Курс: Теория оптимизации и численные методы

Литература

1. Летова Т.А., Пантелеев А.В. Экстремум функции в примерах и задачах. - М.:МАИ, 1998.
2. Киреев В.И., Пантелеев А.В. Численные методы в примерах и задачах. - М.:МАИ, 2000.

Лекция № 1

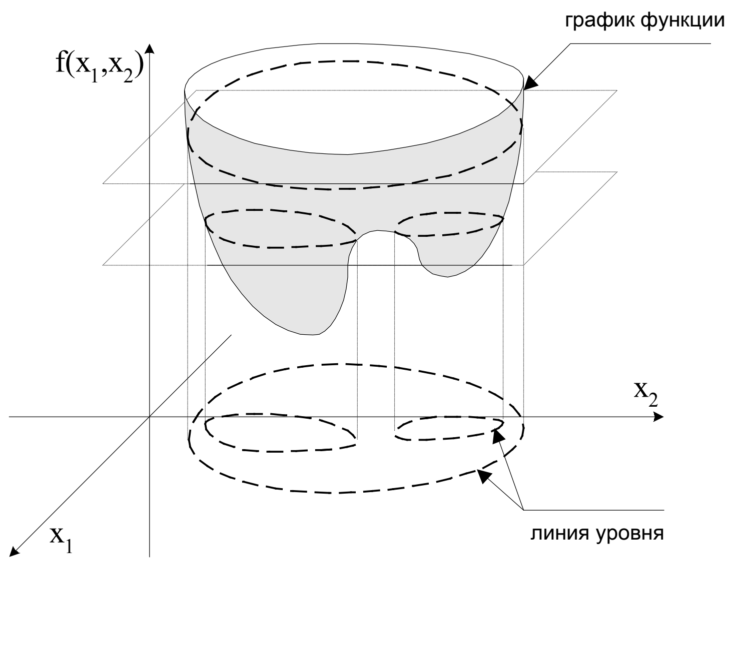
Часть I. Теория оптимизации и численные методы оптимизации

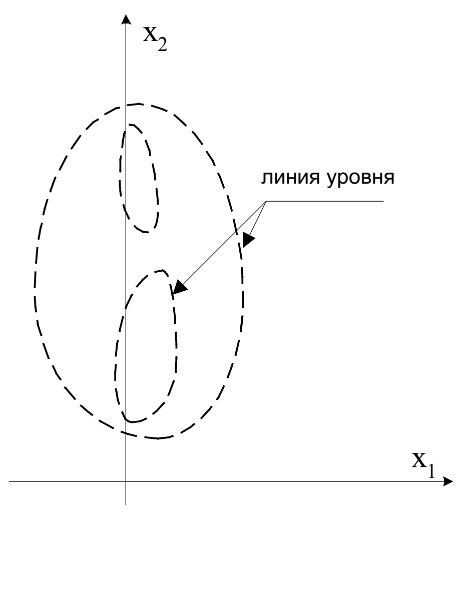
Основные понятия и определения

Рассматривается $f(X)$ - функция многих переменных, здесь $X = (x_{1,} ... x_n)^T$ - вектор переменных. Множество допустимых решений, на котором ищется решение задач, обозначается – $\boldsymbol{X}$.

Определение 1. *Поверхностью уровня* функции $f(X)$ называют геометрическое место точек, такое что $f(X) = C - \text{const}$. В случае 2-х переменных поверхность уровня называют линией уровня.

Определение 2. *Линия уровня* функции $f(x_1, x_2)$ - спроецированная на плоскость переменных x_1, x_2 сечение графика функции плоскостью $f(X) = C - \text{const}$.



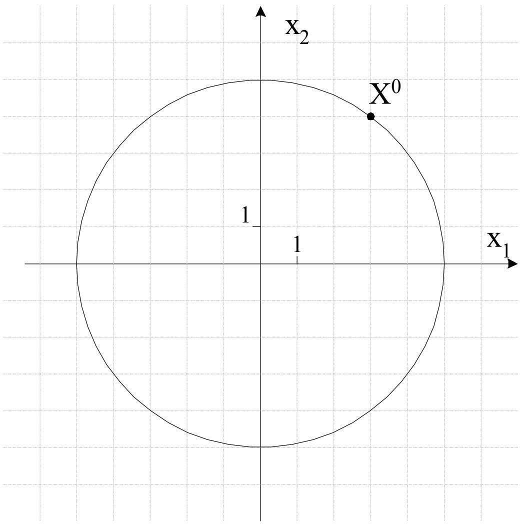


Для построения линии уровня функции $f(X)$ через заданную точку $X^0 = (x_1^0, x_2^0)$ необходимо вычислить значение функции в точке $f(X^0) = C_0$, затем записать уравнение линии уровня – уравнение плоской кривой в неявном виде $f(X) = C_0$ и построить соответствующий график.

