Лекция № 2

Постановка задачи оптимизации

Постановка задачи оптимизации подразумевает:
1. Формулировку цели, ради которой ставится задача.
2. Определение критерия отбора путей достижения цели.
3. Задание множества путей достижения цели: множества допустимых решений (МДР).

Задача оптимизации формулируется следующим образом: найти среди всех путей, ведущих к цели, наилучший, для которого критерий принимает оптимальное значение.

Пример.
Рассчитать оптимальные размеры цилиндрического контейнера для размещения радиоаппаратуры, при условии что площадь поверхности контейнера задана.

Цель: Рассчитываемый контейнер должен обладать наибольшим (максимальным) объемом .

Критерий: $V = S_{осн} \cdot h = \pi \cdot R^2 \cdot h \to \max$, где R и h - неизвестны и должны быть определены.

МДР: $S_{пов} = 2 \cdot S_{осн} + S_{бок} = 2 \cdot \pi \cdot R^2 + 2 \cdot \pi \cdot R \cdot h = L$, где L - известная заданная величина.

Как правило множество допустимых решений содержит огромное (чаще бесконечное) число возможных путей решения задачи (в примере – это различные значения R и h, удовлетворяющих условию $S_{пов} = L$), поэтому перебор, как способ решения, автоматически исключается, и возникает необходимость в применении специальных математических оптимизационных методов, которые работают с определенными математическими моделями задач. В курсе будут рассматриваться так называемые *задачи математического программирования*.

Постановка задачи математического программирования:
1. Задается целевая функция многих переменных $f(X)$, $X = (x_{1,} ... x_n)^T$, определенная на n-мерном евклидовом пространстве R^n.
2. Определяется критерий: найти минимальное значение функции, максимальное значение функции, найти значения переменных, при котором функция примет конкретное значение и т.д.
3. Задается МДР: $\boldsymbol{X} \subseteq R^n$, среди элементов которого осуществляется поиск решения.

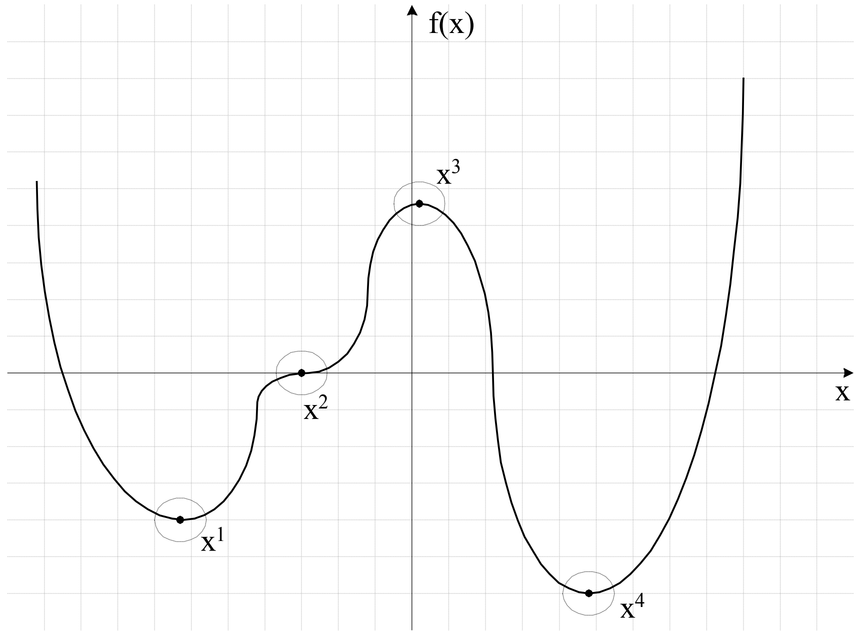
Т.о. задача математического программирования формулируется следующим образом: найти такой вектор X^* из множества допустимых решений $\boldsymbol{X}$, которому соответствует требуемое, с точки зрения цели, значение функции на этом множестве:

