Лекции по Методам Оптимизации

Авторы

2022 — 2023

Содержание

1 Лекция №3 **12**

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 $x^* : f(x^*) \leq f(x)$ $(f(x^*) \geq f(x))$ $\forall x \in \mathcal{U}_\varepsilon$

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 $f(x^*) \leq f(x)$ $(f(x^*) \geq f(x))$ $\forall x \in D$

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 $f(x) \Rightarrow \quad f^{'}(x^*) = 0; (f^{''}(x^*) * h, h) \geq 0 \ ((f^{''}(x^*) * h, h) \leq 0) \forall h \in R^n$

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:

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

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"):

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

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. $A_{1\ldots k} = \begin{vmatrix} a_{11} & a_{ik} \\ a_{k1} & a_{kk} \end{vmatrix} > 0 \ (k = \overline{1,n})$

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

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. $(-1)^k * A_{i_1...i_k} \geq 0$ (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) $\Rightarrow then\ x^0 \in locmin f$;

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

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 $(-1)^k A_{i1...ik} \geq 0 \Rightarrow then\ x^0 \notin locmaxf$;

Лекция №2

Условная оптимизация

Условная оптимизация с ограничениями-равенствами. Функция Лагранжа

def: Задачей условной оптимизации с ограничениями-равенствами называется следующая задача:

$f_1 \to extr$; $f_i(x) = 0 (i = \overline{1,m}), (m < n)$, где $f_i(x) : R^n \to R (i = \overline{0,m})$ и $\forall f_i(x)$ - дифференцируема.

Теорема 0.0.1. *необходимое условие экстремума I порядка.*

If т.$x^* = (x_1^*, ..., x_n^*) \in locextr f_0 \Rightarrow$ then $\exists$ вектор множителей Лагранжа $\lambda^* = (\lambda_1^*, ..., \lambda_m^*)$ $\in R^m$ (λ) $(\lambda^* \neq 0)$ такой, что для функции Лагранжа:

$$\mathcal{L}(x_1, ..., x_n; \lambda_1, ..., \lambda_m) = f_0(x1, ..., x_n) + \sum_{i=1}^{m} \lambda_i f_i(x_1, ..., x_n) \tag{1}$$

выполняется условия стационарности:

$$\begin{aligned} \frac{\partial \mathcal{L}(x^*, \lambda^*)}{\partial x_j} &= \frac{\partial f_1(x^*)}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial f_i(x^*)}{\partial x_j} = 0 \quad (j = \overline{1,n}) \\ \frac{\partial \mathcal{L}(x^*, \lambda^*)}{\partial \lambda_i} &= f_i(x^*) = 0 \quad (i = \overline{1,m}) \end{aligned} \tag{2}$$

т.е. 've систему $n+m$ уравнений для нахождения $n+m$ неизвестных $\{x_1^*, ..., x_n^*; \lambda_1^*, ..., \lambda_m^*\}$

Теорема 0.0.2. *необходимое условия экстремума II порядка.*

If т.$x^* = (x_1^*, ..., x_n^*) \in locmin f_0$ (условие регулярности) и векторы $f_1'(x^*), ..., f_m'(x^*)$ - линейно независимы $\Rightarrow$ then $\exists$ вектор множителей Лагранжа

$\lambda^* = (\lambda_1^*, ..., \lambda_m^*) \in R^m (\lambda^* \neq 0)$ такой, что для функции Лагранжа

$$\mathcal{L}(x,\lambda) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x)$$

выполняются условия:

1. стационарности (2)

2. неотрицательной определенности матрицы вторых производных:

   (3) $(\mathcal{L}''(x^*,\lambda^*)h,h) \geq 0 \ \forall h \in \{(f_i'(x^*),h) = 0(i=\overline{1,m})\}$

Rem: Для т.$x^* \in locmax f_0$ need $\mathcal{L}(x,\lambda) = -f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x)$

Теорема 0.0.3. *достаточное условие экстремума II порядка.*

If в т. $x^* = (x_1^*,...,x_n^*)$ векторы $f_1'(x^*),...,f_m'(x^*)$ - линейно независимы и $\exists$ вектор множителей Лагранжа $\lambda^* = (\lambda_1^*,...,\lambda_m^*) \in R^m (\lambda^* \neq 0)$ такой, что для функции Лагранжа

$$\mathcal{L}(x,\lambda) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x)$$

в т.x^* выполняются условия:

1. стационарности (2)

2. положительной определенности матрицы вторых производных:

   $(\mathcal{L}(x^*,\lambda^*)h,h) > 0 \ \forall h \in \{(f_i'(x^*),h) = 0(i=\overline{1,m}),(h \neq 0)\} \Rightarrow$

   $\Rightarrow$ then т. $x^* \in locmin f_0$

Rem: Для т.$x^* \in locmax f_0$ need $\mathcal{L}(x,\lambda) = -f_0(x) + \sum_{i=1}^{m} \lambda_i f_i$

