Условная оптимизация

Классификация задач на условный экстремум

<u>Задачи на условный экстремум</u>

Классическая задача		Задача математического программирования	
ЦФ	$f(\vec{x}) \to \min$	ЦФ	$f(\vec{x}) \to \min$
СО	$g_i(\vec{x}) = 0 \quad i = \overline{1,m}$	СО	$\begin{cases} g_i(\vec{x}) \le 0 & i = \overline{1,k} \\ g_i(\vec{x}) = 0 & i = \overline{k+1,m} \\ \vec{x} \in P \end{cases}$

СО: ФО и ПО

КЗнаУЭ – частный случай ЗМП

Особенности решения задач условной оптимизации:

1. Оптимальное решение – локальный экстремум внутри ОДР
2. Оптимальное решение – экстремум на границе ОДР

Классическая задача на условный экстремум

$$Z = f(\bar{x}) = f(x_1, x_2, \ldots, x_n) \to \min \qquad \begin{cases} g_1(\bar{x}) = 0 \\ g_2(\bar{x}) = 0 \\ \ldots \\ g_m(\bar{x}) = 0 \end{cases}$$

<u>Функция Лагранжа</u> $L(\bar{x}, \bar{\lambda}) = f(\bar{x}) + \sum_{i=1}^{m} \lambda_i g_i(\bar{x})$, где λ_i $(i = \overline{1, m})$ - множители Лагранжа

<u>Необходимые условия минимума</u> $f(\bar{x})$ при наличии ограничений:

$$\begin{cases} \dfrac{\partial L(\bar{x}, \bar{\lambda})}{\partial x_j} = \dfrac{\partial f(\bar{x})}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \dfrac{\partial g_i(\bar{x})}{\partial x_j} = 0 & j = \overline{1, n} \\ \dfrac{\partial L(\bar{x}, \bar{\lambda})}{\partial \lambda_i} = g_i(x) = 0 & i = \overline{1, m} \end{cases}$$

Решение $(\bar{x}^*, \bar{\lambda}^*)$ - стационарная точка системы

$$L(\bar{x}, \bar{\lambda}) = f(\bar{x}) + \sum_{i=1}^{m} \lambda_i g_i(\bar{x}) = f(\bar{x}) \text{ для } \forall \bar{x} \in \text{ОДР}$$

<u>Достаточное условие минимума</u> $f(\bar{x})$ в $(\bar{x}^*, \bar{\lambda}^*)$

$$\frac{\partial^2 L(\bar{x}^*, \bar{\lambda}^*)}{\partial^2 x_j} > 0 \quad j = \overline{1, n}, \qquad \text{где } \frac{\partial^2 L(\bar{x}^*, \bar{\lambda}^*)}{\partial x_j \partial x_k} = \frac{\partial^2 f(\bar{x}^*)}{\partial x_j \partial x_k} + \sum_{i=1}^{m} \lambda_i^* \frac{\partial^2 g_i(\bar{x}^*)}{\partial x_j \partial x_k}$$

Пример решения классической задачи на условный экстремум

Пример: $f(x_1, x_2) = x_1^2 + x_2^2 \to \min$ при наличии ограничения $x_1 + x_2 = 4$

Функция Лагранжа $L(x_1, x_2, \lambda) = f(x_1, x_2) + \lambda g_1(x_1, x_2) = x_1^2 + x_2^2 + \lambda(4 - x_1 - x_2)$

Необходимые условия минимума $\begin{cases} \dfrac{\partial L(x_1, x_2, \lambda)}{\partial x_1} = 2x_1 - \lambda = 0 \\ \dfrac{\partial L(x_1, x_2, \lambda)}{\partial x_2} = 2x_2 - \lambda = 0 \\ \dfrac{\partial L(x_1, x_2, \lambda)}{\partial \lambda} = 4 - x_1 - x_2 = 0 \end{cases}$

Стационарная точка системы $x_1 = 2$, $x_2 = 2$ и $\lambda = 4$

Достаточное условие минимума $G = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$

Ответ: точка $(2,2)$ минимум $f(x_1, x_2) = x_1^2 + x_2^2 = 8$ с учетом выполнения ограничения $x_1 + x_2 = 2 + 2 = 4$

Задача математического программирования

Распространение метода множителей Лагранжа на ЗМП вида:

$$Z = f(\vec{x}) = f(x_1, x_2, \ldots, x_n) \rightarrow \min \qquad g_i(\vec{x}) \le b_i \quad i = \overline{1, m}$$

Сведение неравенств к равенствам

$$g_i(\vec{x}) + u_i^2 = b_i \quad u_i^2 > 0 \quad \text{т.е } g_i(\vec{x}) + u_i^2 - b_i = 0$$

Сведение к классической задаче на условный экстремум

$$Z = f(\vec{x}) = f(x_1, x_2, \ldots, x_n) \rightarrow \min \qquad g_i(\vec{x}) + u_i^2 - b_i = 0 \quad i = \overline{1, m}$$

Функция Лагранжа $\quad L(\vec{x}, \vec{\lambda}, \vec{u}) = f(\vec{x}) + \sum_{i=1}^{m} \lambda_i \left[g_i(\vec{x}) + u_i^2 - b_i \right]$

