Тема 18 Методы условной оптимизации2

- *Понятие функции Лагранжа*
- *Задача линейного программирования*
- *Задача квадратичного программирования*

Постановка задач

- Найти минимум функции

$$y = f(x_1 \ldots x_n) = f(\vec{x}).$$

- При ограничениях

$$g_k(\vec{x}) \le 0,\; k = 1 \ldots m$$

$$g_k(\vec{x}) = 0,\; k = m+1 .. m+p$$

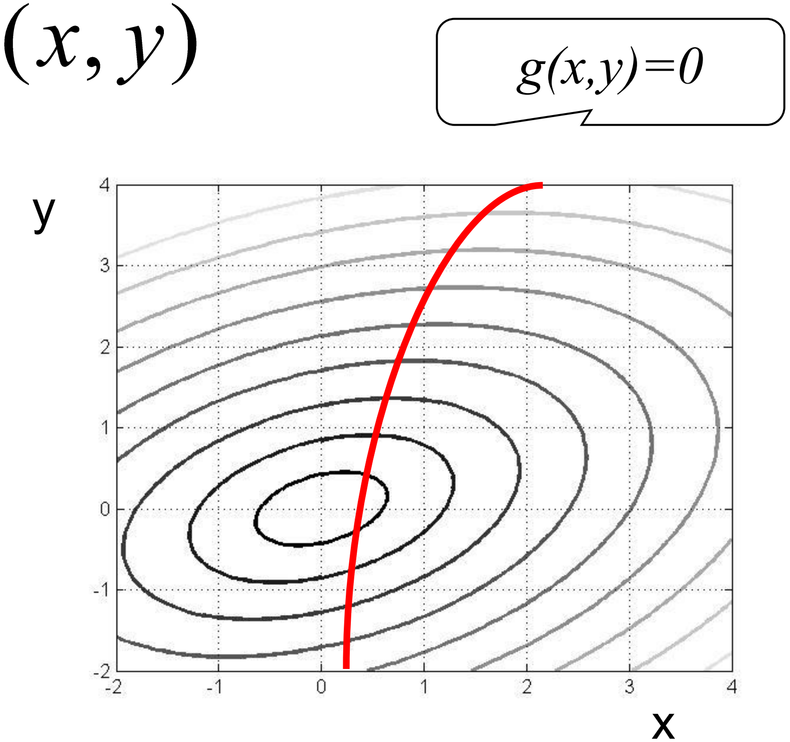
Понятие функции Лагранжа

- Вначале на простом примере функции двух переменных рассмотрим **какие условия в точке минимума имеют место** и как их проще получить

- Целевая функция $z = f(x, y)$
- Ограничение

$$g(x, y) \le 0$$

- Условный минимум лежит
- на кривой, описываемой
- уравнением $g(x, y) = 0$
- которое неявно определяет
- зависимость y=y(x)



Получение Условия минимума

- Вдоль кривой

$$g(x, y) = 0$$

- имеет место очевидное соотношение

$$\frac{dg}{dx} = 0 = \frac{\partial g}{\partial x} + \frac{\partial g}{\partial y}\frac{dy}{\partial x} \quad \Rightarrow \quad \frac{dy}{dx} = -\frac{\frac{\partial g}{\partial x}}{\frac{\partial g}{\partial y}}$$

Получение Условия минимума

- Запишем условие минимума для функции

$$z = f[x, y]$$

Вдоль кривой y=y(x)

$$z = f[x, y(x)]$$

имеем

$$\frac{dz}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx} = 0$$

Условия минимума

- Получаем

$$\frac{\partial f}{\partial x} - \frac{df}{\partial y}\,\frac{\dfrac{\partial g}{\partial x}}{\dfrac{\partial g}{\partial y}} = 0$$

Преобразуем

$$\frac{\dfrac{\partial f}{\partial x}}{\dfrac{\partial g}{\partial x}} - \frac{\dfrac{df}{\partial y}}{\dfrac{\partial g}{\partial y}} = 0$$

обозначим

$$\frac{\dfrac{\partial f}{\partial x}}{\dfrac{\partial g}{\partial x}} = \frac{\dfrac{df}{\partial y}}{\dfrac{\partial g}{\partial y}} = -\lambda$$

