Лекция № 5

Задача нелинейного программирования при ограничениях типа равенств

Постановка задачи:

Решается задача: $f(X) \to \underset{X \in \boldsymbol{X}}{extr}$ (*)

$$\boldsymbol{X} = \{X : \varphi_j(X) = 0, j = 1..m \le n\}$$

Особенностью решения задачи является то, что допустимые решения, на которых ищется экстремум функции, являются решением уравнений $\varphi_j(X) = 0, \quad j = 1..m$.

Графически, экстремумы функции представляют собой точки касания линии уровня функции и поверхности, задаваемой пересечением ограничений.

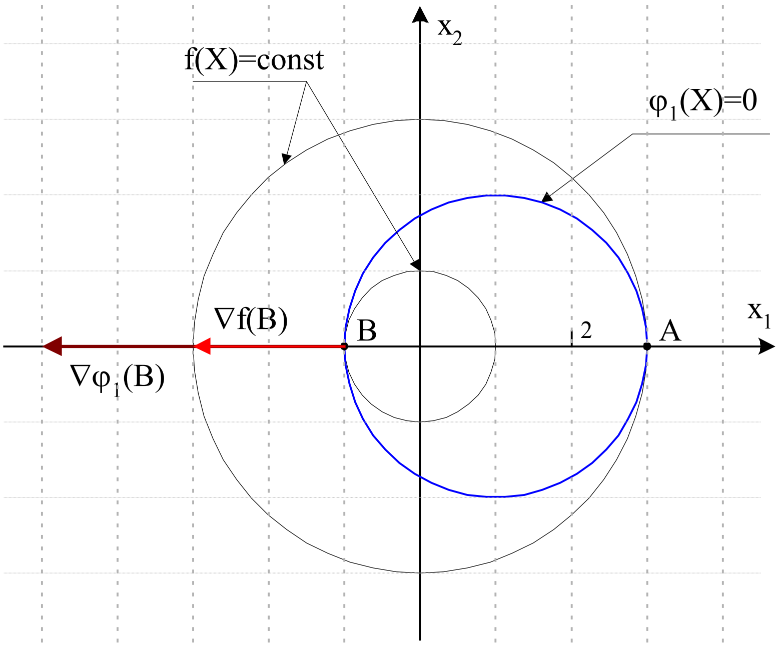
Пример 1

$$f(X) = x_1^2 + x_2^2 \to extr$$

$$\varphi_1(X) = (x_1 - 1)^2 + x_2^2 - 4 = 0$$

$$\nabla f(B) = \begin{pmatrix} -2 \\ 0 \end{pmatrix} \quad \nabla \varphi_1(B) = \begin{pmatrix} -4 \\ 0 \end{pmatrix}$$

Здесь точка A - условный максимум, точка B - условный минимум



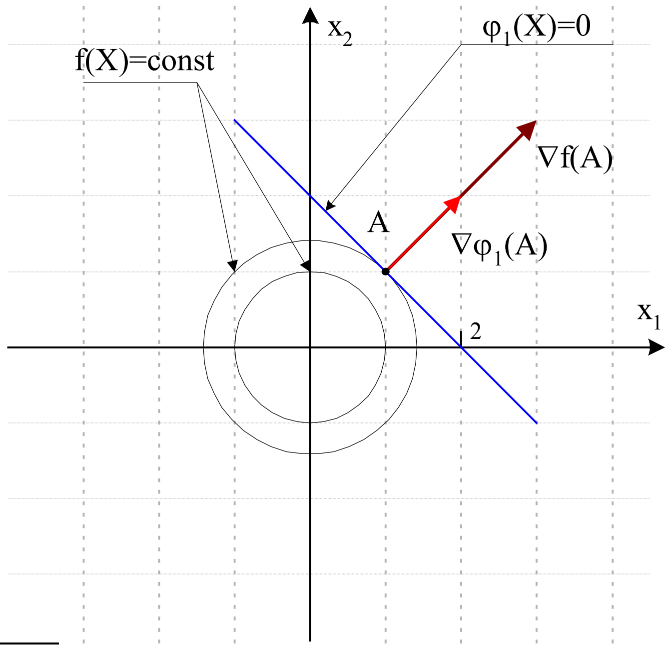
Пример №2

$$f(X) = x_1^2 + x_2^2 \to extr$$

$$\varphi_1(X) = x_1 + x_2 - 2 = 0$$

$$\nabla f(A) = \begin{pmatrix} 2 \\ 2 \end{pmatrix} \quad \nabla \varphi_1(A) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Здесь точка A - условный минимум.