Необходимые условия минимума $f(\vec{x})$ (в стационарной точке $(\vec{x}^*, \vec{\lambda}^*, \vec{u}^*)$)

$$\begin{cases} \dfrac{\partial L(\vec{x}, \vec{\lambda}, \vec{u})}{\partial x_j} = \dfrac{\partial f(\vec{x})}{\partial x_j} + \sum\limits_{i=1}^{m} \lambda_i \dfrac{\partial g_i(\vec{x})}{\partial x_j} = 0 & j = \overline{1, n} \\ \dfrac{\partial L(\vec{x}, \vec{\lambda}, \vec{u})}{\partial \lambda_i} = g_i(\vec{x}) + u_i^2 - b_i = 0 & i = \overline{1, m} \\ \dfrac{\partial L(\vec{x}, \vec{\lambda}, \vec{u})}{\partial u_i} = 2\lambda_i u_i = 0 & i = \overline{1, m} \end{cases}$$

Задача математического программирования

$$\frac{\partial L(\bar{x},\bar{\lambda},\bar{u})}{\partial u_i} = 2\lambda_i u_i = 0 \quad \Big| \times \frac{u_i}{2} \quad \rightarrow \quad \lambda_i u_i^2 = \lambda_i\left[b_i - g_i(\bar{x})\right] = 0 \quad i = \overline{1,m}$$

т.е. $\lambda_i = 0$ или $b_i - g_i(\bar{x}) = 0$

Возможные варианты:

1. $\lambda_i \neq 0$ и $g_i(\bar{x}) = b_i$ - активное ограничение-равенство
2. $\lambda_i = 0$ и $g_i(\bar{x}) < b_i$ - неактивное ограничение (можно пренебречь)

Дополнительное условие в точке минимума $\lambda_i \geq 0$

<u>Условия Куна-Такера</u> – необходимые условия минимума $f(\bar{x})$ при наличии ограничений $g_i(\bar{x}) \leq b_i \quad i = \overline{1,m}$ – поиск $\bar{\lambda}^*$ и $\bar{x}^*$

$$\begin{cases} \dfrac{\partial f(\bar{x})}{\partial x_j} + \sum\limits_{i=1}^{m} \lambda_i \dfrac{\partial g_i(\bar{x})}{\partial x_j} = 0 & j = \overline{1,n} \\ g_i(\bar{x}) \leq b_i & i = \overline{1,m} \\ \lambda_i\left[g_i(\bar{x}) - b_i\right] = 0 & i = \overline{1,m} \\ \lambda_i \geq 0 & i = \overline{1,m} \end{cases}$$

В задаче на максимум λ_i меняет знак на противоположный

Пример на ЗМП

Пример: $f(\vec{x}) = 3x_1^2 + 4x_1x_2 + 5x_2^2 \to \min$

СО $x_1 \ge 0;\ \ x_2 \ge 0;\ \ x_1 + x_2 \ge 4$

Преобразование:

$$f(\vec{x}) = 3x_1^2 + 4x_1x_2 + 5x_2^2 \to \min$$

$$-x_1 + u_1^2 = 0;\ -x_2 + u_2^2 = 0;\ \ -x_1 - x_2 + u_3^2 + 4 = 0$$

Функция Лагранжа:

$$L(\vec{x}, \vec{\lambda}, \vec{u}) = 3x_1^2 + 4x_1x_2 + 5x_2^2 + \lambda_1(u_1^2 - x_1) + \\ + \lambda_2(u_2^2 - x_2) + \lambda_3(u_3^2 - x_1 - x_2 + 4)$$

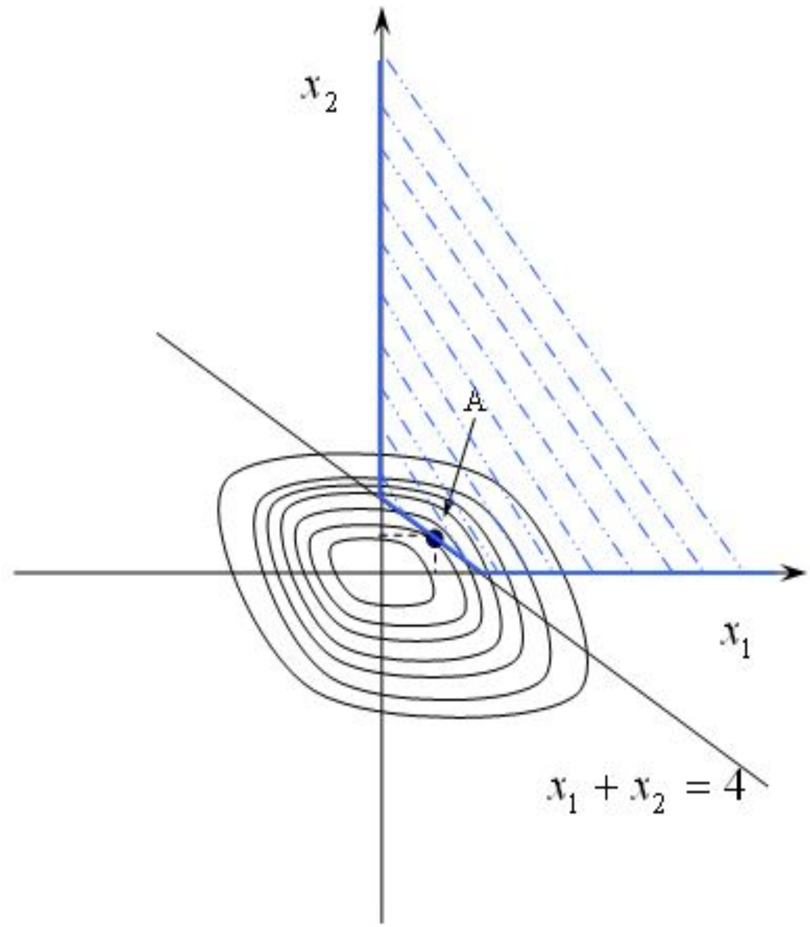
Условия Куна-Такера:

$$\begin{cases} 6x_1 + 4x_2 - \lambda_1 - \lambda_3 = 0 \\ 4x_1 + 10x_2 - \lambda_2 - \lambda_3 = 0 \\ -x_1 \le 0;\ -x_2 \le 0;\ -x_1 - x_2 \le -4; \\ \lambda_1 x_1 = 0;\ \lambda_2 x_2 = 0;\ \lambda_3(4 - x_1 - x_2) = 0 \\ \lambda_1 \ge 0;\ \ \lambda_2 \ge 0;\ \ \lambda_3 \ge 0 \end{cases}$$