Частный случай

$z = f(x,y) \to$ extr; при $\varphi(x,y) = 0 \Rightarrow$ функция Лагранжа 've вид:

$\mathcal{L}(x,y,\lambda) = f(x,y) + \lambda\varphi(x,y)$

Условия стационарности(необх. усл. I порядка)(2):

$$\begin{cases} \mathcal{L}'_x(x^*, y^*, \lambda^*) = f'_x(x^*, y^*) + \lambda\varphi'_x(x^*, y^*) = 0 \\ \mathcal{L}'_y(x^*, y^*, \lambda^*) = f'_y(x^*, y^*) + \lambda\varphi'_y(x^*, y^*) = 0 \\ \mathcal{L}'_\lambda(x^*, y^*, \lambda^*) = \varphi'(x^*, y^*) = 0 \end{cases}$$

$\Rightarrow$ система трех уравнений с тремя неизвестными

$\Rightarrow$ находим стационарные точки (x^*, y^*, λ^*)

Вычисляется в каждой из get-x стационарных точек (x^*, y^*, λ^*) определитель:

$$\Delta = - \begin{vmatrix} 0 & \varphi'_x & \varphi'_y \\ \varphi'_x & \mathcal{L}''_{xx} & \mathcal{L}''_{xy} \\ \varphi'_y & \mathcal{L}''_{yx} & \mathcal{L}''_{yy} \end{vmatrix}$$

if $\Delta > 0 \Rightarrow$ then т.$(x^*, y^*, \lambda^*) \in$ locmin z

if $\Delta < 0 \Rightarrow$ then т.$(x^*, y^*, \lambda^*) \in$ locmax z

Условная оптимизация с ограничениями-равенствами и неравенствами. Математическое программирование

def: Задачей условной оптимизации с ограничениями-равенствами и ограничениями-неравентсвами называется следующая задача:

(4)$f_0(x) \to min$; (5)$f_i(x) \leq 0$ $(i = \overline{1, p})$, (6)$f_i(x) = 0$ $(i = \overline{p+1, m})$,

где $f_i(x) : R^n \to R(i = \overline{0, m})$

def: Эта задача называется задачей математического программирования (ЗМП)

Rem:

1) Обычно в ЗМП присутствуют условия неотрицательности переменных $x_i \geq (i = \overline{1, n})$ эти условия записываются в системе неравенств (5) в виде: $-x_i \leq 0(i = \overline{1, n})$

2) Каждое ограничение-равенство (6) можно заменить двумя неравенствами:

$$f_i(x) = 0 \Leftrightarrow \begin{cases} f_i(x) \le 0 \\ -f_i(x) \le 0 \end{cases} \quad (i = \overline{p+1, m})$$

В силу этих замечаний ЗМП можно записать в виде:

$f_0(x) \to min$

$f_i(x) \le 0$ $(i = \overline{1, m})$ (8) где $x = (x_1, ..., x_n)$

Для ЗМП составляется функция Лагранжа

$$\mathcal{L}(x_1, ..., x_n; \lambda_1, ..., \lambda_m) = f_0(x1, ..., x_n) + \sum_{i=1}^{m} \lambda_i f_i(x_1, ..., x_n)$$

С помощью функции Лагранжа выписываются необходимые и достаточные условия экстремума тип (2)-(3). Однако, проверка выполнения этих условий становится еще более сложной. При этом требуется решать систему, вообще говоря, нелинейных уравнений и неравенств. Для этого применяются итерационные численные методы, формирующие последовательность точек, сходящуюся к точке экстремума. Однако, эта точка может оказаться точкой локального(а не глобального) экстремума. Это объясняется тем, что ЗМП в такой общей постановке, без каких-либо предположений относительной функций $f_i(x)$, является многоэкстремальной задачей. Не существует универсальных методов решения таких задач. Содержательная теория построена лишь для отдельных классов ЗМП, в частности, задач оптимизации выпуклых функций на выпуклом множестве решений систем ограничений-неравенств. Такие задачи, называемые задачами выпуклого программирования(ЗВП), являются, как будет показано ниже, одноэкстремальными задачами.