Определение 1. Точка $X^* \in \boldsymbol{X}$ называется точкой *локального минимума (локального максимума)* функции $f(X)$, если найдется $\delta > 0$, такое что:

$$f(X^*) \le f(X) \text{ при } \left\| X - X^* \right\| < \delta \quad \left(f(X^*) \ge f(X) \text{ при } \left\| X - X^* \right\| < \delta \right)$$

Определение 2. Точка $X^* \in \boldsymbol{X}$ называется точкой *глобального минимума (глобального максимума)* $f(X)$ на $\boldsymbol{X}$, если:

$$f(X^*) \le f(X) \quad \forall X \in \boldsymbol{X} \quad \left(f(X^*) \ge f(X) \quad \forall X \in \boldsymbol{X} \right)$$



На рисунке точки x^1 и x^4 являются точками локальных минимумов; точка x^3 - локальный максимум.

Кроме того, точка x^4 является также и глобальным минимумом.

Определение 3. Задача поиска всех минимумов и максимумов целевой функции называется задачей поиска *экстремума*. Решением задачи поиска экстремума являются пары $(X^*, f(X^*))$, включающие точки и значение целевой функции в них.

Множество точек экстремума может содержать конечное число точек (в том числе и одну), бесконечное число точек или быть пустым.

Если множество допустимых решений задается ограничениями (условиями), накладываемыми на вектор X^*, то решается задача поиска условного экстремума. Если, ограничения (условия) на вектор X^* отсутствуют, решается задача поиска безусловного экстремума.

Замечание. Задачи поиска максимума функции $f(X)$ сводится к задачам поиска минимума путем замены знака перед функцией на противоположный:

$$f(X^*) = \max f(X) = -\min[-f(X)]$$

Задачи поиска безусловного экстремума ФМП

Задачи поиска безусловного экстремума имеют огромное значение в теории оптимизации, т.к. большинство задач на условный экстремум сводятся, путем замены целевой функции, к задачам поиска безусловного экстремума.

Теорема о необходимых условиях экстремума (НУ)

Если точка X^* является точкой безусловного локального экстремума (минимума или максимума) $f(X)$, и функция $f(X)$ непрерывно дифференцируема в ней, то $\nabla f(X^*) = 0$.

Замечание. Точки, в которых выполняются необходимые условия безусловного экстремума функции $\nabla f(X^*) = 0$, называются *стационарными* точками функции, среди них могут быть минимумы, максимумы, а также другие точки, не являющиеся экстремумами функции.

Теорема о необходимых условиях экстремума 2-го порядка (НУ 2-го порядка)

Если точка X^* является точкой безусловного локального минимума (максимума) $f(X)$, и функция $f(X)$ дважды непрерывно дифференцируема в ней, то $H(X^*) \geq 0$ ($H(X^*) \leq 0$).

Теорема о достаточных условиях безусловного экстремума (ДУ)

Если функция $f(X)$ дважды непрерывно дифференцируема в X^*, $\nabla f(X^*) = 0$ и $H(X^*) > 0$, то X^* – точка локального минимума функции $f(X)$, если же при этом $H(X^*) < 0$, то X^*– точка локального максимума функции $f(X)$.

<u>Алгоритм решения задачи</u>
<u>на безусловный экстремум с использованием необходимых и достаточных условий</u>

необходимые условия экстремума	1. Записать градиент функции $f(X)$: $\nabla f(X) = \left(\frac{\partial f(X)}{\partial x_1}, .., \frac{\partial f(X)}{\partial x_n} \right)^T$. 2. Записать необходимые условия безусловного экстремума – составить систему алгебраических уравнений вида: $\left\{ \frac{\partial f(X)}{\partial x_i} = 0, \quad i = 1..n \right.$ 3. Найти стационарные точки функции X^*, решив полученную систему.
достаточные условия экстремума и необходимые условия 2-го порядка	4. Составить матрицу Гессе $H(X)$. 5. Вычислить матрицу Гессе в точках X^*. 6. Проверить знакоопределенность матрицы $H(X^*)$ для каждой точки: ❑ $H(X^*) > 0$ - X^*-локальный минимум функции (ДУ) ❑ $H(X^*) < 0$ - X^*-локальный максимум функции (ДУ) ❑ $H(X^*) \geq 0$ - требуются дополнительная проверка на локальный минимум функции (НУ 2-го порядка) ❑ $H(X^*) \leq 0$ - требуются дополнительная проверка на локальный максимум функции (НУ 2-го порядка) ❑ $H(X^*) <> 0$ - в точке X^* нет экстремума.