Решение:

$x_1 = 3;\ \ x_2 = 1;\ \ \lambda_1 = 0;\ \lambda_2 = 0;\ \ \lambda_3 = 22$

$f(x_1, x_2) = 44$



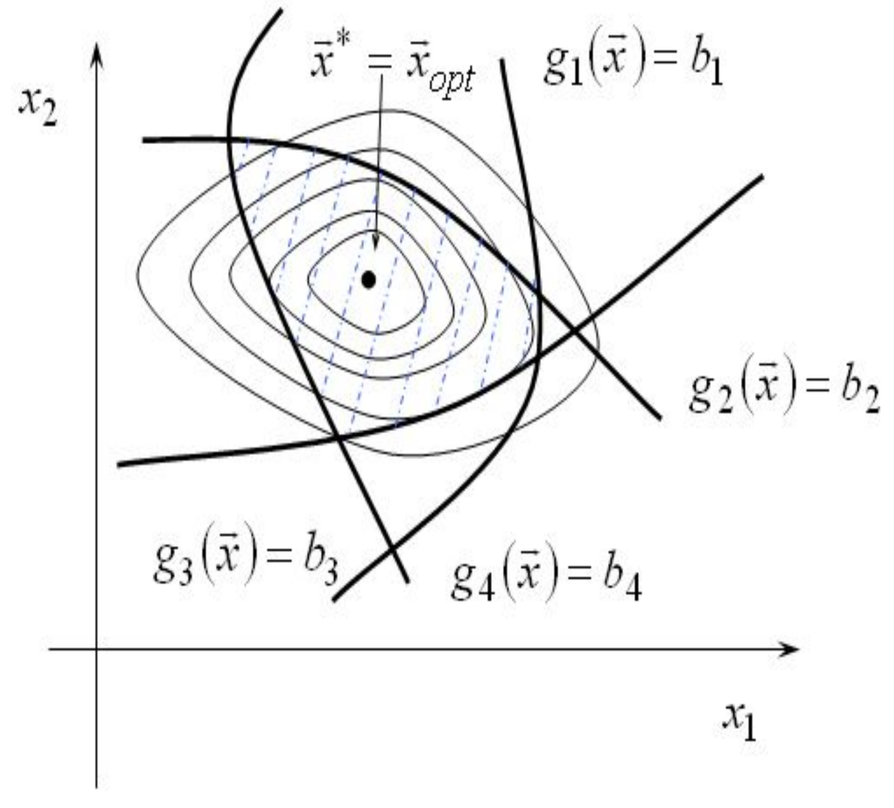
Возможные варианты положения минимума в ЗМП

I

$g_i(\vec{x}) = b_i$ - границы ОДР

минимум $f(\vec{x})$ при наличии ограничений совпадает с минимумом $f(\vec{x})$ при отсутствии ограничений

ограничениями можно пренебречь и для поиска минимума использовать методы минимизации $f(\vec{x})$ без ограничений