1 Лекция №3

Выпуклый анализ

Ввиду важности задачи выпуклого программирования (ЗВП) рассматриваемые основные понятия раздела математики, называемого выпуклым анализом:

1) Выпуклое множество 2) Выпуклая(вогнутая) функция и ее дифференциальные свойства

Выпуклые множества точек

def: Множество точек называется выпуклым, если оно вместе с $\forall$ двумя своими точками содержит весь отрезок, соединяющий эти точки **ex:** — Тут нужно иллюстрацию **ex:** круг, сектор, отрезок, прямая, полуплоскость, куб, пирамида

Теорема 1.0.1. *Пересечение $\forall$ числа выпуклых множеств-выпуклое множеств, too*

Доказательство. Для простоты - пересечение двух выпуклых множеств:
$\forall N, M \in (A \cap B)$. Множество A - выпуклое $\Rightarrow$ отрезок $MN \in A$. Аналогично
$MN \in B \Rightarrow MN \in (A \cap B) \Rightarrow (A \cap B)$ - выпуклое — тут нужно иллюстрацию ■

def: Внутренняя т. множества - $\exists$ окрестность этой точки: в ней - только точек $\in$ множеству.

def: Граничная точка множества - $\forall$ окрестность этой точки содержит как точки $\in$ множеству, так и точек $\notin$ множеству

def: Угловая точка множества - она не является внутренней ни для какого отрезка, целиком $\in$-го множеству.

ex: — Тут нужно иллюстрацию точка M - внутренняя, точка N граничная, точка A - угловая ($AP \in$ множеству целиком, но точка A - не внутренняя точка для AP; точка внутренняя точка для KL, но KL $\notin$ множеству целиком

def: Замкнутое множество точек - if оно включает все свои граничные точки.

def: Ограниченное множество точек - if $\exists$ шар конечного радиуса с центром в $\forall$ точке множества, который полностью содержит в себе данное множество.

Можно показать, что если фигура ограничена only прямыми или их отрезками, то она 've конечное число угловых точек. При криволинейности границ - бесконечное множество угловых точек.

def: Выпуклое замкнутое множество точек пространства(плоскости), имеющее конечное число угловых точек, называется выпуклым многогранником(многоугольником), если оно ограниченное, и выпуклой многогранной(многоугольной) областью, если оно неограниченное.

Rem: Для выпуклого многогранника(многоугольника) угловые точки $\equiv$ его вершинам.
Для невыпуклого - не обязательно

ex: — Тут нужна иллюстрация точка E - вершина, но не угловая точка, т.к. KL $\in$ множеству целиком и точка E - внутренняя точка для KL.

Rem: В ЗЛП часто число параметров объекта $n > 3 \Rightarrow$ имеем дело с гипермногогранниками в гиперпространстве с координатами $x_i(i = (\overline{1, n}, n > 3)$

Геометрический смысл решений СЛН и СЛУ

Теорема 1.0.2. *Множество всех решений линейного неравенства*

$$a_{11}x_1 + ... + a_{1n}x_n \leq b_1$$

- это одно из полупространств, на которые n-мерные гиперпространство делится гиплоскостью

$$a_{11}x_1 + ... + a_{1n}x_n = b_1$$

включая и эту гиперплоскость

Выпуклые множества в n-мерном пространстве. Свойства ЗЛП

Выпуклые множества в n-мерном пространстве

Рассмотрим в n-мерном пространстве k точек (векторов):
$x_1 = (x_1^{(1)}, ..., x_n^{(1)})$, ..., $x_k = (x_1^{(k)}, ..., x_n^{(k)})$

def: Точка(вектор) $X = (x_1, ..., x_n)$ называется линейной комбинацией точек(векторов) $X_1, ..., X_k$, если справедливо соотношение:

$$X = \alpha_1 X_1 + ... + \alpha_k X_k \tag{3}$$

где α_j - const $j = (\overline{1, k})$

def: Точка(вектор) X называется выпуклой линейной комбинацией точек(векторов) $X = \alpha_1 X_1 + ... + \alpha_k X_k$, если:

1. $$X = \alpha_1 X_1 + ... + \alpha_k X_k$$

2. $$\alpha_j \geq 0 j = (\overline{1, k}) \tag{4}$$

3. $$\sum_{j=1}^{k} \alpha_j = 1 \tag{5}$$

Очевидно, что в частном случае при k=2 выпуклой линейной комбинацией двух точек X_1 и X_2 является соединяющий их отрезок, т.к.

$$X = \alpha_1 X_1 + \alpha_2 X_2 = \left\{ \begin{array}{l} \alpha_1 + \alpha_2 = 1 \\ \alpha_2 = \alpha_1 - 1 \end{array} \right\} \alpha X_1 + (1 - \alpha_1) X_2$$

- уравнение точек $X \in [X_1, X_2]$ (см. аналитич. геометр.) — добавить иллюстрацию

Лекция №4

Выпуклые и вогнутые функции

Пусть функция "$f(x)$", def-на на множестве $M \subset R^n$ (т.е. $x = (x_1, ..., x_n) \in M$

def: График функции "$f(x)$" - это множество $\Gamma(f) \subset R^{n+1}$ состоящее из точек $(x, f(x) = (x_1, ..., x_n, f(x))$, где $x \in M$.

def: Надграфик функции "$f(x)$" - это множество $\Gamma(f) \subset R^{n+1}$, состоящее из точек $(x, x_{n+1}) = (x_1, ..., x_n, x_{n+1})$, где $x \in M$, $x_{n+1} \leqslant f(x)$

def: Функция "$f(x)$", заданная на множестве $m \subset R^n$ наз-ся выпуклой, если ее надграфик $\Gamma(f)$ является выпуклым множеством в R^{n+1}

Rem: Функция "$f(x)$" вогнута $\leftrightarrow$ функция -"$f(x)$" выпукла.