λ– множитель Лагранжа

Необходимые Условия минимума

- Таким образом в точке
- минимума f(x,y)
- на кривой g(x,y)=0
- выполняются
- три условия:

$$\begin{cases} \dfrac{\partial f(x^*, y^*)}{\partial x} + \lambda \dfrac{dg(x^*, y^*)}{\partial x} = 0 \\ \dfrac{\partial f(x^*, y^*)}{\partial y} + \lambda \dfrac{dg(x^*, y^*)}{\partial y} = 0 \\ g(x^*, y^*) = 0 \end{cases}$$

введем функцию Лагранжа

$$L(x, y, \lambda) = f(x, y) + \lambda g(x, y)$$

- Для нее условия экстремума которые мы выше вывели получаются естественным образом

$$\frac{\partial L}{\partial x} = \frac{\partial f(x, y)}{\partial x} + \lambda \frac{dg(x, y)}{\partial x} = 0$$

$$\frac{\partial L}{\partial y} = \frac{\partial f(x, y)}{\partial y} + \lambda \frac{dg(x, y)}{\partial y} = 0$$

$$\frac{\partial L}{\partial \lambda} = g(x, y) = 0$$

Рассмотрим простой пример

$$z = f(x, y) = x^2 + y^2 \qquad g(x, y) = x + y - 4 = 0$$

- Функция Лагранжа

$$L(x, y, \lambda) = x^2 + y^2 + \lambda(4 - x - y)$$

- Условия экстремума

$$\frac{\partial L}{\partial x} = 2x - \lambda = 0$$

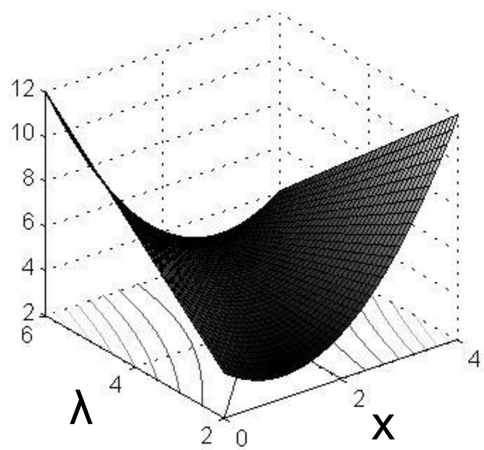
$$\frac{\partial L}{\partial y} = 2y - \lambda = 0$$

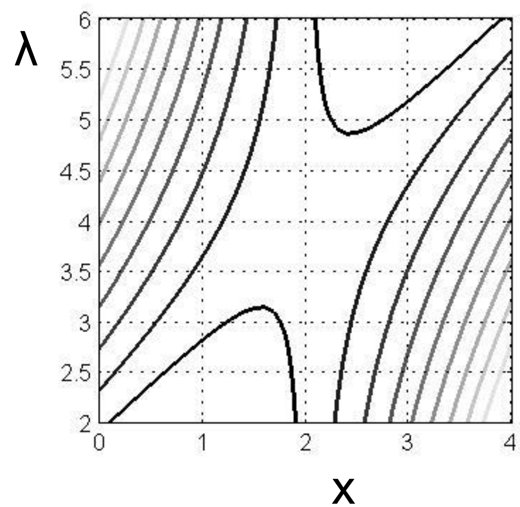
$$\frac{\partial L}{\partial \lambda} = 4 - x - y = 0$$

- Решение **x=2; y=2; λ=4.**

анализ, вблизи **точки экстремума**

- **точка экстремума** функции Лагранжа **представляет** **седловую точку**, в которой достигается минимум по переменным xy и максимум по переменной λ





Сечение функции Лагранжа
при y=2

$$L(x, y, \lambda) = x^2 + y^2 + \lambda(4 - x - y)$$

функция Лагранжа для нескольких ограничений в виде неравенств

$$\min f(\vec{x}) \qquad g_k(\vec{x}) \le 0,\ k = 1 \ldots m$$

$$L(\vec{x}, \vec{\lambda}) = f(\vec{x}) + \sum_{k=1}^{m} \lambda_k g_k(\vec{x}); \quad \lambda_k \ge 0; \ \vec{\lambda} = (\lambda_1 \ldots \lambda_m)$$

