	-1	-1/2	-1/2
2	U_4	0	47/4	0	-1	0	1	-1/4	0	-1/2	-1/4	-1/4
3	U_1	0	5/4	1	0	0	0	-1/4	0	1/2	1/4	1/4
4	U_6	0	5	0	4	0	0	0	1	-3	1	0
m+1	$\varphi_j - C_j$		0	0	0	0	0	0	0	0	0	0

Не выписывая полностью решения вспомогательной ЗЛП из последней симплексной таблицы выпишем решение данной задачи квадратичного программирования

$$X_{\text{опт}} = \left(\frac{5}{4},\, 0\right),$$

$$f_{\max} = 3\frac{1}{8}.$$

Контрольные вопросы

1. При каких условиях оптимизационная задача может быть отнесена к классу нелинейных?
2. Перечислите основные трудности, возникающие в процессе решения задачи нелинейного программирования.
3. Какой смысл вкладывается в понятие «условная оптимизация»?
4. Для чего предназначен метод множителей Лагранжа и в чем он состоит?
5. Дайте определение выпуклой (вогнутой) функции.
6. Сформулируйте достаточное условие выпуклости (вогнутости) функции.
7. В чем заключена специфика задач выпуклого программирования?
8. Сформулируйте необходимое и достаточное условия теоремы Куна-Таккера. Какое значение они имеют для решения задач нелинейного программирования?
9. Приведите пример пары двойственных задач нелинейного программирования.
10. Какие свойства пары нелинейных двойственных задач могут быть применены для их решения?