Кроме данного геометрического def-я выпуклых (вогнутых) функций, часто use-ся следующее аналитическое def-е: def: Функция "$f(x)$", заданная на множестве $M \subset R^n$ наз-ся выпуклой(вогнутой), если 1) M - выпуклое, 2) $\forall x_1 \in M, x_2 \in M$ и числа $t \in [0, 1]$ 've:

$f((1-t) * x_1 + tx_2) \leqslant (1-t) * f(x_1) + t * f(x_2)$ (1)

($\geqslant$ - для вогнутой функции)

Rem: 1) Можно показать, что геометрические и аналитические def- я эквивалентны; (Rem: $(1-t) * f(x_1) + t * f(x_2)$ - уравнение отрезка $[x_1, x_2](t \in [0, 1])$ при $x_1 \neq x_2$ и $t \in (0, 1)$)

2) Если (1) 've строгое неравенство $\Rightarrow$, то функция "$f(x)$" наз-ся строго выпуклой (стро-

го вогнутой).

3)

Свойства выпуклых (вогнутых) функций

1. $f(x) = const$ и $f(x) = ax + b$ всюду выпуклы (вогнуты).

2. *if* функции $f_i(x)(i = \overline{1, m})$, заданные на $M \subset R^n$, выпуклы $\Rightarrow$

then функция $f(x) = \sum_{i=1}^{m} \alpha_i f_i(x)$ выпукла (при $\forall \alpha_i \geqslant 0$)

3. *if* функция "$f(x)$", заданная на выпуклом множестве $M \subset R^n$, выпукла $\to$

then $\forall$"α" множество решений неравенства $f(x) \leqslant \alpha$, т.е. множество $M_\alpha = \{x \in M : f(x) \leqslant \alpha\}$, является выпуклым.

3.1 *if* функции $f_1(x), ..., f_m(x)$, заданные на выпуклом множестве $M \subset R^n$, выпуклы $\Rightarrow$

then множество решений системы неравенств $f_i(x) \leqslant \alpha_i$ $(i = \overline{1, m})$ является выпуклым.

4. Выпуклая (вогнутая) функция, заданная на выпуклом множестве $M \subset R^n$, непрерывна в $\forall$ внутренней точке множества.

(вставка на страницу 25)

def: Выпуклая оболочка множества - это совокупность всех выпуклых линейных комбинаций его конечных подмножеств $\{x_1, ..., x_k$ (где $x_i \in M$). (Конечное подмножество - это конечный набор точек $x_1, ..., x_k$)

Теорема (Крейна-Мильмана)

Выпуклый компакт в нормированном пространстве является выпуклой оболочкой своих угловых (крайних) точек.

Теорема

Пусть выпуклая (вогнутая) функция "$f(x)$" задана на выпуклом множестве $M \subset R^n \Rightarrow$ Каждый локальный минимум (максимум) функции "$f(x)$" является её глобальным минимумом (максимумом) на множестве M.

Доказательство (для выпуклой функции)

Пусть точка $x^* \in M$ - точка локального min-a. Пусть точка точка $x \in M$ - произвольная точка множества M. Need доказать: $f(x) \geqslant f(x^*)$.

Отрезок $[x^*, x] = (1-t_0)x^* + tx$ $(t \in [0,1])$ принадлежит "M". При малом значении $t_0 \in (0,1)$ /-щая точка отрезка $[x^*, x]$ находится в малой окресности т.x^*, в котором имеем:

$f((1-t_0)x^* + t_0 x) \geqslant f(x^*)$

Из def-я выпуклой функции $f(x)$ 've:

$f((1-t_0)x^* + t_0 x) \leqslant (1-t_0)f(x^*) + t_0 f(x)$

Т.о. :

$(1-t_0) * f(x^*) + t_0 * f(x) \geqslant f(x^*) \Rightarrow$

$\Rightarrow f(x) \geqslant f(x^*)$ ч. и т.д. (для вогнутой "$f(x)$" доказательство аналогичное.)

Rem:

1) Задачи выпуклой оптимизации называются одноэкстремальными. В многоэкстремальных задачах может $\exists$-ть локальные экстремумы, не совпадающие в глобальными.

2) Одноэкстремальность задач выпуклой оптимизации не означает, что каждая такая задача имеет решение и при том единственное. f.e.: 1) Выпуклая функция одной переменной $f(x) = x$, при $x \in (0,1)$ не достигает min-a (и max-a) на $(0,1)$; 2) Множество точек min-a выпуклой фукнции $f(x) = C$ - const, $x \in M$, совпадает со всем M.

