Постановка задач нелинейного программирования.
Выпуклые и вогнутые множества и функции.

Цели:

1. Формирование знаний об основных понятиях и определениях задач нелинейного программирования.
2. Формирование знаний об основных способах решения задач нелинейного программирования.

Задачи:

1. Сформировать теоретические знания необходимые при составлении и решении задач нелинейного программирования.
2. Содействовать расширению профессиональной компетенции в области основных понятий и способов решений задач нелинейного программирования.

В течение последних двух десятилетий из *нелинейного программирования* выделились самостоятельные разделы:

- выпуклое программирование,
- квадратичное программирование,
- *целочисленное программирование,*
- стохастическое программирование,
- динамическое программирование и др.

Задачи выпуклого программирования – это задачи, в которых определяется минимум выпуклой функции (или максимум вогнутой), заданной на выпуклом замкнутом множестве. Эти задачи среди задач *нелинейного программирования* наиболее изучены.

Среди *задач выпуклого программирования* более подробно изучены задачи квадратичного программирования. В этих задачах целевая функция – квадратична, а ограничения – линейны.

В задачах *целочисленного программирования* неизвестные параметры могут принимать только целочисленные значения.

В задачах стохастического программирования в целевой функции или в функциях ограничений содержатся случайные величины, которые подчиняются законам теории вероятностей.

В задачах динамического программирования ограничения содержат как параметр время и при этом описываются дифференциальными уравнениями. Процесс нахождения решений в задачах динамического программирования является многоэтапным.

Классификация методов нелинейного программирования

Для решения задачи *нелинейного программирования* было предложено много методов, которые можно классифицировать по различным признакам.

По количеству локальных критериев в целевой функции методы *нелинейного программирования* делятся на:

- однокритериальные,
- многокритериальные.

По длине вектора $\overline{X}$ методы делятся на:

- однопараметрические или одномерные (n=1),
- многопараметрические или многомерные (n>1).

По наличию ограничений методы *нелинейного программирования* делятся на:

- без ограничений (безусловная оптимизация),
- с ограничениями (*условная оптимизация*).

По типу информации, используемой в алгоритме поиска экстремума методы делятся на:

- методы прямого поиска, т.е. методы, в которых при поиске экстремума целевой функции используются только ее значения;
- *градиентные методы* первого порядка, в которых при поиске *экстремума функции* используются значения ее первых производных;
- *градиентные методы* второго порядка, в которых при поиске *экстремума функции* наряду с первыми производными используются и*вторые производные*.

Ни один метод *нелинейного программирования* не является универсальным. В каждом конкретном случае необходимо приспосабливать применяемый метод к особенностям решаемой задачи.

3. Классический метод определения условного экстремума

Задача *нелинейного программирования* (задача НП) в общем виде формулируется так:

максимизировать $f(x_1, x_2, \ldots, x_n)$

при ограничениях

$g_1(x_1, x_2, \ldots, x_n) \geq 0;$

$g_2(x_1, x_2, \ldots, x_n) \geq 0;$

$\ldots\ldots\ldots\ldots\ldots\ldots\ldots$

$g_m(x_1, x_2, \ldots, x_n) \geq 0;$

где функции $f(x_1, x_2, \ldots, x_n),\ g_i(x_1, x_2, \ldots, x_n) \geq 0,\ i = \overline{1, m}$ нелинейны.

В отличие от задачи ЛП для задач НП нет универсального метода решения.

В задаче ЛП допустимое множество R всегда является выпуклым с конечным числом крайних точек. Поэтому воспользовавшись *симплекс-методом* и перебрав только крайние точки, можно за конечное число шагов найти *оптимальное решение. В задачах* НП, наоборот, выпуклость допустимого множества и конечность числа его крайних точек совсем необязательны. Это и служит причиной основной трудности решения задач НП.

Для определения условного экстремума (то есть экстремума при ограничениях) можно воспользоваться методами дифференциального исчисления, когда функция $f(x_1, x_2, \ldots, x_n)$ имеет не ниже *второй производной*. Рассмотрим некоторые важные понятия и теоремы классического анализа, которые лежат в основе классических методов поиска условного экстремума .

Теорема 1 (теорема существования экстремума). *Если* $f(x_1, x_2, \ldots, x_n)$ - *непрерывная функция, определенная на замкнутом и ограниченном множестве, то она достигает на этом множестве, по крайней мере один раз, своих максимального и минимального значений>.*

Следующая теорема определяет возможные местоположения максимума (или минимума).

