Лекция №1. Безусловная оптимизация

Постановка задачи и определения

def: Методы оптимизации — это раздел математики, посвящённый решению (экстремальных) задач, то есть задач на нахождение минимумов и максимумов функций.

Rem: Задачу на нахождение максимума функции "f(x)"можно свести к задаче на нахождение минимума функции "$f_1(x) = -f(x)$" и наоборот.

Общий вид оптимизационной задачи

Найти экстремум(минимум или максимум) функции $f : X \to R$ определенной на некотором множестве $X \in R^n$ при ограничении $X \in D(D \subset X)$ то есть $f(x) \to extr, X \in D$(у функции есть экстремум на промежутке D).
В большинстве задач область определения функции "$f(x)$" $X = R^n$. Ограничение $X \in D$ записывается, как правило, в виде уравнений или неравенств . Если множество $D = X$, то имеет место задача без ограничений или задача безусловной оптимизации.
При решении оптимизационной задачи находятся как локальные, так и глобальные экстремумы функции.

def: Точка "x^*" является точкой локального минимума(максимума) функции "$f(x)$" if $\exists$ "ε" - окрестность $\mathcal{U}_\varepsilon = \{x : |x - x^*|\} < \varepsilon$. Точка $x^* : f(x^*) \le f(x)$ $(f(x^*) \ge f(x))$ $\forall x \in \mathcal{U}_\varepsilon$

Rem: то что пишется в скобках - для максимума, а то что без - для минимума

def: Точка "x^*" является точкой глобального минимума(максимума) функции "$f(x)$" if $f(x^*) \le f(x)$ $(f(x^*) \ge f(x))$ $\forall x \in D$

Необходимые и достаточные условия экстремума

def: Точка $X^0 = (x_1^0, ..., x_n^0)$ называется стационарной точкой дифференцируемой функции $f(x) = f(x_1...x_n)$, если в ней все частные производные равны нулю, то есть $f'(x^0) = 0$ или $\frac{\partial f(x^0)}{\partial x_1} = ... = \frac{\partial f(x^0)}{\partial x_n} = 0$

Необходимые условия экстремума I порядка

Теорема: if точка $x^* = (x_1^*, ..., x_n^*)$ - точка локального extr дифференцируемой в точке x^* функции $f(x_1, ...x_n) \Rightarrow then$ $\frac{\partial f(x^*)}{\partial x_1} = ... = \frac{\partial f(x^*)}{\partial x_n} = 0$ (1) (то есть - точка x^* - точка локального экстремума $\Rightarrow$ точка x^* - стационарная точка (обратное утверждение неверно))

Доказательство: Рассмотрим функцию одной переменной:
$\varphi(x_i) = f(x_1^*, ..., x_{i-1}^*, x_i^*, x_{i+1}^*, x_n^*)$ точка $x^* = (x_1^*, ..., x_n^*)$ - т. локального extr функции "f" $\Rightarrow$ x_i^* - т. локального extr функции "φ" $\Rightarrow$ по необходимому условию для функции одной переменной (по т. Ферма) 've:

$$\varphi(x_i^*) = 0 \Leftrightarrow \frac{\partial f(x^*)}{\partial x_i} = 0 \text{ — что и требовалось доказать :)}$$

Для формулировки достаточных условий extr, позволяющих отобрать среди стационарных точек именно точки локального extr(среди стационарных точек могут быть также точки перегиба, седловые точки и т.д.), рассмотрим матрицу вторых производных функции - матрицу Гессе(гессиан):

$$A = f''(x^*) = (\frac{\partial^2 f(x^*)}{\partial x_i \partial x_j})_{i,j=\overline{1,n}} = (a_{i,j})_{i,j=\overline{1,n}} \text{ (от 1 до n)}$$

def: Матрица "A"называется неотрицательно определённой ($A \geq 0$), если $\forall h = (h_1, ..., h_n) \in R^n$ неотрицательной является квадратичная форма:

$$(A * h, h) = \Sigma_{i,j=1}^n a_{ij} h_i h_j \geq 0$$

def: Матрица "A"называется положительно определённой ($A > 0$), если $(A * h, h) > 0, \forall h \in R^n (h \neq 0)$