3) Если функция "$f(x)$" - строго выпуклая (строго вогнутая), то разрешимая задача выпуклой оптимизации (т.е. множество $M \neq \varnothing$ и ограничено) имеет единственное решение.

def: Производной $\frac{\partial f(x)}{\partial \ell}$ функции $f(x) = f(x_1, ..., x_n)$ по направлению ненулевого единичного вектора $\ell = (\ell_1, ..., \ell_n)$ в т. $x = (x_1, ..., x_n)$ называется предел

$$\frac{\partial f(x)}{\partial \ell} = \lim_{\lambda \to +0} \frac{f(x + \lambda \ell) - f(x)}{\lambda}$$

Если функция дифференциируема в т. x, то она в этой точке производную по $\forall$ направлению $\ell = (\ell_1, ..., \ell_n)$, которая выражается через частные производные по следующей формуле:

$$\frac{\partial f(x)}{\partial \ell} = \sum_{i=1}^{n} \frac{\partial f(x)}{\partial x_i} \ell_i$$

Производная по направлению равна скалярному произведению вектора "ℓ" и вектора градиента функции "$f(x)$" в т. x

$$\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, ..., \frac{\partial f x)}{\partial f(x_n)} \right) : \frac{\partial f(x)}{\partial \ell} = \nabla f(x) * \ell = \nabla f(x) * |\ell|$$

или

$$\frac{\partial f(x)}{\partial \ell} = |f(x)| * |\ell| * \cos\phi$$

Т.о.:

$\forall$ направления "ℓ" производная $\frac{\partial f}{\partial \ell} \leqslant |\nabla f(x)|$.

$\Rightarrow$ Вывод:

1) Производная по направлению "$\frac{\partial f}{\partial \ell}$" - это скорость изменения функции "$f(x)$" по направлению "ℓ"(знак "$\frac{\partial f}{\partial \ell}$" - это характер изменения функции(возрастание или убывание)).

2) Направление градиента $\nabla f(x)$ - это направление наибольшего возрастания функции "$f(x)$" в т. x; длина градиента $\nabla f(x)|$" равна наибольшей скорости возрастания функции в этой точке.

Теорема

Пусть 've дифференциируемую выпуклую функцию "$f(x)$", def-ую на выпуклом множестве $M \subset R^n$.

1) $\forall x, y \in M$ 've: $\lambda f(x) * (y - x) \leqslant f(y) - f(x)$

2) т. $x^* \in absminf \Leftrightarrow \nabla f(x^*) = 0$ (2)

Теорема (дифференциальный признак выпуклых функций)

Дважды дифференциируемая функция "$f(x)$", def-ная на выпуклом множестве $M \subset R^n$

является выпуклой $\Leftrightarrow \forall x = (x_1, ..., x_n) \in M$ и $\forall \ell = (ell_1, ..., \ell_n) \in R^n$

've:

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f(x)}{\partial x_i \partial x_j} \ell_i \ell_j \geqslant 0, (3)$$

т.е. гессиан функции всюду неотрицательно определён:

$$A = f''(x) = \left(\frac{\partial^2 f(x)}{\partial x_i \partial x_j} \right) \geqslant 0, (\forall x \in M)$$

Rem:

1) Функция является строго выпуклой $\Leftrightarrow$ гессиан положительно определён, т.е.

$$\left(\frac{\partial^2 f(x)}{\partial x_i \partial x_j} \right) > 0, (\forall x \in M)$$

2) Знакоопределённость гессиана устанавливается с помощью критерия Сильвестра.

ex 1:

Проверить выпуклость функции:

$f(x_1, x_2) = 4x_1 + x_2^2 - 2x_1x_2 + 6x_1 - 5x_2 - 2$

$f'_{x1} = 8x_1 = 2x_2 + 6$; $f''x_1x_1 = 8$; $f''_{x_1x_2} = -2$; $f''x_2x_2 = 2$; $f'_{x2} = 2x_2 - 2x_1 - 5$

Гессиан A $= \begin{pmatrix} 8 & -2 \\ -2 & 2 \end{pmatrix}$; $\Delta_1 = 8 > 0; \Delta_2 = 12 > 0$

$\Rightarrow$ функция является строго выпуклой.

ex 2:

Проверить выпуклость функции: $f(x) = -\sqrt{x_1x_2}$ на множестве

$M = \{(x_1, x_2) | x_1 > 0, x_2 > 0\}$.