Возможные варианты положения минимума в ЗМП

II

$\vec{x}_{opt} \in g_2(\vec{x}) = b_2$

ограничения $g_1(\vec{x}) = b_1$ и

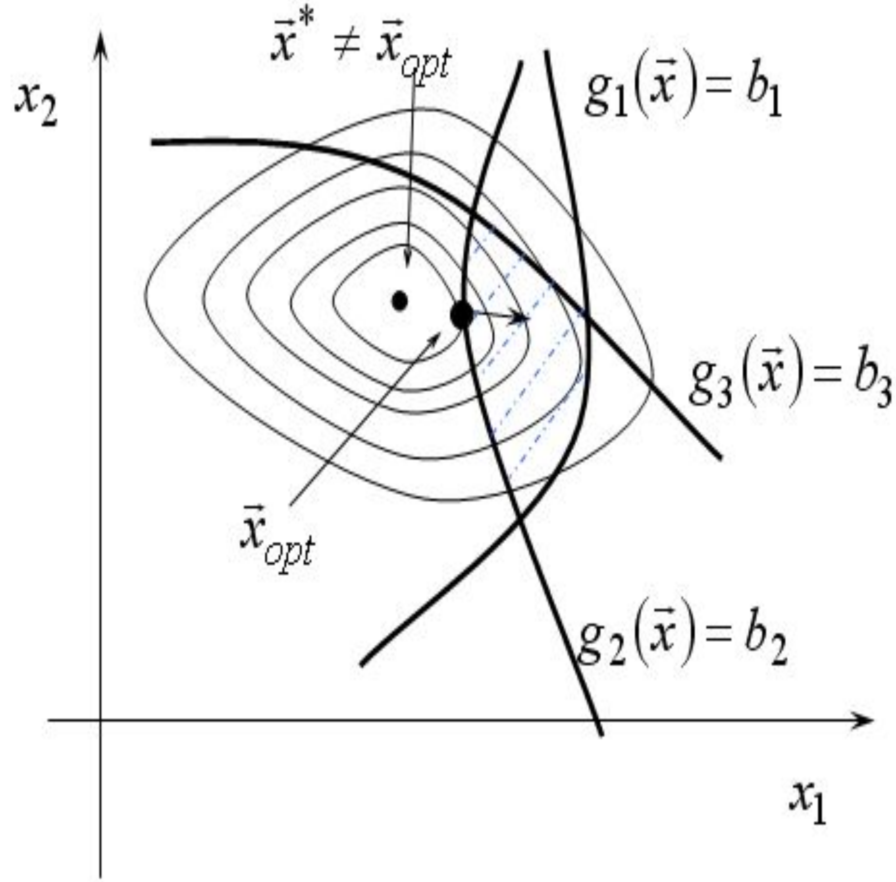
$g_3(\vec{x}) = b_3$ неактивные

исходная задача переходит к виду

$$f(\vec{x}) \to \min$$
$$g_2(\vec{x}) = b_2$$

в $\vec{x}_{opt}$ справедливо условие

$$\frac{\partial f(\vec{x})}{\partial x_j} = \lambda \frac{\partial g_2(\vec{x})}{\partial x_j} \quad j = \overline{1, n}$$

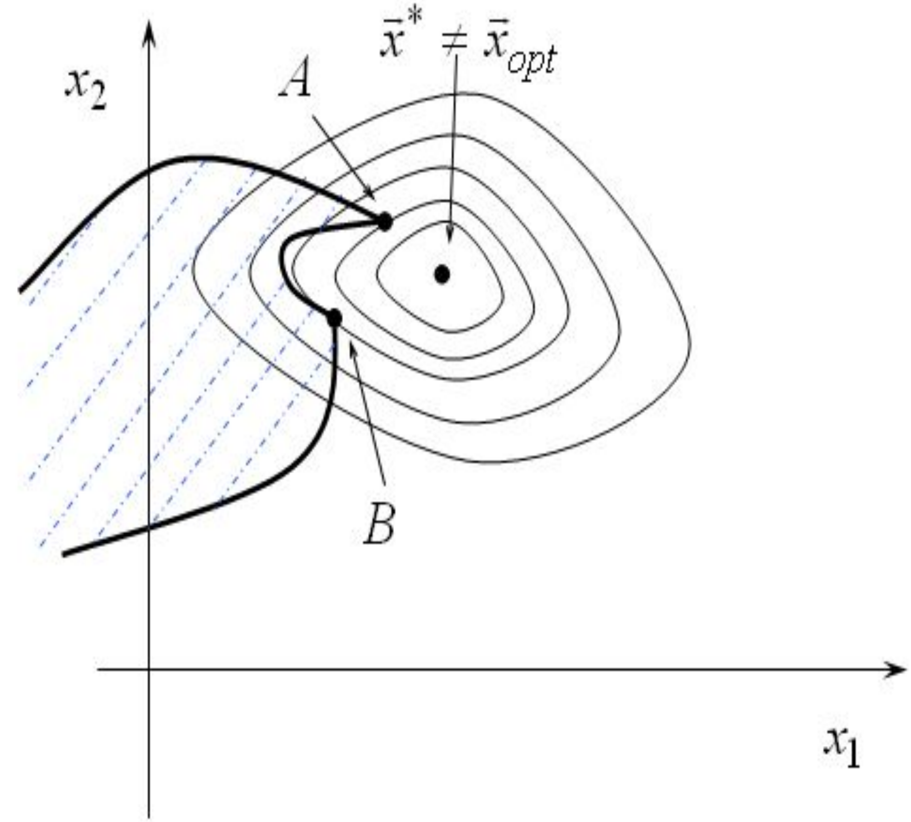


Возможные варианты положения минимума в ЗМП

III

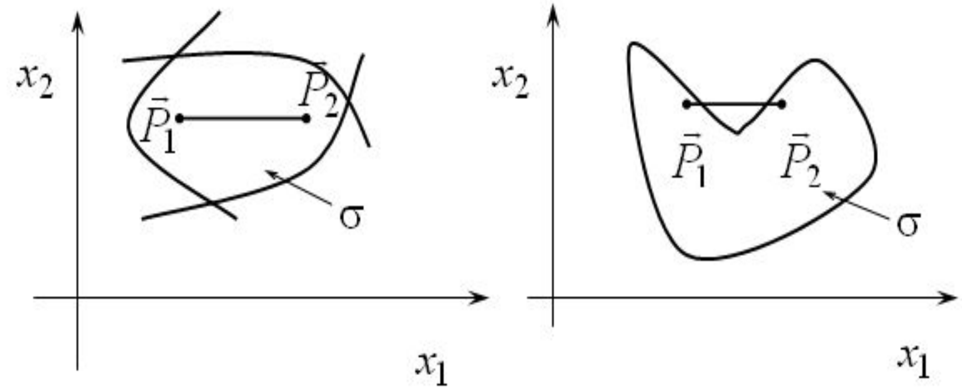
Функция имеет одну точку минимума в отсутствии ограничений

Для задачи с ограничениями A и B являются локальными минимумами



Задачи выпуклого программирования

1. ОДР – выпукла
2. Минимизируемая (максимизируемая) функция выпукла (вогнута)

