Лекция 4. Нелинейное программирование.

Классификация методов нелинейного программирования. Классический метод определения условного экстремума. Метод множителей Лагранжа

Аннотация: Данная лекция раскрывает отличия и преимущества задач нелинейного программирования перед классическими задачами математического анализа, классифицирует разделы нелинейного программирования; формулирует задачи и классифицирует методы решения задач нелинейного программирования. Наиболее полно раскрыты такие методы, как классический метод определения условного экстремума и метод множителей Лагранжа.

1. Понятие нелинейного программирования

В большинстве инженерных задач построение математической модели не удается свести к *задаче линейного программирования*.

Математические модели в задачах проектирования реальных объектов или технологических процессов должны отражать реальные протекающие в них физические и, как правило, нелинейные процессы. Переменные этих объектов или процессов связанны между собой физическими нелинейными законами, такими, как законы сохранения массы или энергии. Они ограничены предельными диапазонами, обеспечивающими физическую реализуемость данного объекта или процесса. В результате, большинство задач *математического программирования*, которые встречаются в научно-исследовательских проектах и в задачах проектирования – это задачи *нелинейного программирования* (НП).

Пусть в математической модели проектируемого объекта или процесса непрерывная функция $F(\overline{X})$ представляет собой функцию цели (функцию качества),

$$h_1(\overline{X}), h_2(\overline{X}), h_3(\overline{X}), \ldots, h_m(\overline{X}), \quad i = \overline{1, m}$$

задают ограничения в виде равенств

$$g_{m+1}(\overline{X}), g_{m+2}(\overline{X}), g_{m+3}(\overline{X}), \ldots, g_p(\overline{X}), \quad j = \overline{m+1, p},$$

задают *ограничения в виде неравенств*,

где $\overline{X} = [x_1, x_2, x_3, \ldots, x_n], \ \overline{X} \in E^n$ - вектор параметров проектируемого объекта, процесса или системы, оптимальные значения которых должны быть найдены.

Тогда задача *нелинейного программирования* может быть сформулирована следующим образом:

найти вектор $\overline{X} = [x_1, x_2, x_3, \ldots, x_n], \ \overline{X} \in E^n$, доставляющий минимум (максимум) целевой функции $F(\overline{X})$ при m линейных и (или) нелинейных ограничений в виде равенств

$$h_i(\overline{X}) = 0, \qquad i = \overline{1, m}$$

и (p-m) линейных и (или) нелинейных ограничений в виде неравенств $g_j(\overline{X}) > 0, \qquad j = \overline{m+1, p}.$

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.

2. 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:

$$\text{максимизировать } f(x_1, x_2, \ldots, x_n)$$

при ограничениях

$$\begin{aligned} &g_1(x_1, x_2, \ldots, x_n) \geq 0; \\ &g_2(x_1, x_2, \ldots, x_n) \geq 0; \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &g_m(x_1, x_2, \ldots, x_n) \geq 0; \end{aligned}$$

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 3.1 (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 3.2. *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* $f(x_1, x_2, \ldots, x_n)$ *недифференцируема.*

Определение 3.1. Множество точек $S_1(x_1, x_2, \ldots, x_n)$ функции f(x) называется ***множеством стационарных точек***, если они удовлетворяют условию

$$\frac{\partial f(x)}{\partial x_j} = 0, \ j = \overline{1, n} \quad (3.1)$$

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

$$f\left(x_1^0, x_2^0, \ldots, x_n^0\right) \geq f\left(x_1, x_2, \ldots, x_n\right) \quad (3.2)$$

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* $(x_1^0, x_2^0, \ldots, x_n^0)$ *области* R *функция* f(x) *достигает относительного максимума, то*

$$\frac{\partial f(x_0)}{\partial x_j} = 0, \ j = \overline{1, n} \quad (3.3)$$

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 3.4. 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

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

Если

$$f[kx_1 + (1-k)x_2] \leq kf(x_1) + (1-k)f(x_2) \quad (3.5)$$

то функция называется вогнутой.

Если (3.4) или (3.5) 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 3.4. *Дифференцируемая функция* f(x) *строго вогнутая в некоторой окрестности точки* $x^0\ (x_1^0, x_2^0, \ldots, x_n^0)$, *если выполняются следующие условия*:

$$f_{11}(x_0) < 0; \quad \begin{vmatrix} f_{11}(x_0) & f_{12}(x_0) \\ f_{21}(x_0) & f_{22}(x_0) \end{vmatrix} > 0; \quad \begin{vmatrix} f_{11}(x_0) & f_{12}(x_0) & f_{13}(x_0) \\ f_{21}(x_0) & f_{22}(x_0) & f_{23}(x_0) \\ f_{31}(x_0) & f_{32}(x_0) & f_{33}(x_0) \end{vmatrix} < 0 \quad <0\ (3.6)$$

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 3.5. *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 3.6 (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 множество* $R(x) = \{x: \ g_1(x) \geq 0, g_2(x) \geq 0, \ldots, g_m(x) \geq 0, \ x \geq 0\}$ *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).