$f'_{x1} = -\frac{1 * x_2}{2\sqrt{x_1x_2}} = -\frac{1}{2}\sqrt{\frac{x_2}{x_1}}$;

$f'_{x2} = -\frac{1}{2}\sqrt{\frac{x_1}{x_2}}$;

$f''x_1x_1 = (-\frac{\sqrt{x_2}}{2} * x_1^{-\frac{1}{2}})' = \frac{1}{4}\frac{\sqrt{x_2}}{\sqrt{x_1^3}} = \frac{1}{4x_1}\sqrt{\frac{x_2}{x_1}};$

$f''_{x_1x_2} = -(\frac{1}{2\sqrt{x_1}} * x_2^{\frac{1}{2}})' = \frac{1}{4\sqrt{x_1x_2}};$

$f''x_2x_2 = (-\frac{1*\sqrt{x_1}}{2} * x_2^{-\frac{1}{2}})' = \frac{1\sqrt{x_1}}{4\sqrt{x_2^3}} = \frac{1}{4x_2} * \sqrt{\frac{x_1}{x_2}}$

Гессиан H $= \begin{pmatrix} \frac{1}{4x_1} * \sqrt{\frac{x_2}{x_1}} & -\frac{1}{4\sqrt{x_1x_2}} \\ -\frac{1}{4\sqrt{x_1x_2}} & \frac{1}{4x_1} * \sqrt{\frac{x_1}{x_2}} \end{pmatrix}$

$\Delta_1 = \frac{1}{4x_1}\sqrt{\frac{x_2}{x_1}} > 0(\forall x_1, x_2 > 0);$

$\Delta_2 = \frac{1}{16x_1x_2}\sqrt{\frac{x_2x_1}{x_1x_2}} - \frac{1}{16x_1x_2} = 0 \Rightarrow$

Гессиан H $\geqslant 0 \Rightarrow$ функция является на множестве $M = \{(x_1, x_2)|x_1 > 0, x_2 > 0\}$ выпуклой (но не является строго выпуклой).

Лекция №5. Общая постановка выпуклого программирования(ЗВП).Теорем Куна-Таккера

Общая постановка ЗВП

ЗВП называется следующей задачей математического программирования(ЗМП)

$f(x_1, ..., x_n) \to min$

при ограничениях:

$$\begin{cases} g_1(x_1, ..., x_n) \leq 0 \\ g_m(x_1, ..., x_n) \leq 0 \end{cases},$$

где $f, g_1, ..., g_m$ - выпуклые функции, def-ные на некотором выпуклом множестве $M \subset R^n$.

При этом множество M содержит допустимую область значений, то есть множество, удовлетворяющее системе ограничений. Заметим, что свойству 3.1 выпуклых функций допустимая область ЗВП также выпукла

Rem:

1) Аналогично формулируется задача максимизации вогнутой функции "f" при вогнутых функциях "$g_1, ..., g_m$", def-х на некотором выпуклом множестве $M \subset R^n$. При этом знак неравенств — " $\geq$ ".

Аналогично допустимая область ЗВП будет также выпуклым множеством(свойство для вогнутых функций)

2) В ??? с теор. на с.??? в ЗВП каждый локальный минимум является глобальным минимумом функции "f" в допустимой области. Причём, если функция "f" строго выпуклая и ограниченная снизу на ограниченном непустом множестве M, то ЗВП 've единственное решение, то есть минимумом функции "f" достигается в одной точке

$$x^0 = (x_1^0, ..., x_n^0)$$

(см с.???)

Теорема Куна-Таккера.

Теорема: Пусть на ЗВП налагаются следующие требования:

1. Множество, удовлетворяющее системе строгих неравенств:

$$\begin{cases} g_1(x_1,...,x_n) < 0 \\ g_m(x_1,...,x_n) < 0 \end{cases}$$

и называемое внутренней частью допустимой области, не пусто (т.н. условие Слейтера)

2. Часть допустимой области, в которой некоторые ограничения обращаются в равенства, называется границей допустимой области, а эти ограничения - активными. Пусть градиенты активных ограничений в отвечающих им точках границы линейно независимы(т.н. условие регулярности (см с.???))

Теорема Куна-Таккера (необходимые и достаточные условия решения (глобального минимума) ЗВП)

Точка $x^0 = (x_1^0, ..., x_n^0)$ является решением ЗВП (то есть точкой глобального минимума функции "f") $\Leftrightarrow$(Н. и Д.) в ней выполнены следующие условия(Условия Куна-Таккера):

1)Условие стационарности функции Лагранжа (см.с.???)

$$\lambda(x_1,...,x_n;\lambda_1,...,\lambda_m) = f(x_1,...x_n)\Sigma_{i=1}^m) * \lambda_i * g_i * (x_1,...,x_n)$$

$$\frac{\partial L}{\partial x_1} = ... = \frac{\partial L}{\partial x_n} = 0 \text{ (то есть } \nabla(x^0) = 0);$$