Как видно из чертежа, точки касания обладают следующими свойствами:

- точка касания принадлежит поверхности, определяемой пересечением ограничений;
- градиенты функции и поверхности, определяемой пересечением ограничений, в точке касания являются линейно-зависимыми.

<u>Последовательность графического решения задачи:</u>

1. Построить ограничения и определить множество допустимых решений X.
2. Вычислить точку касания, пользуясь условиями касания.
3. Вычислить функцию в точке касания, определить конфигурацию и построить соответствующую линию уровня.

Если система ограничений $\varphi_j(X) = 0, \quad j = 1..m$ разрешима относительно любых m переменных, то решение задачи может быть получено методом исключений.

(1) Метод исключений

Допусти, что система ограничений $\varphi_j(X) = 0, \quad j = 1..m$ может быть разрешена относительно некоторых m переменных:

$$\begin{aligned}
x_1 &= \psi_1(x_{m+1}, x_{m+2}, \ldots x_n) \\
x_2 &= \psi_2(x_{m+1}, x_{m+2}, \ldots x_n) \\
x_3 &= \psi_3(x_{m+1}, x_{m+2}, \ldots x_n) \\
&\ldots \\
x_m &= \psi_m(x_{m+1}, x_{m+2}, \ldots x_n)
\end{aligned} \quad (1.1)$$

тогда можно утверждать, что переменные $x_{m+1}, x_{m+2}, \ldots, x_n$ являются независимыми (свободными для выбора), а это означает, что решение задачи (*) может быть сведено к решению задачи на безусловный экстремум.
Подставим выражения в исходную функцию, получим:

$$f(X) = f(x_1, x_2, \ldots, x_m, x_{m+1}, \ldots, x_n) =$$
$$= f(\psi_1(x_{m+1}, x_{m+2}, \ldots, x_n), \psi_2(x_{m+1}, x_{m+2}, \ldots, x_n), \ldots, x_{m+1}, \ldots, x_n) == \tilde{f}(x_{m+1}, x_{m+2}, \ldots x_n)$$

Получена новая функция переменных $x_{m+1}, x_{m+2}, \ldots, x_n$, экстремум которой требуется найти. Поскольку переменные $x_{m+1}, x_{m+2}, \ldots, x_n$ являются независимыми, будем искать безусловный экстремум. Оптимальные значения зависимых переменных $x_1, x_2, \ldots, x_m$ могут быть получены из соотношений (1.1).

<u>Алгоритм решения задачи методом исключений:</u>

1. Разрешить систему ограничений относительно любых m переменных.
2. Подставить полученные выражения в исходную функцию и перейти к задаче безусловной оптимизации.
3. Решить полученную задачу безусловной оптимизации - найти стационарные точки и проверить достаточные условия.
4. Вернуться к исходной задаче и, используя решение задачи безусловной оптимизации, найти значения недостающих переменных.

Возможности применения метода исключения ограничены тем, что система ограничений в большинстве задач носит нелинейный характер, а, следовательно, неразрешима относительно нужного числа переменных.

(2) Метод множителей Лагранжа (аппарат необходимых и достаточных условий)

Определение 1.

Функция $L(X,\lambda)=f(X)+\sum_{j=1}^{m}\lambda_j\varphi_j(X)$ называется *классической функцией Лагранжа*.

Функция $L(X,\lambda)$ зависит от $n+m$ переменных: n штук x_i и m штук λ_j - называемых *множителями Лагранжа*.

Определение 2.

Вторым дифференциалом функции Лагранжа называется функция: $d^2L(X,\lambda)=\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 L}{\partial x_i \partial x_j}dx_i dx_j$

Определение 3.