Теорема 2. *Если* $f(x_1, x_2, \ldots, x_n)$ *является непрерывной функцией нескольких переменных, определенной на допустимом множестве* R, *то максимальное значение* $f(x_1, x_2, \ldots, x_n)$, *если оно существует, достигается в одной или нескольких точках, которые принадлежат одному из следующих множеств*: 1) S_1 - *множество стационарных точек* ; 2) S_2 - *множество точек границы* ; 3)S_3 - *множество точек, где функция* $f(x_1, x_2, \ldots, x_n)$ *не дифференцируема*.

Определение 1. Множество точек $S_1(x_1, x_2, \ldots, x_n)$ функции f(x) называется ***множеством стационарных точек***, если они удовлетворяют условию

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

Определение 3.2. Функция f(x) достигает *локального максимума* в точке $x^0 = \left(x_1^0, x_2^0, \ldots, x_n^0\right)$, если для всех точек x, лежащих в малой окрестности точки $\left[x_1^0, x_2^0, \ldots, x_n^0\right]$ имеет место неравенство

$$f\left(x_1^0, x_2^0, \ldots, x_n^0\right) \geq f\left(x_1, x_2, \right. \quad (2)$$

Определение 3.3. Функция f(x) достигает *глобального (абсолютного) максимума* в точке x^0, если для всех точек $x \in R$ справедливо неравенство

$$f(x^0) \geq f(x)$$

Для нахождения стационарных точек функции f(x) можно использовать следующую теорему.

Теорема 3.3. *Пусть* $f(x_1, x_2, \ldots, x_n)$ *дифференцируема в некоторой допустимой области* R. *Если в некоторой внутренней точке* $\left(x_1^0, x_2^0, \ldots, x_n^0\right)$ *области* R *функция* f(x) *достигает относительного максимума, то*

$$\frac{\partial f(x_0)}{\partial x_j} = 0 \quad 3)$$

Для того чтобы определить, являются ли найденные стационарные точки точками максимума или минимума, необходимо исследовать функцию $f(x_1, x_2, \ldots, x_n)$ в окрестности стационарных точек и определить, является она выпуклой или вогнутой.

Определение 4. Пусть R - *выпуклое множество* точек n - мерного пространства. Функция f, определенная на R, называется выпуклой вверх, если для любой пары точек $x_1, x_2 \in R$ и произвольного $0 \leq k \leq 1$ выполняется неравенство

$f[kx_1 + (1-k)x_2] \geq kf(x_1) + (1 -$ 3.4)

Если

$f[kx_1 + (1-k)x_2] \leq kf(x_1) + (1 -$ 5)

то функция называется вогнутой.

Если (4) или (5) выполняются как строгие неравенства, то функция называется строго вогнутой или строго выпуклой соответственно.

Критерий выпуклости и вогнутости функции n - переменных можно сформулировать в виде следующей теоремы.

Теорема 4. *Дифференцируемая функция* f(x) *строго вогнутая в некоторой окрестности точки* $x^0\left(x_1^0, x_2^0, \ldots, x_n^0\right)$, *если выполняются следующие условия*:

$$f_{11}(x_0) < 0; \quad \begin{vmatrix} f_{11}(x_0) & f_{12}(x_0) \\ f_{21}(x_0) & f_{22}(x_0) \end{vmatrix} > 0; \quad \begin{vmatrix} f_{11}(x_0) & f_{12}(x_0) & f_{1} \\ f_{21}(x_0) & f_{22}(x_0) & f_{2} \\ f_{31}(x_0) & f_{32}(x_0) & f_{3} \end{vmatrix} \quad 6)$$

0

И так далее, то есть если знаки определителей чередуются начиная с < 0, где

$$f_{ij}(x_0) = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}\bigg| x = x_0$$

Функция f(x) *строго выпукла в окрестности точки* x_0, *если все определители (выписанные выше) положительные*.

Имеет место следующая теорема.

Теорема 5. *Для того чтобы в точке* x_0 *достигался внутренний относительный минимум, достаточно, чтобы эта точка была стационарной, а самая функция в окрестности точки* x_0 *была строго выпуклой*.

Справедливо следующее **утверждение**: если f(x) строго выпуклая (вогнутая) функция на всем множестве решений R, то f имеет только один относительный минимум (максимум), который является и абсолютным.

Теорема 6 (о выпуклости допустимого множества решений).