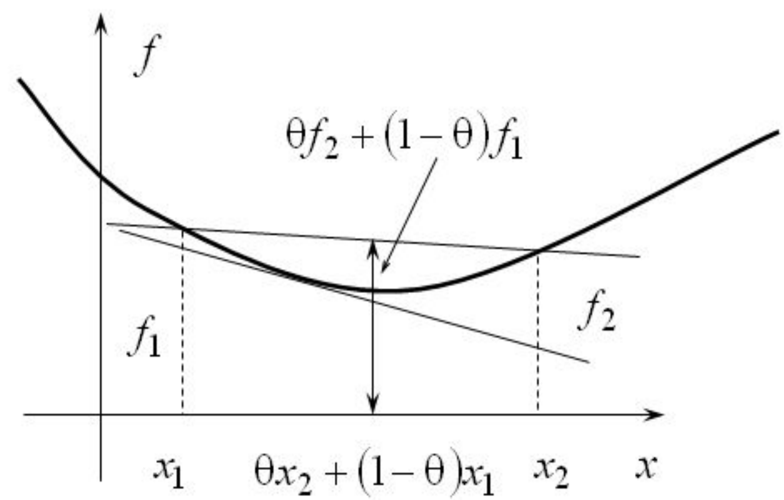


<u>Выпуклая область</u>:

$$\vec{P}_1, \vec{P}_2 \in \sigma$$
$$\vec{P}_\theta = \theta\vec{P}_2 + (1-\theta)\vec{P}_1 \in \sigma$$
$$0 < \theta < 1$$

Выпуклая функция $f(\vec{x})$ на выпуклой области X

$$f(\theta\vec{x}_2 + (1-\theta)\vec{x}_1) \le \theta f(\vec{x}_2) + (1-\theta) f(\vec{x}_1)$$
$0 < \theta < 1$ для $\forall \vec{x}_1, \vec{x}_2 \in X$



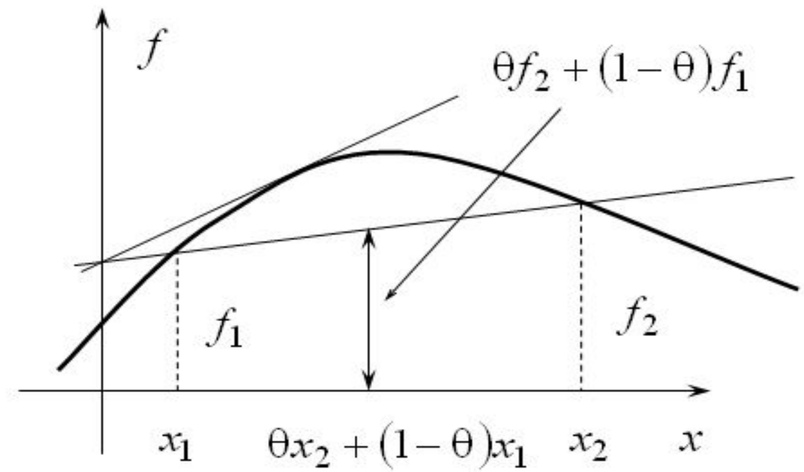
У выпуклой функции гессиан положительно определен

Задачи выпуклого программирования

Вогнутая функция $f(\vec{x})$ на выпуклой области X

$$f(\theta\vec{x}_2 + (1-\theta)\vec{x}_1) \geq \theta f(\vec{x}_2) + (1-\theta) f(\vec{x}_1)$$

$0 < \theta < 1$ для $\forall \vec{x}_1, \vec{x}_2 \in X$



Для задачи минимизации $f(\vec{x})$ при ограничениях $g_i(\vec{x}) \leq b_i \quad i = \overline{1,m}$, где $f(\vec{x})$ и $g_i(\vec{x})$ - выпуклые функции, <u>необходимые условия Куна-Такера</u> являются также <u>и достаточными</u>

В этом случае функция Лагранжа

$$L(\vec{x}, \vec{\lambda}, \vec{u}) = f(\vec{x}) + \sum_{i=1}^{m} \lambda_i \left[g_i(\vec{x}) + u_i^2 - b_i \right]$$

как сумма выпуклых функций является также выпуклой, а множители $\lambda_i \geq 0$, поэтому функция Лагранжа имеет глобальный минимум в точке, где ее производные исчезающе малы, и такая точка является единственной

Пример не единственного решения

$$f(x,y) = -\left(x^2 + y^2\right) \to \min$$

$$\begin{cases} x \ge 0 \\ y \ge 0 \\ x + 2y \le 3 \end{cases} \to \begin{cases} -x + u_1^2 = 0 \\ -y + u_2^2 = 0 \\ x + 2y + u_3^2 = 3 \end{cases}$$