Условие регулярности:

Существует точка $\vec{x} \in D$ в которой все ограничения обращаются в строгие неравенства

$$g_k(\vec{x}) < 0,\ k = 1 \ldots m$$

Теорема о седловой точке Куна-Таккера

- В точке минимума $\vec{x}^*$ при указанных ограничениях существует такой набор

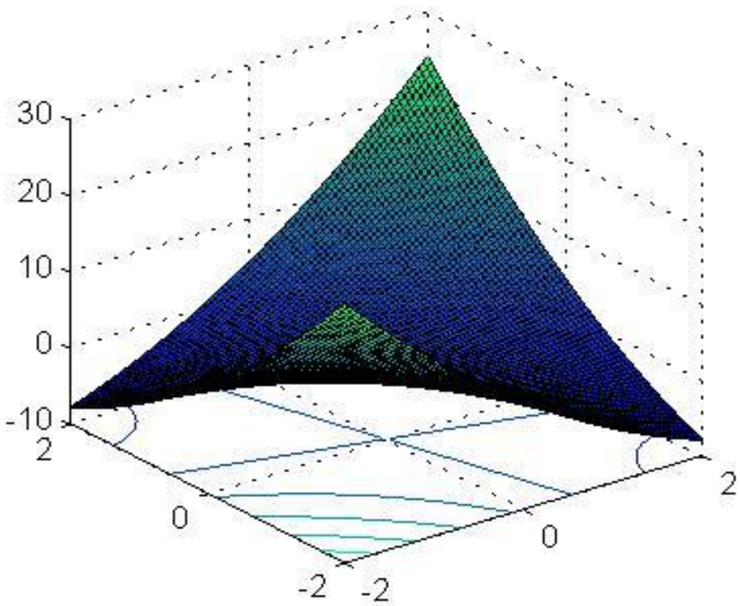
$$\vec{\lambda}^* = (\vec{\lambda}_1^*, \vec{\lambda}_2^*, \ldots \vec{\lambda}_m^*)$$

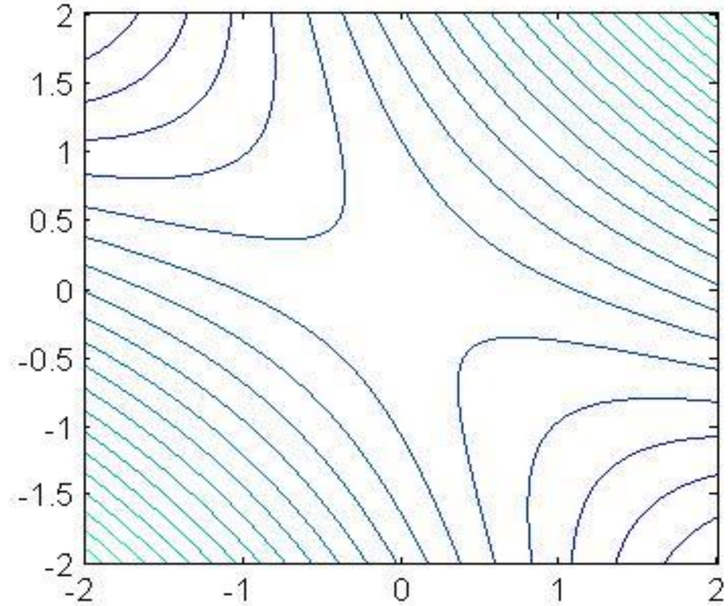
- при котором для всех $\vec{x}, \vec{\lambda}$ выполняется

$$L(\vec{x}^*, \vec{\lambda}) \le L(\vec{x}^*, \vec{\lambda}^*) \le L(\vec{x}, \vec{\lambda}^*)$$