2)Условия "дополняющей нежёсткости":

$$\lambda_i * g_i * (x_1^0, ..., x_m^0) = 0 \ (i = \overline{1,m});$$

3)Условия неотрицательности:

$$\lambda_i \geq 0 \ (i = \overline{1,m})$$

Доказательство:

Заметим, что в условиях К.-Т. множители Лагранжа $\lambda_1, ..., \lambda_m$ являются "выключателями"делящих ограничений. Если, например, $\lambda_k = 0$, то ограничение $g_k \leq 0$ исключается из условий К.-Т., так как : 1) это ограничение не входит в функцию Лагранжа (слагаемое $\lambda_k * g_k = 0$) и,??? не входит в условие 1; 2) условия 2 и 3 для этого ограничения выполняются автоматически

1. Необходимость

Пусть точка $x^0 = (x_1^0, ..., x_n^0)$ - решение ЗВП, то есть $min f(x_1, ..., x_n) = f(x_1^0, ...x_n^0)$. Целевая функция "$f$" может достигать минимума внутри допустимой области или на её границе.

Пусть минимум достигается внутри области $\Rightarrow$ в точке x^0 выполняются необходимые условия локального(безусловного экстремума функции "f"):

$$\frac{\partial f}{\partial x_1} = ... = \frac{\partial f}{\partial x_n}$$

Это частный случай условия 1 К.-Т.(условия стационарности функции Лагранжа) при $\lambda_1 = ... = \lambda_m = 0$ все ограничения выключены, так как 've безусловную оптимизацию)

Условия 2 и 3 при этом (при $\lambda_i = 0 (i = \overline{1, m})$ выполняются автоматически.

Пусть минимум достигается на границе допустимой области, то есть является условным extr функции "f" с активными ограничениями задачи для точки минимума точки x^0. Необходимые условия условного extr (см. с. ???) ($\frac{\partial L}{\partial x_j} = 0; (\frac{\partial L}{\partial \lambda_j} = 0 (j = \overline{1, n}; i = \overline{1, m}))$ влекут за собой выполнение условий 1 и 2 К.-Т., если положительно равынми нулю множит. Лагранжа для неактивных ограничений (так как реально они не входят в функцию Лагранжа). Остаётся доказать, что выполняется 3 условие К.-Т. : $\lambda_i \geq 0 (i = \overline{1, m})$

Для наглядности ограничимся случаем двух переменных (n=2).

Уравнение $g(x_1, x_2)$ означает плоскую кривую, а неравенство $g(x_1, x_2) \leq 0$ - одну из частей плоскости, ограниченную этой кривой

Система ограничений

$$\begin{cases} g_1(x_1, x2) \leq 0 \\ \dots \\ g_m(x_1, x_2) \leq 0 \end{cases}$$

задает криволинейный многоугольник, в каждой вершине которого пересекаются две кривые.

Минимум на границе допустимой области достигается либо на строке этого многоугольника, либо в его вершине.

Пусть минимум - на стороне многоугольника $\Rightarrow$ активным является только одно ограничение и только отвечающий ему множитель Лагранжа $\lambda_k \neq 0$. Функция Лагранжа 've вид:

$$L = f(x_1, ...x_n) + \lambda_k * g_k(x_1, ..., x_n),$$

где $g_k = 0$ - активное ограничение.

В точке минимума выполняются необходимые условия. Условия условного extr(см. с. ???):

$$\frac{\partial L}{\partial x_1} = ... \frac{\partial L}{\partial x_n} = \frac{\partial L}{\partial \lambda_k}$$

или

$$\begin{cases} \nabla L = \nabla f + \lambda_k * \nabla g_k = 0 \\ g_k = 0, \end{cases}$$

$\Rightarrow \nabla f = -\lambda_k * \nabla g_k, (*)$,

то есть векторы ∇f и ∇g_k — коллинеарны

Функция $g_k < 0$ внутри допустимой области и $g_k > 0$ вне её $\Rightarrow$ вектор ∇g_k направлен из допустимой области (в сторону возрастания функции "g_k")

Значение функции "f" внутри допустимой области больше, чем в точке минимума $\Rightarrow \nabla f$ направлен внутрь допустимой области.

Таким образом, векторы ∇f и ∇g_k противонаправлены и в равенстве (*) коэффиценты коллинеарности $\lambda_k \leq 0 \Rightarrow \lambda_k \geq 0$, то есть условие 3 К.-Т. доказано для случая, когда минимум достигается на стороне допустимого многоугольника.

Можно показать, что условие 3 К.-Т. ($\lambda_i \geq 0 (i = \overline{1, m})$) выполняется и для случая, когда минимум достигается в вершине допустимого многоугольинка.

Таким образом, необходимость К.-Т. доказана.

Лекция №6

Доказательство теоремы Куна-Таккера
2. Достаточность

Пусть в точке $x^0 = (x_1^0, ..., x_n^0)$ выполняются условия 1-3 теоремы Куна-Таккера.