Функция Лагранжа

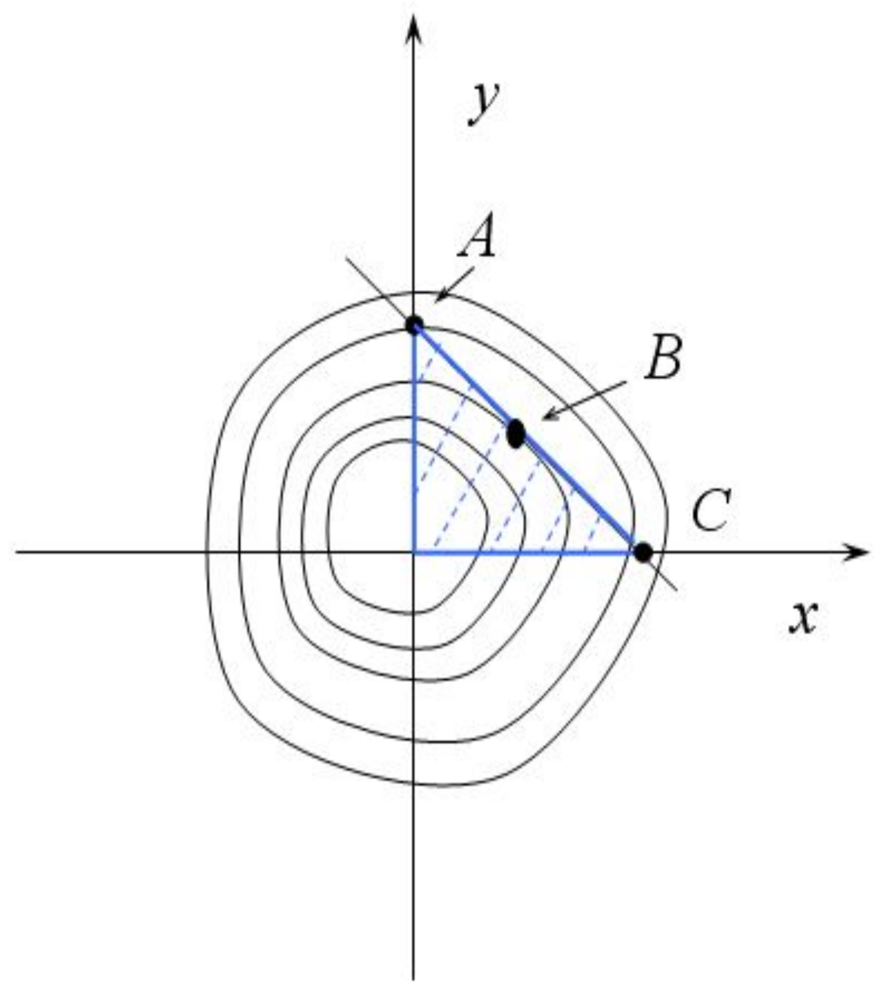
$$L\left(x, y, \vec{\lambda}, \vec{u}\right) = -x^2 - y^2 + \lambda_1\left(-x + u_1^2\right) + \lambda_2\left(-y + u_2^2\right) + \\ + \lambda_3\left(x + 2y + u_3^2 - 3\right)$$

Условия Куна-Такера

$$\begin{cases} -2x - \lambda_1 + \lambda_3 = 0 \\ -2y - \lambda_2 + 2\lambda_3 = 0 \\ \lambda_1 x = 0;\ \ \lambda_2 y = 0;\ \ \lambda_3(x + 2y - 3) = 0 \\ x \ge 0;\ y \ge 0;\ x + 2y \le 3 \\ \lambda_1 \ge 0;\ \lambda_2 \ge 0;\ \lambda_3 \ge 0 \end{cases}$$

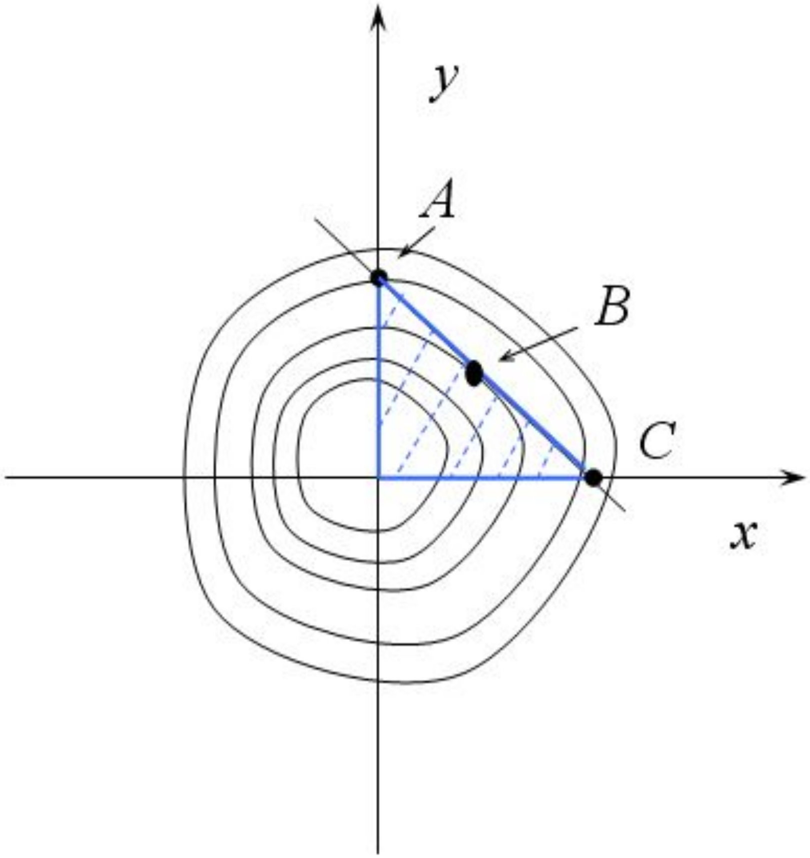
Решение:

1. при $x > 0$ и $y > 0$ $\lambda_1, \lambda_2 = 0$
2. если $\lambda_3 = 0$ $x = y = 0$ - это максимум $f(x,y)$



Пример не единственного решения

3. при $\lambda_1 = \lambda_2 = 0$ $\lambda_3 > 0$ имеем
   $2x = \lambda_3 = y$ и $x + 2y - 3 = 0$
   локальный минимум в точке $B(3/5, 6/5)$
   при $\lambda_3 = 6/5$ с $f = -45/25$
4. при $\lambda_1 > 0$ $\lambda_2 = 0$ $\lambda_3 > 0$ имеем
   $x = 0$ $y > 0$ $x + 2y = 3$
   локальный минимум в точке $A(0, 3/2)$
   при $\lambda_3 = 3/2$ $\lambda_1 = 3/2$ с $f = -9/4$
5. при $\lambda_1 = 0$ $\lambda_2 > 0$ $\lambda_3 > 0$ имеем
   $x > 0$ $y = 0$ $x + 2y = 3$
   глобальный минимум в точке $C(3,0)$
   при $\lambda_3 = 6$ $\lambda_2 = 12$ с $f = -9$