Необходимые и достаточные условия extr II порядка

Теорема: Пусть $f(x)$ - дважды дифференцируема в точке x^*. Необходимые условия условия extr:

if точка x^* - точка локального минимума (максимума) функции $f(x) \Rightarrow$ $f^{'}(x^*) = 0; (f^{''}(x^*) * h, h) \geq 0 \ ((f^{''}(x^*) * h, h) \leq 0) \forall h \in R^n$

Достаточные условия extr

$f^{'}(x*) = 0; (f^{''}(x^*) * h, h) >0 \ ((f^{''}(x^*) * h, h) <0) \forall h \in R^n (h \neq 0) \Rightarrow$ точка x^* - т. локального минимума (максимума) функции $f(x)$

Доказательство:

Для случая минимума (для максимума аналогично)

По формуле Тейлора 've:

$$f(x^* + h) = f(x^*) + (f^{'}(x^*), h) + \tfrac{1}{2}(f^{''}(x^*) * h, h) + r(h),$$
$$\text{где } r(h) = o(|h|^2).(*)$$

1) Необходимость:

Пусть точка x^* - точка локального минимума $\Rightarrow$ по необходимому условию I порядка $f^{'}(x^*) = 0$, а также $f(x^* + \lambda h) \geq f(x^*)$ (при достаточно малых "λ") $\Rightarrow$ из (*) get (g при малых "λ" и фиксированном "h"):

$$f(x^* + \lambda h) - f(x^*) = 0 + \tfrac{\lambda^2}{2}(f^{''}(x^*) * h, h) + r(\lambda * h) \geq 0| : \lambda^2$$
$$(\text{где } r(\lambda * h) = 0(|\lambda|^2)).\ \tfrac{1}{2}(f''(x^*) * h, h) + \tfrac{r(\lambda * h)}{\lambda^2} \geq 0$$

При $\lambda \to 0\ 've : (f''(x^*) * h) \geq 0 (\forall h \in R^n)$ $\Rightarrow$ необходимость доказана

2) Достаточность:

Можно показать, что в R^n имеет место эквивалентность условий:

$$(A * h, h) > 0 \forall h \in R^n (h \neq 0) \Leftrightarrow \exists \alpha > 0 : (A * h, h) \geq \alpha * |R|^2 \ (\forall h \in R^n)$$

Учитывая, что $f'(x^*) = 0$ и $(f''(x^*) * h, h) \geq \alpha * |h|^2$

По формуле Тейлора 've:

$f(x^* + h) - f(x^*) = 0 + \frac{1}{2}(f''(x^*) * h, h) + r(h) \geq \frac{\alpha}{2}|h|^2 + r(h) \geq 0))$, то есть

$f(x^* + h) \geq f(x^*) \Rightarrow$ точка x^* - точка локального extr функции $f(x) \Rightarrow$ достаточность доказана.

что и требовалось доказать

Rem: Для квадратичной функции
$Q(x) = \Sigma_{i,j=1}^{n} a_{ij} x_i x_j$ условие положительной(отрицательной) определённости матрицы $A = (a_{ij})_{i,j=1}^{n} > 0$ - это достаточное условие абсолютного минимума(максимума) $Q(x)$ в стационарной точке.

Теорема Вейерштрасса

Теорема:

Непрерывная функция $f : R^n \to R$ на непустом ограниченном замкнутом подмножестве(компакте) множества R^n достигает своих абсолютных минимума и максимума[или 1)в стационарной точке внутри; 2) в граничной точке] - без доказательства

Следствие:

if функция "$f(x)$" непрерывна на R^n и $\lim\limits_{|x|\to\infty} f(x) = \infty$
($\lim\limits_{|x|\to\infty} f(x) = -\infty$) $\Rightarrow$ then она достигается своего абсолютного минимума (максимума) на $\forall$ замкнутом подмножестве и R^n. (без доказательства).