Исследование знакоопределенности матрицы $H(X^*)$:

- $H(X^*) > 0$ - либо по критерию Сильвестра, либо на основании определения (все $\lambda_j > 0$)
- $H(X^*) < 0$ - либо по критерию Сильвестра, либо на основании определения (все $\lambda_j < 0$)
- $H(X^*) \geq 0$ - на основании определения (все $\lambda_j \geq 0$), либо из условия, что все главные миноры матрицы неотрицательны.
- $H(X^*) \leq 0$ - на основании определения (все $\lambda_j \leq 0$), либо из условия, что все главные миноры матрицы неположительны.
- $H(X^*) <> 0$ - на основании определения (λ_j разных знаков).

Пример 1

<u>Дано:</u> $f(X) = 3x_1^3 + x_2^2 - 9x_1 + 4x_2 \to extr$

<u>Решение:</u>

1. $\dfrac{\partial f}{\partial x_1} = 9x_1^2 - 9 \qquad \dfrac{\partial f}{\partial x_2} = 2x_2 + 4$

$\nabla f(X) = (9x_1^2 - 9; \quad 2x_2 + 4)^T$

2. $\begin{cases} 9x_1^2 - 9 = 0 \\ 2x_2 + 4 = 0 \end{cases}$

3. Решаем систему:

$$\begin{cases} 9x_1^2 - 9 = 0 \\ 2x_2 + 4 = 0 \end{cases} \Rightarrow \begin{cases} x_1^2 = 1 = 0 \\ x_2 = -2 \end{cases} \Rightarrow \left[\begin{array}{l} \begin{cases} x_1 = 1 \\ x_2 = -2 \end{cases} \\ \\ \begin{cases} x_1 = -1 \\ x_2 = -2 \end{cases} \end{array} \right.$$

Таким образом получены 2 стационарные точки функции: $A = (1, -2); \quad B = (-1, -2)$

4. $\dfrac{\partial^2 f}{\partial x_1^2} = 18x_1 \qquad \dfrac{\partial^2 f}{\partial x_1 \partial x_2} = \dfrac{\partial^2 f}{\partial x_2 \partial x_1} = 0 \qquad \dfrac{\partial^2 f}{\partial x_2^2} = 2$

$$H(X) = \begin{pmatrix} 18x_1 & 0 \\ 0 & 2 \end{pmatrix}$$

5. $H(A) = \begin{pmatrix} 18 & 0 \\ 0 & 2 \end{pmatrix} \qquad H(B) = \begin{pmatrix} -18 & 0 \\ 0 & 2 \end{pmatrix}$

6. Исследуем знакоопределенность матриц по критерию Сильвестра:

В точке A : $\Delta_1 = 18 > 0 \qquad \Delta_2 = 18 \cdot 2 - 0 = 36 > 0$, значит $H(A) > 0$, следовательно, A – локальный минимум

В точке B : $\Delta_1 = -18 < 0 \qquad \Delta_2 = -18 \cdot 2 - 0 = -36 < 0$, значит не выполняются достаточные условия экстремума.

Проверим необходимые условия экстремума 2-го порядка, вычислив собственные числа :

$$\begin{vmatrix} -18-\lambda & 0 \\ 0 & 2-\lambda \end{vmatrix} = 0 \quad \Rightarrow \quad (-18-\lambda)\cdot(2-\lambda) - 0 = 0 \quad \Rightarrow \quad \lambda_1 = -18 < 0 \qquad \lambda_2 = 2 > 0$$

значит матрица $H(B) <> 0$, следовательно в B экстремума нет.