Здесь $g_i(x, y)$ - выпуклые функции
$f(x, y)$ - вогнутая функция
Условия Куна-Такера – необходимые (но не достаточные), поэтому решение не единственное

Комплексный метод Бокса

модификация симплексного метода Нельдера-Мида

<u>Применение</u>: для поиска минимума $f(\bar{x}) = f(x_1, x_2, \ldots, x_n)$ при ограничениях

$$g_{i1}(\bar{x}) \le b_{i1} \quad i1 = \overline{1, m} \qquad l_j \le x_j \le u_j \quad j = \overline{1, n}$$

При $f(\bar{x})$ и $g_{i1}(\bar{x})$ - выпуклых, задача имеет единственное решение

<u>Начальные условия</u> для реализации алгоритма: заданы n, m, l_j, u_j и начальное приближение $\bar{x}_1$ (удовлетворяет системе ограничений)

<u>I Построение начального комплекса</u>

комплекс имеет k вершин ($k > n+1$, рекомендовано $k = 2n$)

вершины комплекса $\bar{x}_i$ $\left(i = \overline{1, k}\right)$ - удовлетворяют системе ограничений, $f_i = f(\bar{x}_i)$

Для выбора вершин – итерационная процедура

$$x_{ij} = l_j + r\left(u_j - l_j\right) \quad j = \overline{1, n} \quad i = \overline{2, k}$$

r - псевдослучайная равномерно распределенная переменная в интервале $(0,1)$

если $\bar{x}_i$ не удовлетворяет ограничениям, то $\bar{x}_i' = \dfrac{\bar{x}_i + \bar{x}_c}{2}$

где $\bar{x}_c = \dfrac{1}{i-1}\sum\limits_{l=1}^{i-1} \bar{x}_l$ - центр тяжести заданных вершин комплекса

если $\bar{x}_i'$ - недопустима, расчет повторяют

Комплексный метод Бокса

II Поиск минимума $f(\bar{x})$ перемещением внутри ОДР

а) выбор $\bar{x}_h$ $\quad f_h = \max\limits_{1 \le i \le k} f_i$

расчет центра тяжести $(k-1)$ вершин комплекса $\bar{x}_0 = \dfrac{1}{k} \sum\limits_{i\ (i \ne h)} \bar{x}_{i-1}$

б) процедура отражения $\bar{x}_r = (1+\alpha)\bar{x}_0 - \alpha\bar{x}_h$, где $\alpha > 0$ - коэффициент отражения

в) проверка $\bar{x}_r$ на допустимость

если $\bar{x}_r$ недопустима $\quad \bar{x}_{rn} = \dfrac{\bar{x}_r + \bar{x}_0}{2}$

г) при допустимом $\bar{x}_r$ расчет f_r

если $f_r > f_h$ расчет $\bar{x}_{rn} = \dfrac{\bar{x}_r + \bar{x}_0}{2}$ и переход к в)

если $f_r < f_h$, то $\bar{x}_h \to \bar{x}_r$ и переход к а)

Проверка сходимости к минимуму:

1. среднеквадратичное отклонение σ для k значений функции

$$\sigma = \sqrt{\frac{1}{k}\sum_{l=1}^{k}\left[f(\bar{x}_l) - \bar{f}\right]^2} \quad \text{где } \bar{f} = \frac{1}{k}\sum_{l=1}^{k} f_l$$

2. d_m - максимальное расстояние между двумя вершинами комплекса

если σ и d_m - достаточно малы, минимум найден, иначе – переход к IIa)

Комплексный метод Бокса

Особенности метода:

1. $k = 2n$ - предотвращение преждевременного сжатия комплекса
2. $\alpha = 1.3$ - перемещение и расширение комплекса в направлении минимума функции
3. сходимость метода – при $f(\vec{x})$ выпуклой и ОДР выпуклой
4. Проверка получения глобального минимума – процедуру минимизации $f(\vec{x})$ повторяют при разных $\vec{x}_1$

Последовательная оптимизация без ограничений - метод штрафной функции

Основная идея:

<u>исходная задача с ограничениями</u> <u>эквивалентная задача без ограничений</u>

$$z = f(\vec{x}) \to \min \qquad \leftrightarrow \qquad Z = f(\vec{x}) + P(\vec{x}) \to \min$$

$$g_i(\vec{x}) \quad i = \overline{1,m} \qquad\qquad P(\vec{x}) \text{ - штрафная функция}$$

Вид $P(\vec{x})$ - множество вариантов представления

Пример:

$$z = f(\vec{x}) \to \min \qquad \textit{Пусть } P(\vec{x}) = r\sum_{j=1}^{m}\frac{1}{c_j(\vec{x})} \quad r > 0$$

$$c_j(\vec{x}) \geq 0 \quad j = \overline{1,m} \qquad Z = f(\vec{x}) + P(\vec{x}) = \varphi(\vec{x}, r) \to \min$$

1. $\vec{x}$ - допустимое решение $\rightarrow$ при $r\downarrow \quad (Z - z)\downarrow$

2. $\vec{x}$ - допустимое решение (близко к границе ОДР) $P(\vec{x})\uparrow \quad Z\uparrow$

Минимум Z - внутри ОДР задачи с ограничениями

при $r \to 0$ влияние $P(\vec{x})$ мало и $\min\left(Z = \varphi(\vec{x}, r)\right) \to \min(z)$

Решение ЗЛП методом штрафной функции

Дано: $f(x) = x \to \min \quad x \geq 2$

Эквивалентная задача

$$Z = \varphi(x, r) = x + \frac{r}{x-2} \to \min$$

Необходимое условие минимума

$$\frac{d\varphi(x,r)}{dx} = 1 - \frac{r}{(x-2)^2} = 0 \qquad \text{т.е. } x = 2 \pm \sqrt{r}$$