Пусть $g_1(x), g_2(x), \ldots, g_m(x), \geq 0$ *и* $x \geq 0$ *-ограничения задачи нелинейного программирования*. *Если функции* $g_1(x), g_2(x), \ldots, g_m(x)$ *- вогнуты, то допустимое*

$$R(x) = \{x: \; g_1(x) \geq 0, g_2(x) \geq 0, \ldots, g_m(x) \geq 0, \; x \geq 0\}$$

множество является выпуклым.

Доказательство. Для доказательства теоремы достаточно показать, что множество $R(x) = \{x: \; g_1(x) \geq 0, \; x \geq 0\}$ при каждом $i = \overline{1, m}$ будет выпуклым. Тогда множество $R = R_1 \cap R_2 \cap \ldots \cap R_m$ также выпукло, так как пересечение конечного числа выпуклых множеств R_i.

Рассмотрим некоторую вогнутую функцию $g_i(x) \geq 0$. Выберем две произвольных точки $x_1 \geq 0$ и $x_2 \geq 0$ (рис.7.1). Тогда $x_2 = \lambda x_1 + (1-\lambda)x_3 \geq 0, \; 0 < \lambda < 1$. Поскольку $x_1 \in R_i, x_3 \in R_i$, то и точка x_2 принадлежит R_i. Из условия вогнутости g_i следует, что $g_i[\lambda x + 1 + (1-\lambda)] \geq g_i(x_1)\lambda + (1-\lambda)g_i(x_1) \geq 0$.

Следовательно, множество R_i содержит отрезок $\lambda g_i(x_1) + (1-\lambda)g_i(x_1)$, и поэтому оно выпукло (рис.7.1).

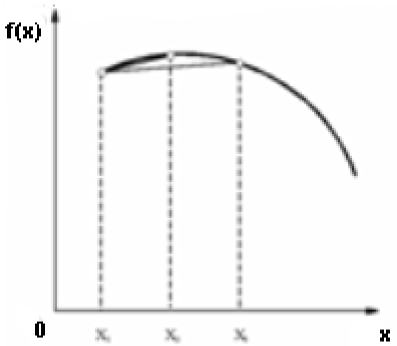

Рис. 1.

Справедливое такое **утверждение**: если функции $f_1(x), f_2(x), \ldots, f_p(x)$ - выпуклы (вогнуты) на множестве R_i, то функция $g(x) = \sum_{i=1}^{p} k_i f_i(x)$ - также выпукла (вогнута) при условии, что все $k_i \geq 0,\ i = 1, 2, \ldots, p$.

Рассмотрим метод поиска условного экстремума. Он состоит из следующих процедур.

1.Отыскивают *множество всех стационарных точек* $S_1(x)$ функции f(x) на выпуклом допустимом множестве R. Найденные точки далее исследуют на максимум (минимум) и определяют точку наибольшего максимума $x_0(x_0 \in S_1(x))$.

2. Переходят к исследованию точек границы $S_2(x)$ и отысканию тех из них, где f(x) достигает максимума. Этот процесс состоит в следующем. Выбирают произвольную границу, определяемую, например, условием $g_1(x)=0$. Если функция

$g_i(x) = g_i(x_1, x_2, \ldots,$ 3.7)

является сепарабельной, то можно, определив из (7) переменную

$x_i = \varphi(\{x_j\}),\ j = \overline{1, n},\ j \neq i$

подставить ее в выражение для f(x). Тем самым задача сведется к поиску безусловного экстремума, для чего можно использовать процедуру, описанную в п.1.

Обозначим через x_i^+ точку границы $g_i(x) = 0,\ x_i^+ \in R$, в которой f(x) достигает максимума. Повторив вышеописанную процедуру по всем остальным границам, найдем соответственно точки максимума (минимума) для всех границ $x_k^+, k = \overline{1, m}$.

3. Непосредственным сравнением значений функции f(x) для всех точек $x_0^+, x_1^+, \ldots, x_m^+$ определяют точку *абсолютного максимума* (минимума) x_{opt} на множестве решений R.

Такой подход требует значительных вычислительных затрат и может применяться лишь в простейших случаях при небольшом числе ограничений m и для случая сепарабельных функций $g_1(x)$, поэтому область его применения очень ограничена, и ниже рассматриваются более эффективные методы решения задач *условной оптимизации*.

Обобщение понятия выпуклой функции. Рассмотрим некоторые классы функций, которые не являются полностью выпуклыми, но обладают лишь отдельными их свойствами.