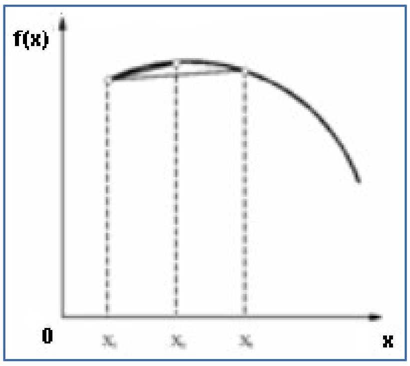

Рис. 7.1.

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

$$g_i(x) = g_i(x_1, x_2, \ldots, x_n) = 0 \quad (3.7)$$

является сепарабельной, то можно, определив из (3.7) 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 3.5. *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*

$$f(\lambda x_1 + (1-\lambda)x_2) \le \max\{f(x_1), f(x_2)\}. \quad (3.8)$$

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

Из этого определения следует, что функция f(x) - квазивыпукла, если из неравенства $f(x_2) \ge f(x_1)$ 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 $f(x_2) \ge f(x_1)$ 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.

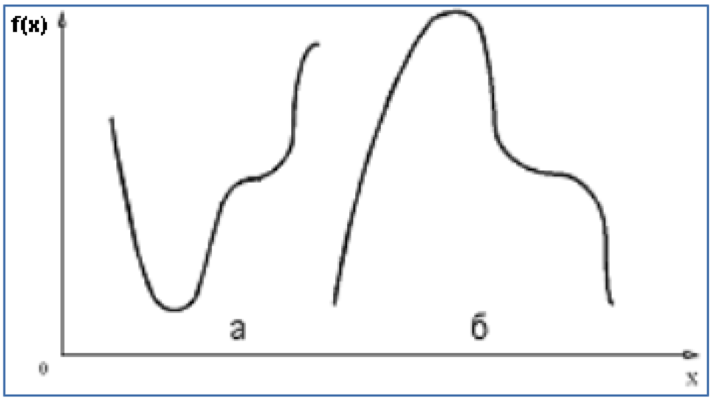

Рис. 7.2.

Определение 3.6. *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* $f(x_1) \ne f(x_2)$ *и* $\lambda \in (0;1)$ *выполняется неравенство*

$$f(\lambda x_1 + (1-\lambda)x_2) < \max\{f(x_1), f(x_2)\}. \quad (3.9)$$

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.

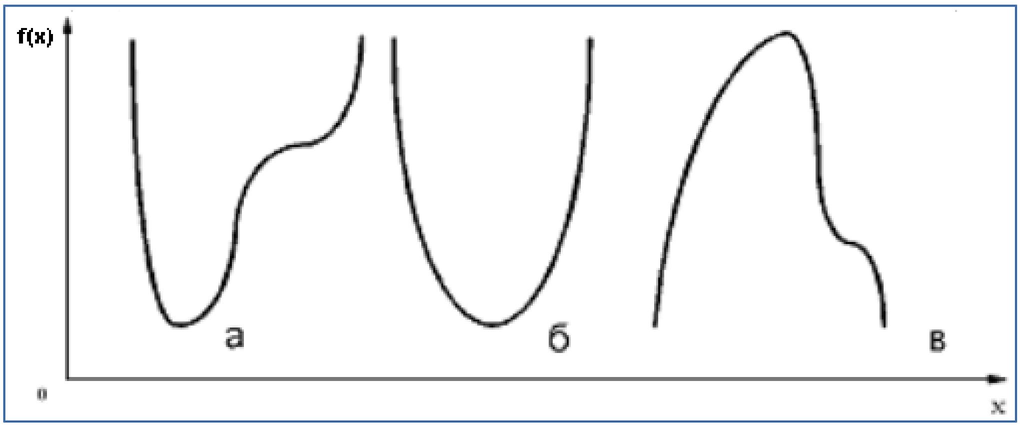

Рис. 7.3.

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 $f(x^+) < f(\overline{x})$. Поскольку R - 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

$f(\overline{x}) \le f[\lambda x^+ + (1-\lambda)\overline{x}]$ (3.10)

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 (3.10).