Критерий Сильвестра

Rem: В необходимых и достаточных условиях extr-a II порядка use-ся знакоопределённость матрицы вторых производных (гессиана) $A = f''(x)$.

Знакоопределённость матрицы устанавливается с помощью критерия Сильвестра.

Теорема:

Пусть A - симметричная матрица

1) Матрица "A"положительно определена $(A > 0) \Leftrightarrow$ все её последовательные гл. миноры положительны, т.е. $A_{1...k} = \begin{vmatrix} a_{11} & a_{ik} \\ a_{k1} & a_{kk} \end{vmatrix} > 0 \ (k = \overline{1,n})$

2) Матрица "A"отрицательно определена $(A < 0) \Leftrightarrow$ все её последовательные главные миноры чередуют знак, начиная с отрицательного, т.е. $(-1)^k * A_k > 0 \ \ (\text{k} = \overline{1,n})$

3) Матрица "A"неотрицательно определена $(A \geq 0) \Leftrightarrow$ все её гл. миноры (необязательно только последовательные) неотрицательны, т.е. $A_{1...k} = \begin{vmatrix} a_{i_1 i_1} & a_{i_1 i_k} \\ a_{i_k i_1} & a_{i_k i_k} \end{vmatrix} \geq 0 \ (1 \geq i_1 \geq ... \geq i_k \geq n)(k = \overline{1,n})$

4) Матрица "A"неположительно определена $(A \leq 0) \Leftrightarrow$ все её последовательные главные миноры чередуют знак, начиная с неположительного, т.е. $(-1)^k * A_{i_1...i_k} \geq 0 \ \ (\text{k} = \overline{1,n})$

(Теорема без доказательства)

Rem:

1) Можно показать, что $A > 0 (A \geq 0) \Leftrightarrow \forall \lambda_i > 0 (\lambda \geq 0)$, где λ_i - собственные значения матрицы.

2)

2.1) $A_{1...k} > 0 \Leftrightarrow A_{i_1...i_k} > 0$

2.2) $A_{1...k} \geq 0 \nRightarrow A_{i_1...i_k} \geq 0$ (т.е. $\nRightarrow A \geq 0$)

ex:

$A = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} \Rightarrow$ последовательные главные миноры $A_1 = 0; A_{12} = \begin{vmatrix} 0 & 0 \\ 0 & -1 \end{vmatrix} = 0$, но A не является неотрицательно определённой, так как $(Ah, h) = ((0; -h), (h, h)) = -h^2 < 0 (\forall h \neq 0)$

Правило решения задачи безусловной оптимизации

1) Найти стационарные точки, то есть точки, удовлетворяющие необх. усл. extr I порядка

$$\begin{cases} \frac{\partial f}{\partial x_1} = 0 \\ \frac{\partial f}{\partial x_n} = 0 \end{cases}$$

2) Во всех стационарных точках "x^0" проверяем достаточное условие extr II порядка, то есть проверяем знаки последовательных главных миноров гессиана "$f''(x^0)$" :

2.1) if $A_{1..k} > 0$ (k от 1 до n) $\Rightarrow then\ x^0 \in locminf$;

2.2) if $(-1)^k * A_{1...k} > 0$ (при k от 1 до n) $\Rightarrow then\ x^0 \in locmaxf$;

3) Если достаточное условие extr II порядка не выполняется $\Rightarrow$, то проверяем в стационарной точке необходимое условие extr II порядка, то есть проверяем знаки главных миноров гессиана "$f''(x^0)$"

3.1) if гессиан $f''(x^0) \ngeq 0$ не является неотрицательным отрезком, то есть не выполняется условие $A_{i1...ik} \geq 0 \Rightarrow then\ x^0 \notin locminf$;

3.2) if гессиан $f''(x^0) \nleq 0$ не является неположительным отрезком, то есть не выполняется условие что все $(-1)^k A_{i1...ik} \geq 0 \Rightarrow then\ x^0 \notin locmaxf$;