Достаточное условие минимума

$$\frac{d^2\varphi(x,r)}{dx^2} = \frac{2r}{(x-2)^3} > 0 \quad \text{при} \quad x = 2 + \sqrt{r}$$

Минимум исходной задачи

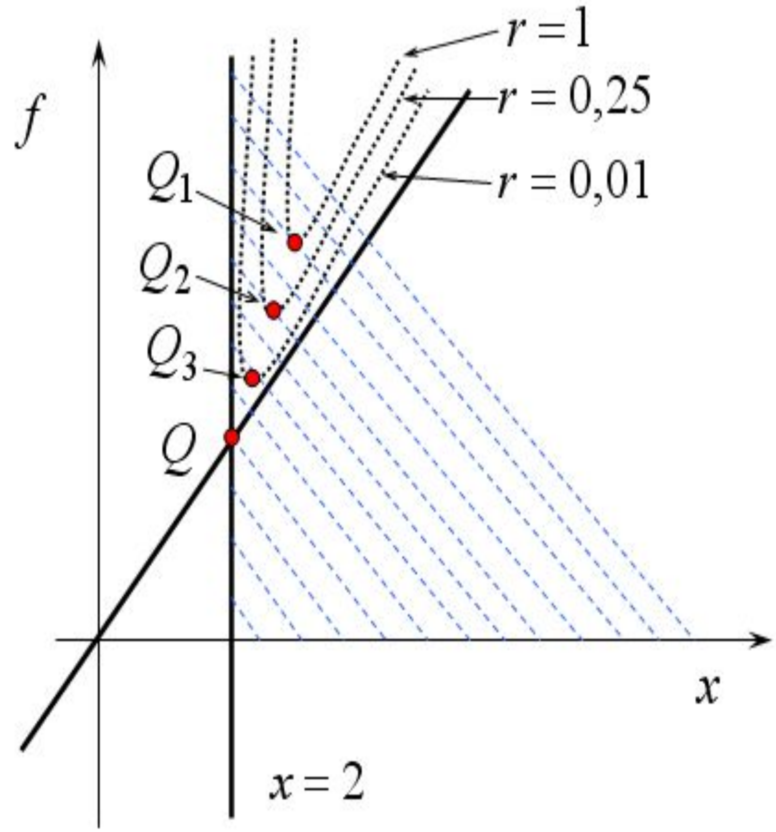
$Q \qquad x = 2 \quad f = 2$

Решение эквивалентной задачи

$Q_1 \quad r = 1 \qquad x = 3 \qquad f = 4$

$Q_2 \quad r = 0{,}25 \quad x = 2{,}5 \quad f = 3$

$Q_3 \quad r = 0{,}01 \quad x = 2{,}1 \quad f = 2{,}2$



Численный алгоритм реализации метода штрафной функции

<u>Условия реализации метода</u>: $f(\bar{x}) \to \min$ - выпукла, $c_j(\bar{x})$ - вогнуты

$\varphi(\bar{x}, r) = f(\bar{x}) + r \sum_j 1/c_j(\bar{x})$ выпукла в области ограничений

(ОДР выпукла)

$\varphi(\bar{x}, r)$ при r - имеет единственный минимум

$$r_1 > r_2 > \ldots > r_n \quad \Leftrightarrow \quad \bar{x}_1^*,\ \bar{x}_2^*,\ \ldots,\ \bar{x}_n^*$$

минимумы $\varphi(\bar{x}, r)$

$$\lim_{r_k \to 0} \bar{x}_k^* = \bar{x}^* \qquad \lim_{r_k \to 0} \left[\min \varphi(\bar{x}, r_k)\right] = f\left(\bar{x}^*\right)$$

где $\bar{x}^*$ - минимум $f(\bar{x})$ при наличии ограничений $c_j(\bar{x}) \geq 0$

<u>Алгоритм Кэрролла-Фиакко-Маккормика</u>

1. Для заданных $f(\bar{x})$ и $c_j(\bar{x}) \geq 0$ $j = \overline{1, m}$ - выбор $r = r_0$ и формирование $\varphi(\bar{x}, r_0)$

2. Поиск минимума $\varphi(\bar{x}, r_0)$ методом Давидона-Флетчера-Пауэла $\quad \to \quad \bar{x}_0^*$

Численный алгоритм реализации метода штрафной функции

3. $r_1 = \dfrac{r_0}{C}$ где $C = const > 1$

Поиск от начальной точки $\bar{x}_0^*$ минимума $\varphi(\bar{x}, r_1)$ методом Давидона-Флетчера-Пауэла $\rightarrow$ $\bar{x}_1^*$

. . . .

$r_k = \dfrac{r_{k-1}}{C}$ где $C = const > 1$

Поиск от начальной точки $\bar{x}_{k-1}^*$ минимума $\varphi(\bar{x}, r_k)$ методом Давидона-Флетчера-Пауэла $\rightarrow$ $\bar{x}_k^*$

Условие сходимости к минимуму

$$\left| \frac{\varphi(\bar{x}_k^*, r_k) - \varphi(\bar{x}_{k+1}^*, r_{k+1})}{\varphi(\bar{x}_k^*, r_k)} \right| < \varepsilon$$

где ε погрешность расчета $\bar{x}^*$ и $f(\bar{x}^*)$

Выбор r_0: $r_0 = 1$ и проверка $\bar{x}_k^*$ на каждом шаге расчета на допустимость (начало алгоритма всегда от начальной допустимой $\bar{x}_0$)