Определение 5. *Пусть функция* f(x) *определена на непустом и выпуклом множестве* R. *Функция* f(x) *квазивыпукла, если для любых* $x_1, x_2 \in R$ *и* $\lambda \in [0, 1]$ *выполняется неравенство*

$f(\lambda x_1 + (1 - \lambda) x_2) \leq \max\{f(x_1),$ 8)

Функция f(x) называется *квазивогнутой*, если -f(x) - *квазивыпуклая функция*.

Из этого определения следует, что функция f(x) - квазивыпукла, если из неравенства $f(x_2) \geq f(x_1)$ следует, что $f(x_2)$ не меньше значения функции f(x) в любой точке, являющейся выпуклой комбинацией точек x_1 и x_2. И наоборот, функция f(x)квазивогнута, если из неравенства $f(x_2) \geq f(x_1)$ следует, что $f(x_1)$ не больше значения f(x) в любой точке, которая есть выпуклой комбинацией точек x_1 и x_2.

На рис. 7.2 приведены примеры квазивыпуклых и *квазивогнутых функций*, где а - квазивыпуклая, б - *квазивогнутая функции*.

Введем понятия строгой квазивыпуклости и квазивогнутости.

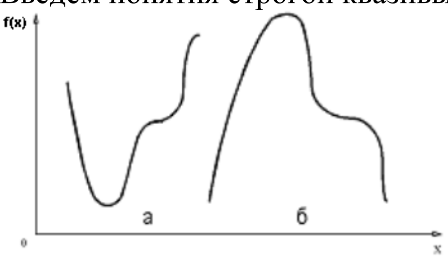

Рис. 2.

Определение 6. *Пусть функция* f(x) *определена на непустом и выпуклом множестве* R. *Функция* f(x) *строго квазивыпукла, если для любых* $x_1, x_2 \in R$ *таких, что* $f(x_1) \neq f(x_2)$ *и* $\lambda \in (0;1)$ *выполняется неравенство*

$$f(\lambda x_1 + (1-\lambda)x_2) < \max\{f(x_1), j$$ 9)

Функция f(x) называется строго квазивогнутой, если -f(x) - строго *квазивыпуклая функция*. На рис. 7.3 изображены: а, б - строго*квазивыпуклые функции*, в - *квазивогнутая функция*. Из приведенного определения следует, что любая выпуклая функция является в тоже время и строго квазивыпуклой.

Строго квазивыпуклые и *квазивогнутые функции* играют важную роль в *нелинейном программировании*, поскольку для них *локальный минимум* и *локальный максимум* являются глобальным минимумом и максимумом соответственно.

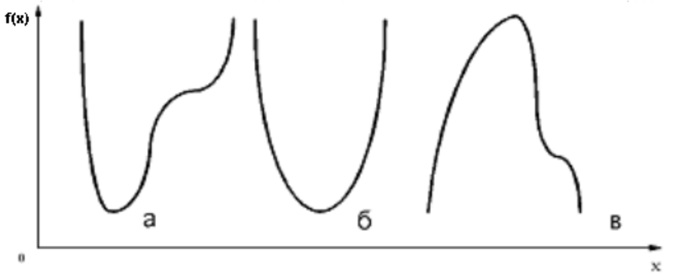

Рис. 3.

Утверждение. Пусть f(x) - строго *квазивыпуклая функция*. Рассмотрим задачу минимизаци f(x) при условии, что $x \in R$, где R - непустое *выпуклое множество* в $E^{(n)}$. Пусть $\overline{x}$ - точка *локального минимума* рассматриваемой задачи. Тогда она является и точкой глобального минимума.

Доказательство. Предположим противное, то есть пусть существует точка $x^+ \in R$, для которой $f(x^+) < f(\overline{x})$. ПосколькуR - выпуклое, то точка $\lambda x^+ + (1-\lambda)\overline{x} \in R$ при любой $\lambda \in (0;1)$. Так как $\overline{x}$ - точка локального минимума, то

$$f(\overline{x}) \leq f[\lambda x^+ + ($$ 10)

для всех $\lambda \in (0, \delta)$ для некоторого $\delta \in (0,1)$.

Поскольку f(x) - *квазивыпуклая функция* и выполняется неравенство $f(x^+) < f(\overline{x})$, то мы получим, что $f[\lambda x^+ + (1-\lambda)\overline{x}] < f(\overline{x})$ при всех $\in (0;1)$. Однако это соотношение противоречит (10).

Заметим, что строго квазивыпуклые и *квазивогнутые функции* называются унимодальными.