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

4. Метод множителей Лагранжа

Метод *множителей Лагранжа* позволяет отыскивать максимум $\langle$ или минимум $\rangle$ функции при ограничениях-равенствах. Основная идея метода состоит в переходе от задачи на условный экстремум к задаче отыскания безусловного экстремума некоторой построенной *функции Лагранжа*. Пусть задана задача НП при ограничениях-равенствах вида

$\text{минимизировать } f(x_1, x_2, \ldots, x_n)$ (4.1)

при ограничениях

$$\begin{aligned} &h_1(x_1, x_2, \ldots, x_n) = 0; \\ &h_2(x_1, x_2, \ldots, x_n) = 0; \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &h_m(x_1, x_2, \ldots, x_n) = 0; \end{aligned} \quad (4.2)$$

Предположим, что все функции f, h_1, h_2, ..., h_m - дифференцируемы. Введем набор переменных $\lambda_1, \lambda_2, \ldots, \lambda_m$ (число которых равняется числу ограничений), которые называются *множителями Лагранжа*, и составим *функцию Лагранжа* такого вида:

$$\begin{aligned} &L(x_1, x_2, \ldots, x_m, \lambda_1, \lambda_2, \ldots, \lambda_m) = \\ &= f(x_1, x_2, \ldots, x_n) + \sum_{i=1}^{m} \lambda_i h_i(x_1, x_2, \ldots, x_n) \end{aligned} \quad (4.3)$$

Справедливо такое утверждение: для того чтобы вектор $x^0 = \{x_1^0, x_2^0, \ldots, x_n^0\}$ являлся решением задачи (4.1) при ограничениях (4.2), необходимо, чтобы существовал такой вектор $\Lambda^0 = \{\lambda_1^0, \lambda_2^0, \ldots, \lambda_m^0\}$, что пара векторов удовлетворяла бы системе уравнений

$$\frac{\partial L(x^0, \Lambda^0)}{\partial x_j} = 0, \quad j = \overline{1, n}; \quad (4.4)$$

$$\frac{\partial L(x^0, \Lambda^0)}{\partial \lambda_i} = 0, \quad i = \overline{1, m}; \quad (4.5)$$

Покажем необходимость условий (4.4), (4.5) на простом примере:

$$\text{минимизировать } f(x_1, x_2, x_3) \quad (4.6)$$

при ограничениях

$$\begin{aligned} &h_1(x_1, x_2, x_3) = 0, \\ &h_2(x_1, x_2, x_3) = 0. \end{aligned} \quad (4.7)$$

Ограничения (4.7) определяют *допустимую область* S, которая представляет собой кривую в пространстве $R^{(2)}$ и является результатом пересечения $h_1(x)$ и $h_2(x)$.

Допустимим, что рассматриваемая задача имеет точку минимума в S_1: $x^+ = \{x_1^+, x_2^+, , x_n^+\}$, функции f, h_1, h_2 имеют непрерывные производные первого порядка на некотором открытом множестве и градиенты

$$\nabla h_1(x) = \left[\frac{\partial h_1}{\partial x_1}; \frac{\partial h_1}{\partial x_2}; \frac{\partial h_1}{\partial x_3};\right]^T; \quad \nabla h_1(x) = \left[\frac{\partial h_2}{\partial x_1}; \frac{\partial h_2}{\partial x_2}; \frac{\partial h_2}{\partial x_3};\right]^T$$

линейно независимы.

Если две переменные в уравнениях (4.7) можно выразить через третью в виде $x_2=U(x_1)$, $x_3=V(x_1)$, то подставив их в целевую функцию (4.6), преобразуем исходную задачу в следующую задачу без ограничений, которая содержит лишь одну переменную x_1:

$$\text{минимизировать} f(x_1, U(x_1), V(x_1)). \quad (4.8)$$

Поскольку градиенты $\nabla h_1(x_1, x_2, x_3),\ i = 1, 2$, непрерывны и линейно независимы, то можно применить известную теорему математического анализа о неявной функции и найти стационарную точку x_1^+, а потом $x_2^+ = U(x_1^+), x_3^+ = V(x_1^+)$.

Приведенный подход можно в принципе распространить и на случай функции n переменных $f(x), x = [x_1, x_2, \dots, x_n]^T$ при наличии m ограничений-равенств:

$$h_1(x) = 0, h_2(x) = 0, \dots, h_m(x) = 0. \quad (4.9)$$

Если функции $h_1(x), \dots, h_m(x)$ удовлетворяют условиям теоремы о неявной функции, то m из n переменных уравнений (4.9) можно выразить через остальные (n-m) переменных, подставить их в f(x) и таким образом преобразовать задачу минимизаци с ограничениями в задачу безусловной минимизации с (n-m) переменными. Однако такой подход трудно реализовать на практике, поскольку очень трудно разрешить уравнения (4.9) относительно некоторых переменных. В общем случае это совсем невозможно.

Поэтому рассмотрим другой подход, который базируется на методе *множителей Лагранжа*.