- Точка $(\vec{x}^*, \vec{\lambda}^*)$ является седловой

Графическое представление седловой точки





Условия дополнительности

- Если пара $(\vec{x}^*, \vec{\lambda}^*)$ является седловой точкой функции Лагранжа, то выполняются условия дополнительности:

$$\lambda_k^* g_k(\vec{x}^*) = 0, \; k = 1 \ldots m$$

- Это значит, что для тех ограничений, на которых не лежит точка минимума, соответствующий множитель лагранжа равен нулю
- При этом

$$L(\vec{x}^*, \vec{\lambda}^*) = f(\vec{x}^*)$$

Понятие двойственности

- Допустим, что у функции Лагранжа седловая точка существует $(\vec{x}^*, \vec{\lambda}^*)$
- Положим $\varphi(\vec{\lambda}) = \min\limits_{\vec{x}} L(\vec{x}, \vec{\lambda}); \quad \lambda \geq 0$
- Аналогично $f(\vec{x}) = \max\limits_{\vec{\lambda}} L(\vec{x}, \vec{\lambda}); \quad \lambda \geq 0;$
- Из свойства седловой точки $\varphi(\vec{\lambda}^*) = f(\vec{x}^*)$
- Или $f(\vec{x}^*) = \min f(\vec{x}) = \max \varphi(\vec{\lambda})$
- Т.е можно вместо минимума находить максимум $\varphi(\vec{\lambda})$

Рассмотрим простой пример

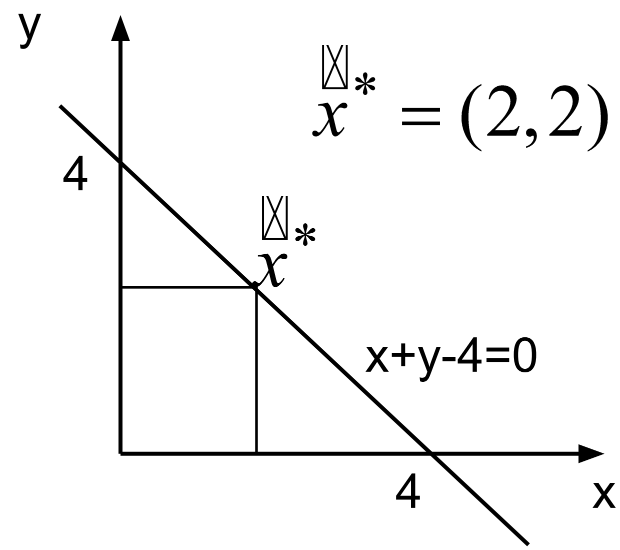
$$z = f(x, y) = x^2 + y^2$$

$$g_1 = 4 - x - y \le 0$$

$$g_2 = -x \le 0$$

$$g_3 = -y \le 0$$

$$\vec{x}^* = (2, 2)$$



$$L(x, y, \lambda) = x^2 + y^2 + \lambda_1(4 - x - y) - \lambda_2 x - \lambda_3 y$$

Из условия дополнительности $g_2(\vec{x}^*) < 0 \quad g_3(\vec{x}^*) < 0 \;\Rightarrow \lambda_2 = \lambda_3 = 0$

Двойственная задача

$$\varphi(\vec{\lambda}) = \min_{x,y}\left[x^2 + y^2 + \lambda_1(4 - x - y) - \lambda_2 x - \lambda_3 y\right]$$

$$\frac{\partial}{\partial x}\left[x^2 + y^2 + \lambda_1(4 - x - y) - \lambda_2 x - \lambda_3 y\right] = 0 \;\Rightarrow\; x = \frac{\lambda_1 + \lambda_2}{2}$$

$$\frac{\partial}{\partial y}\left[x^2 + y^2 + \lambda_1(4 - x - y) - \lambda_2 x - \lambda_3 y\right] = 0 \;\Rightarrow\; y = \frac{\lambda_1 + \lambda_3}{2}$$

$$\varphi(\vec{\lambda}) = \frac{(\lambda_1 + \lambda_2)^2}{4} + \frac{(\lambda_1 + \lambda_3)^2}{4} + \lambda_1\left(4 - \frac{\lambda_1 + \lambda_2}{2} - \frac{\lambda_1 + \lambda_3}{2}\right) - \lambda_2\frac{\lambda_1 + \lambda_2}{2} - \lambda_3\frac{\lambda_1 + \lambda_3}{2}$$

$$\max \varphi(\lambda_1, \lambda_2, \lambda_3);\; \lambda_i \ge 0 \quad \Rightarrow \quad \lambda_1 = 4 \quad \lambda_2 = 0 \quad \lambda_3 = 0$$