Первым дифференциалом ограничения $\varphi_j(X)$ называется функция: $d\varphi_j(X)=\sum_{i=1}^{n}\frac{\partial\varphi_j(X)}{\partial x_i}dx_i$

Теорема 1. (о необходимых условиях экстремума)

Пусть X^* есть точка локального условного экстремума в задаче (*), и при этом $\nabla\varphi_j(X^*),\quad j=1..m$ являются линейно-независимыми, то найдутся такие λ_j^*, что:

- $\frac{\partial L(X^*,\lambda^*)}{\partial x_i}=0,\quad i=1..n$ (условие стационарности функции Лагранжа по X)
- $\varphi_j(X^*)=0,\quad j=1..m$ (условие допустимости решения)

Теорема 2. (о достаточных условиях экстремума)

Если в точке (X^*,λ^*) выполняются необходимые условия экстремума и $d^2L(X^*,\lambda^*)>0$ при всех ненулевых dx_i, таких, что $d\varphi_j(X^*)=0,\quad j=1..m$, то X^* - точка условного локального минимума функции, если же при всех тех же условиях $d^2L(X^*,\lambda^*)<0$, то X^* - точка условного локального максимума функции.

<u>Алгоритм решения задачи методом множителей Лагранжа:</u>

1. Записать классическую функцию Лагранжа: $L(X,\lambda)=f(X)+\sum_{j=1}^{m}\lambda_j\varphi_j(X)$
2. Записать необходимые условия экстремума ФМП при ограничениях типа равенств:

$$\begin{cases}\frac{\partial L(X,\lambda)}{\partial x_i}=\frac{\partial f(X)}{\partial x_i}+\sum_{j=1}^{m}\lambda_j\frac{\partial\varphi_j(X)}{\partial x_i}=0, & i=1..n\\ \varphi_j(X)=0, & j=1..m\end{cases}$$

3. Решить полученную систему. Решение системы – условно-стационарные точки (X^*,λ^*).
4. Проверить достаточные условия экстремума в каждой точке (X^*,λ^*), для этого

Записать второй дифференциал функции Лагранжа: $d^2L=\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial^2 L}{\partial x_i\partial x_j}dx_i dx_j$

Записать дифференциалы ограничений $d\varphi_j(X)=\sum_{i=1}^{n}\frac{\partial\varphi_j(X)}{\partial x_i}dx_i$

В каждой точке (X^*,λ^*)

4.1. Вычислить второй дифференциал $d^2L(X^*,\lambda^*)$

4.2. Записать условия равенства 0 дифференциалов ограничений в каждой точке X^*:

$$d\varphi_j(X^*) = \sum_{i=1}^{n}\left(\frac{\partial\varphi_j}{\partial x_i}\right)_{X^*} \cdot dx_i = 0, \quad j = 1..m$$

4.3. Используя уравнения из п. 4.2, выразить любые m дифференциалов переменных через оставшиеся $(n-m)$ и подставить их в выражение для $d^2L(X^*,\lambda^*)$.

4.4. Определить знак $d^2L(X^*,\lambda^*)$:

- если $d^2L(X^*,\lambda^*)>0$ при $dx_i \neq 0$, то точка X^* - точка условного локального минимума в задаче;
- если $d^2L(X^*,\lambda^*)<0$ при $dx_i \neq 0$, то точка X^* - точка условного локального максимума в задаче.

(3) Метод штрафной функции

Метод штрафной функции относится к численным методам решения задачи (*).
Метод штрафной функции предусматривает поиск решения задачи в результате решения последовательности задач безусловной минимизации вида:

$$F(X,r^k) = f(X) + \Phi(X,r^k) \to \min_{X\in R^n}, \qquad (3.1)$$

здесь $F(X,r^k)$ - штрафная функция

$\Phi(X,r^k)$ - штраф

$r^k > 0$ - штрафной параметр, при решении каждой задачи (3.1) фиксируется

Функция штрафа конструируется из условия:

$$\Phi(X,r^k) = \begin{cases} =0, & \varphi_j(X) = 0 \\ >0, & \varphi_j(X) \neq 0 \end{cases}$$

причем, чем больше $|\varphi_j(X)|$, тем больше штраф. Кроме того, штраф должен быть таким, чтобы при $r^k \to \infty$ штраф $\Phi(X,r^k) \to \infty$ при невыполнении ограничений, т.е. становился тем больше, чем больше не выполняются ограничения.

Обычно, в качестве штрафа используют функцию вида: $\Phi(X,r^k) = \frac{r^k}{2}\sum_{j=1}^{m}\varphi_j^2(X)$

<u>Идея метода штрафной функции</u>: при каждом значении k ищется точка $X^*(r^k)$ минимума в задаче (3.1) при заданном значении параметра r^k с помощью одного из методов безусловной минимизации. Полученная точка $X^*(r^k)$ используется в качестве начальной в следующей задаче (3.1) при возрастающем значении параметра r^k.