Имеем выпуклые функции "$f, g_1, ..., g_n$" и выпуклую допустимую область $\Rightarrow$

$\Rightarrow$ Функция Лагранжа

$L = f + \lambda_1 g_1 + ... + \lambda_m g_m$ $(\lambda_1 \geqslant 0, ..., \lambda_m \geqslant 0)$ также выпукла (см. св-во 2 выпуклых ф-ий)

Если в точке $x^0 == (x_1^0, ..., x_n^0)$ выполняется условие 1 теоремы Куна-Таккера $\frac{\partial L}{\partial x_1} = ... = \frac{\partial L}{\partial x_n} = 0$, (т.е. $\nabla L(x^0) = 0$) то эта точка - точка min выпуклой функции Лагранжа "L"в допустимой области.

Обозначим $minL(x) = L(x^0) = L^0 \Rightarrow L^0 \leqslant L$

Но вследствие условия 2 теоремы Куна-Таккера

$\lambda_1 g_1^0 = ... = \lambda_m g_m^0 = 0 \Rightarrow L^0 = f^0, f^0 \leqslant L$

Кроме того, в допустимой области $g_1 \leqslant 0, ..., g_m \leqslant 0$ и по условию 3 теоремы Куна-Таккера, $\lambda_1 \geqslant 0, ..., \lambda_m \leqslant 0 \Rightarrow \lambda_1 g_1 = ... = \lambda_m g_m \leqslant 0 \Rightarrow$

в допустимой области

$L = f + \lambda_1 g_1 + ... + \lambda_m g_m \leqslant f \Rightarrow$ в допустимой области:

$f^0 \leqslant L \leqslant f$, т.е. $f^0 \leqslant f \Rightarrow minf = f^0 = f(x^0)$, т.е. точка $x^0 = (x_1^0, ..., x_n^0)$ - точка минимума (глобального) целевой функции в допустимой области. Т. О. достаточность теоремы Куна-Таккера доказана.

ч и т.д.

Теорема Куна-Таккера (в "седловом"варианте)

Точка $x^0 = (x_1^0, ..., x_n^0)$ является решением ЗВП (т.е. точкой глобального min-а функции "f") $\Leftrightarrow \exists$ неотрицательный вектор множителей Лагранжа

$\lambda^0 = (\lambda_1^0, ..., \lambda_m^0)(\lambda_i^0 \geqslant 0; i = \overline{1, m})$ такой, что для функции Лагранжа

$L = (x_1, ..., x_n; \lambda_1, ..., \lambda_m) = f(x_1, ..., x_n) + \Sigma_{i=1}^{m} \lambda_i g_i(x_1, ..., x_n)$ точка $(x^0; \lambda^0)$ является седловой точкой, т.е.

$L(x^0, \lambda) \leqslant (x^0, \lambda^0) \leqslant L(x, \lambda^0)$ (*)

$\forall x \in$ допустимой области, $\lambda \geqslant 0$

(без доказательства)

Rem:

1. Из неравенства (*) следует, что точка - точка минимума функции, а точка - точка максимума функции по. Существование седловой точки означает равенство минимакса максимуму.
2. Можно показать, что необходимые и достаточные условия того, чтобы являлась седловой точкой функции Лагранжа, имеют вид:

Rem: 1. При доказательстве необходимости условий теоремы Куна-Таккера выпуклость функций и допустимой области не требовалась $\Rightarrow$ условия Куна-Таккера являются необходимыми для произвольной (необязательно выпуклой) ЗМП.
2. Множители Лагранжа "λ_k" представляют собой "цену"ограничения "$g_k \leqslant 0$", т.е. чувствительность оптимального значения целевой функции "f^0" к нарушению этого ограничения. В функции Лагранжа слагаемые "$\lambda_k g_k$" представляет собой как бы "штраф"за нарушение ограничения "$g_k \leqslant 0$". Чем больше значение множителя Лагранжа "λ_k" и чем больше нарушено /-щее ограничение (т.е. стало "$g_k > 0$"), тем больше величина этого "штрафа". На этой идее основан один из методов решения ЗМП - метод штрафных функций, который был рассмотрен ниже.

3. Условия теоремы Куна-Таккера представляют собой систему нелинейных (в общем случае) уравнений и неравенств, которая допускает лишь приближенное решение

численными методами. Для решения ЗВП разработали специальные итерационные численные методы, которые заключаются в построении последовательных приближений к точке целевой функции в допустимой области. Критерием окончания итерационного процесса является достижение заданной точности вычислений. В результате определяются приближенные значения.

4. Точное решение ЗВП допускает в случае, когда и целевая функция и все ограничения являются линейными функциями переменных (линейные функции являются выпуклыми в /-ии со свойством 1 выпуклых функций). Такая ЗВП называется задачей линейного программирования (ЗЛП).

Линейное программирование