Пусть x^+ - точка минимума f(x), *определяемого выражением* (4.8). В соответствии с известной теоремой математического анализа о неявной функции можно записать

$$\frac{df}{dx_1} = \frac{\partial f}{\partial x_1} + \frac{\partial f}{\partial x_2} \cdot \frac{dU}{dx_1} + \frac{\partial f}{\partial x_3} \cdot \frac{dV}{dx_1} = 0 \quad (4.10)$$

Аналогичные соотношения получим для ограничений

$$\frac{dh_i}{dx_1} = \frac{\partial h_i}{\partial x_1} + \frac{\partial h_i}{\partial x_2} \cdot \frac{dU}{dx_1} + \frac{\partial h_i}{\partial x_3} \cdot \frac{dV}{dx_1} = 0,\ i = 1, 2 \quad (4.11)$$

Запишем уравнения (4.10), (4.11) совместно в виде

$$A \times \begin{bmatrix} 1 \\ \frac{dU}{dx_1} \\ \frac{dV}{dx_1} \end{bmatrix} = 0, \quad (4.12)$$

где

$$A = \begin{bmatrix} \nabla f(x^+) \\ \nabla h_1(x^+) \\ \nabla h_2(x^+) \end{bmatrix}.$$

Поскольку вектор $\left[1, \frac{dU}{dx_1}, \frac{dV}{dx_1}\right]$ не является нулевым, то из (4.12) следует, что $\det A = 0$. Из этого следует, что вектора-строки матрицы A должны быть линейно зависимы. Следовательно, существуют три таких скаляра a, b, c не все равные 0, что

$$a\nabla f(x^+) + b\nabla h_1(x^+) + c\nabla h_2(x^+) = 0 \quad (4.13)$$

Скаляр a не может равняться 0, так как в соответствии с предположением ∇h_1 и ∇h_2 - линейно независимы. Поэтому после деления (4.13) на a, получим

$$\nabla f(x^+) + \lambda_1\nabla h_1(x^+) + \lambda_2\nabla h_2(x^+) = 0 \quad (4.14)$$

Таким образом, для задачи минимизации с ограничениями (4.6) существуют такие λ_1, λ_2, для которых справедливо уравнение (4.14) и которые одновременно не обращаются в нуль. Итак, справедливость условий (4.4) для случая $n=3$ показана.

Таким образом, для отыскания минимума (4.6) при условиях (4.7) необходимо найти стационарную точку *функции Лагранжа*:

$$L(x, \Lambda) = f(x) + \lambda_1 h_1(x) + \lambda_2 h_2(x).$$

Для того чтобы найти искомые значения $\lambda_1, \lambda_2, x^+$, необходимо решить совместно систему уравнений (4.14), (4.5). С геометрической точки зрения условие (4.14) означает, что $f(x^+)$ лежит в плоскости, натянутой на векторы $h_1(x^+)$.

Теперь рассмотрим общий случай для произвольных n. Пусть задана задача НП в виде (4.1), (4.2), все функции $f(x), h_1(x), i = \overline{1,m}\ (m < n)$, имеют непрерывные частные производные на множестве $R^{(n)}$. Пусть $S(x)$ - подмножество множества $R^{(n)}$, на котором все функции $h_1(x) = 0, \quad i = \overline{1,m}$, то есть $S = \{x : h_1(x) = 0,\ i = \overline{1,m}\}$. Тогда справедлива такая теорема о *множителях Лагранжа*.

Теорема 4.7. *Допустим, что существует такая точка* x^+, *в которой достигается относительный экстремум задачи НП (4.1) при условиях (4.2). Если ранг матрицы* $I = \left[\frac{\partial h_j(x)}{\partial x_j}\right], \; i = \overline{1,m}, \; j = \overline{1,n}$ *в точке* x^+ *равен* m, *то существуют* m *чисел* $\lambda_1, \lambda_2, \ldots, \lambda_m$, *не все из которых равны нулю одновременно, при которых*

$$\nabla f(x^+) + \sum_{i=1}^{m} \lambda_i \nabla h_i(x^+) = 0. \quad (4.15)$$

Эта теорема обосновывает метод *множителей Лагранжа*, который состоит из следующих шагов.

Составляют *функцию Лагранжа* $L(x, \Lambda)$

Находят частные производные $\frac{\partial L(x,\Lambda)}{\partial x_j}, \; j = \overline{1,n}; \; \frac{\partial L(x,\Lambda)}{\partial \lambda_i}, \; i = \overline{1,m};$

Решают систему уравнений

$$\begin{aligned} &\frac{\partial L(x,\Lambda)}{\partial x_j} = 0, \; j = \overline{1,n}; \\ &\frac{\partial L(x,\Lambda)}{\partial \lambda_i} = h_i(x) = 0, \; i = \overline{1,m}. \end{aligned} \quad (4.16)$$

и отыскивают точки $x^0 = [x_j^0]$, удовлетворяющие системе (4.16).

Найденные точки x^0 дальше исследуют на максимум (или минимум).