Доказано, что при $r^k \to \infty$, последовательность получаемых точек $\{X^*(r^k)\}$ стремится к X^*:

$$\lim_{r^k\to\infty} X^*(r^k) = X^*$$

Существует связь между значением параметра штрафа r^k и множителями Лагранжа:

$$\lambda_j(r^k) = r^k \cdot \varphi_j(X(r^k))$$

$$\lambda_j^* = \lim_{r^k\to\infty} \lambda_j(r^k)$$

Замечание. В случае поиска условного экстремума квадратичной функции при линейном ограничении задача (3.1) может быть решена аналитически.

<u>Алгоритм аналитического решения задачи методом штрафной функции:</u>

1. Записать штрафную функцию: $F(X,r) = f(X) + \frac{r}{2} \cdot \sum_{j=1}^{m} \varphi_j^2(X)$
2. Записать необходимые условия экстремума для штрафной функции:
   $$\left\{ \frac{\partial F(X,r)}{\partial x_i} = 0 \quad i = 1..n \right.$$
3. Найти решение полученной системы: $X^*(r)$. Решение системы зависит от параметра r.
4. Найти условно-стационарную точку в задаче: $X^* = \lim_{r \to \infty} X^*(r)$.
5. Составить матрицу Гессе для штрафной функции: $H(X^*(r)) = \left(\frac{\partial^2 F(X,r)}{\partial x_i \partial x_j} \right)_{\substack{i=1..n \\ j=1..n}}$
6. Исследовать знакоопределенность матрицы при $r \to \infty$ по критерию Сильвестра.
7. Записать оценку множителей Лагранжа: $\lambda_j^* = \lim_{r \to \infty} r \cdot \varphi_j(X^*(r)) \quad j = 1..m$.

Замечание.

В случае поиска максимума, используют штрафную функцию вида: $F(X,r) = f(X) - \frac{r}{2} \cdot \sum_{j=1}^{m} \varphi_j^2(X)$, и оценку множителей Лагранжа вычисляют по формуле: $\lambda_j^* = -\lim_{r \to \infty} r \cdot \varphi_j(X^*(r)) \quad j = 1..m$

Рассмотрим решение ***Примера 2*** различными методами:

Дано:

$$f(X) = x_1^2 + x_2^2 \to \text{extr}$$

$$\varphi_1(X) = x_1 + x_2 = 2$$

Найдем решение задачи **методом исключений**:

Решение:

1. Выразим одну из переменных из ограничения: $x_1 = 2 - x_2$
2. Подставим полученное выражение в функцию: $f(X) = (2 - x_2)^2 + x_2^2 = 4 - 4x_2 + 2x_2^2 = \tilde{f}(x_2)$
3. Найдем экстремум полученной функции:

$$\tilde{f}(x_2) = 4 - 4x_2 + 2x_2^2 \to \text{extr}$$

$$\frac{d\tilde{f}}{dx_2} = -4 + 4x_2 = 0 \quad \Rightarrow \quad x_2^{\ *} = 1$$

$$\frac{d^2\tilde{f}}{d(x_2)^2} = 4 > 0 \quad \Rightarrow \quad \min$$

4. Найдем значение оставшейся переменной: $x_1^{\ *} = 2 - x_2^{\ *} = 2 - 1 = 1$

Ответ: получена точка $A = (1, 1)$ - условный локальный минимум.

Найдем решение задачи **методом множителей Лагранжа**:

Решение:

1. Запишем функцию Лагранжа: $L(X, \lambda) = x_1^2 + x_2^2 + \lambda_1(x_1 + x_2 - 2)$
2. Запишем необходимые условия экстремума.

$$\left.\begin{aligned} \frac{\partial L}{\partial x_1} &= 2x_1 + \lambda_1 = 0 \\ \frac{\partial L}{\partial x_1} &= 2x_2 + \lambda_1 = 0 \\ x_1 + x_2 - 2 &= 0 \end{aligned}\right\}$$

3. Найдем координаты условно-стационарных точек.

$$\begin{cases} 2x_1 + \lambda_1 = 0 \\ 2x_2 + \lambda_1 = 0 \\ x_1 + x_2 - 2 = 0 \end{cases} \quad \Rightarrow \quad \begin{cases} 2x_1 + \lambda_1 = 0 \\ x_1 - x_2 = 0 \\ x_1 + x_2 - 2 = 0 \end{cases} \quad \Rightarrow \quad \begin{cases} x_1^* = 1 \\ x_2^* = 1 \\ \lambda_1^* = -2 \end{cases}$$

Получена условно-стационарная точка $A = (1, 1, -2)$

4. Установим тип полученной точки с помощью достаточных условий экстремума.