Задача линейного программирования

$$\sum_{j=1}^{n} c_j x_j \rightarrow \min \qquad \vec{c}^{\,\mathrm{T}} \vec{x} \rightarrow \min$$

$$\sum_{j=1}^{n} a_{ij} x_j \le b_i \qquad A\vec{x} \le \vec{b}$$

$$A^e \vec{x} = \vec{b}^e$$

$$x_j^{\min} \le x_j \qquad \vec{x}^{\min} \le \vec{x}$$

Пример: Имеются три продукта П1,П2,П3 разной цены, каждый из которых содержит определенное количество питательных ингридиентов И1,И2,И3,И4.
Известно, что в день требуется употребить :
И1 – не менее 250, И2 ≥ 60, И3 ≥100, И4≥200.
цена: П1=44, П2=35, П3=100.
Требуется минимизировать затраты на приобретение П

Затраты: x_j - количество

$$\sum_{j=1}^{n} c_j x_j = 44x_1 + 35x_2 + 100x_3$$

Питательность, или содержание Иi в Пj

$$\alpha_{ij} = \begin{bmatrix} 4 & 6 & 15 \\ 2 & 2 & 0 \\ 5 & 3 & 4 \\ 7 & 3 & 12 \end{bmatrix}$$

Решение в МатЛаб

- function LinProgr1;
- %Задание цены продуктов
- c=[44; 35; 100];
- % Матрица ограничений
- A=[4 6 15
- 2 2 0
- 5 3 4
- 7 3 12];
- A=-A;
- b=[250; 60; 100; 220];
- b=-b;
- xm=[0; 0; 0];
- % Обращение к стандартной программе
- [x,p]=linprog(c,A,b,[],[],xm)
- return

Результат

x =

13.2143

16.7857

6.4286

p =

1.8118e+003

Задача квадратичного программирования

$$f(\vec{x}) = \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} c_{ij} x_j x_k + \sum_{j=1}^{n} d_j x_j \to \min$$

$$f = \frac{1}{2}\vec{x}^{\mathrm{T}} C \vec{x} + \vec{d}^{\mathrm{T}} \vec{x}$$

$$A\vec{x} \le \vec{b} \qquad A^e \vec{x} = \vec{b}^e \qquad \vec{x}^{\min} \le \vec{x}$$

Задача о рисках

- Нужно вложить некоторую сумму в различные предприятия A1,A2,A3,A4 с целью получить желаемую доходность с минимальным риском.
- Доли вкладов : x1+x2+x3+x4=1; xi≥0
- Доходности Ai: y1 y2 y3 y4 и матрица ковариации V
- Тогда доходность и риски вычисляются по формулам:

$$D = \vec{y}^{\mathrm{T}}\vec{x} \qquad \sigma^2 = \vec{x}\ V\vec{x}$$

Задача: $$f = \vec{x}^{\mathrm{T}}V\vec{x} \qquad \vec{x} \geq 0$$

$$\vec{y}^{\mathrm{T}}\vec{x} = b_1^e \qquad (1111)\vec{x} = x_1 + x_2 + x_3 + x_4 = 1$$

Решение МатЛаб

```
function quadrogr1;
% матрица квадратичной формы
C=[102  27 -52  66
   27  148 42  -66
  -52  42  246  57
   66 -66  57   272];
% матрицы ограничений
Ae=[11 13 16 17.5;
     1  1  1  1 ];
be=[15; 1];
xm=[0; 0; 0];
% Обращение к стандартной пр-ме
[x,p]=quadprog(C,[],[],[],Ae,be,xm)
return
```

Результат:

```
Optimization
terminated.
x =
   0.0413
   0.4511
   0.1341
   0.3734
p =
  31.8349
```

Конец