Составляем второй дифференциал функции Лагранжа:

$$\frac{\partial^2 L}{\partial x_1^2} = 2 \qquad \frac{\partial^2 L}{\partial x_1 \partial x_2} = \frac{\partial^2 L}{\partial x_1 \partial x_2} = 0 \qquad \frac{\partial^2 L}{\partial x_2^2} = 2$$

$$d^2 L(X, \lambda) = 2(dx_1)^2 + 2(dx_2)^2$$

Составляем дифференциал ограничения:

$$\frac{\partial \varphi_1}{\partial x_1} = 1 \qquad \frac{\partial \varphi_1}{\partial x_2} = 1$$

$$d\varphi_1(X) = dx_1 + dx_2$$

В точке $A = (1, 1, -2)$ имеем: $d^2 L(A) = 2(dx_1)^2 + 2(dx_2)^2$ при условии $d\varphi_1(A) = dx_1 + dx_2 = 0$. Значит $dx_1 = -dx_2$, тогда получим $d^2 L(A) = 4(dx_2)^2 > 0$ при $dx_2 \neq 0$, следовательно точка A – условный локальный минимум.

<u>Ответ:</u> получена точка $A = (1, 1)$ - условный локальный минимум.

Найдем решение задачи **<u>методом штрафной функции</u>**:

<u>Решение:</u>

1. Запишем штрафную функцию: $F(X, r) = x_1^2 + x_2^2 + \frac{r}{2}(x_1 + x_2 - 2)^2$

2. Запишем необходимые условия экстремума.

$$\left.\begin{aligned} \frac{\partial F(X,r)}{\partial x_1} &= 2x_1 + r(x_1 + x_2 - 2) = 0 \\ \frac{\partial F(X,r)}{\partial x_1} &= 2x_2 + r(x_1 + x_2 - 2) = 0 \end{aligned}\right\}$$

3. Найдем координаты стационарных точек штрафной функции:

$$\begin{cases} (2+r)x_1 + r \cdot x_2 = 2r \\ r \cdot x_1 + (2+r)x_2 = 2r \end{cases} \Rightarrow \begin{cases} x_1^* = \dfrac{4r}{4+4r} \\ x_2^* = \dfrac{4r}{4+4r} \end{cases}$$

Получена условно-стационарная точка $X^*(r) = \left(\frac{4r}{4+4r}, \frac{4r}{4+4r}\right)$

4. Найдем координаты условно-стационарных точек.

$x_1^* = x_2^* = \lim\limits_{r \to \infty} \frac{4r}{4+4r} = 1$, следовательно получена условно-стационарная точка $A = (1, 1)$

5. Составим матрицу Гессе для штрафной функции: $H(X^*(r)) = \begin{pmatrix} 2+r & r \\ r & 2+r \end{pmatrix}$

6. По критерию Сильвестра: $\Delta_1 = 2 + r > 0$ при $r \to \infty$

$\Delta_2 = 4 + 4r > 0$ при $r \to \infty$,

значит матрица $H(X^*(r)) > 0$ и точка $A = (1, 1)$ - условный минимум.

7. Запишем оценку множителя Лагранжа:

$$\lambda_1 = \lim_{r \to \infty} r \cdot \left(\frac{4r}{4+4r} + \frac{4r}{4+4r} - 2 \right) = \lim_{r \to \infty} r \cdot \left(\frac{8r - 2 \cdot (4+4r)}{4+4r} \right) = \lim_{r \to \infty} r \cdot \left(\frac{-8}{4+4r} \right) = \lim_{r \to \infty} \left(\frac{-8r}{4+4r} \right) = -2$$

<u>Ответ:</u> получена точка $A = (1, 1)$ - условный локальный минимум.