Пример 1 Построить линию уровня функции $f(X) = x_1^2 + x_2^2$, проходящую через точку $X^0 = (3,4)$.

Последовательность действий:

1. Вычислим значение функции в точке $f(X^0) = 3^2 + 4^2 = 25$.
2. Запишем уравнение линии уровня: $x_1^2 + x_2^2 = 25$ - это уравнение окружности с центром в точке $(0,0)$ и радиусом 5.
3. Построим чертеж линии уровня.



Определение 3. *Градиентом* функции многих переменных $f(X)$ называется вектор, составленный из первых частных производных функции по всем переменным:

$$\nabla f(X) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, .., \frac{\partial f}{\partial x_n} \right)^T$$

Градиент - это вектор-столбец размерности $(n \times 1)$, где n - число переменных функции.

Свойства градиента:

(1) градиент функции перпендикулярен касательной к линии уровня функции $f(X)$

(2) направление градиента есть направление наиболее быстрого роста функции.

Для построения градиента функции двух переменных $f(X)$ в заданной точке $X^0 = (x_1^0, x_2^0)$ необходимо:

- найти частные производные функции $f(X)$ и записать градиент функции $\nabla f(X) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2} \right)^T$
- вычислить значения частных производных функции в точке X^0 и составить полученный вектор градиента $\nabla f(X^0) = \left(\left.\frac{\partial f}{\partial x_1}\right|_{X=X^0}, \left.\frac{\partial f}{\partial x_2}\right|_{X=X^0} \right)^T$

- построить полученный вектор на координатной плоскости из точки $(0, 0)$ и затем перенести его в заданную точку X^0.

Пример 2 Построить градиент функции $f(X) = x_1^2 + 3x_2^2$ в точке $X^0 = (0.5, 0.5)$.

Последовательность действий:

1. Построим линию уровня функции в точке $X^0 = (0.5, 0.5)$.

Получим $f(X^0) = 0.5^2 + 3 \cdot 0.5^2 = 1$, значит, уравнение линии уровня: $x_1^2 + 3x_2^2 = 1$ - это уравнение эллипса с центром в точке $(0, 0)$.

2. Составим градиент функции.

Получим $\dfrac{\partial f}{\partial x_1} = 2x_1$; $\dfrac{\partial f}{\partial x_2} = 6x_2$, следовательно $\nabla f(X) = (2x_1, \quad 6x_2)^T$.

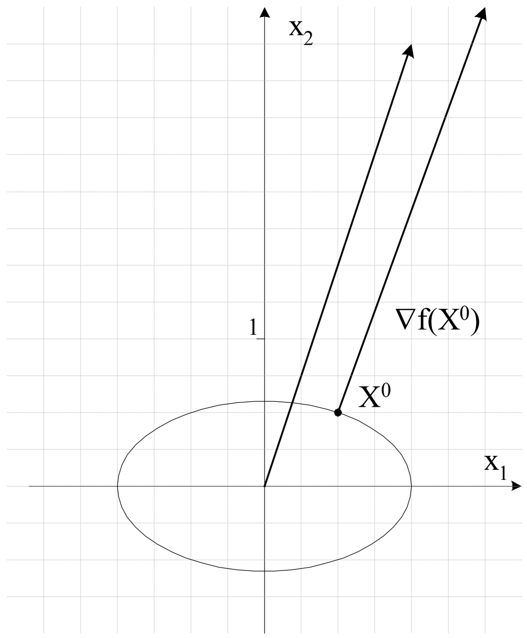
3. Вычислим значение частных производных функции в точке $X^0 = (0.5, 0.5)$:

$$\left.\frac{\partial f}{\partial x_1}\right|_{X^0} = 2 \cdot 0.5 = 1$$

$$\left.\frac{\partial f}{\partial x_2}\right|_{X^0} = 6 \cdot 0.5 = 3$$

Следовательно $\nabla f(X^0) = (1, \quad 3)^T$.

4. Построим полученный вектор на координатной плоскости из точки $(0, 0)$ и затем перенесем его в заданную точку X^0.



Определение 4. *Матрицей Гессе* называется квадратная матрица, составленная из вторых частных производных функции $f(X)$ по всем переменным, матрица имеет размерность $(n \times n)$, где n - число переменных функции:

$$H(X) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 x_2} & ..... & \frac{\partial^2 f}{\partial x_1 x_n} \\ \frac{\partial^2 f}{\partial x_2 x_1} & \frac{\partial^2 f}{\partial x_2^2} & ..... & \frac{\partial^2 f}{\partial x_1 x_n} \\ ..... & ..... & ..... & ..... \\ \frac{\partial^2 f}{\partial x_n x_1} & \frac{\partial^2 f}{\partial x_n x_2} & ..... & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

Определение 5. *Собственные числа* матрицы A - это корни характеристического уравнения вида:

$$\det(A - \lambda E) = 0$$

Определение 6.

Матрица A называется	*положительно определенной* $A > 0$	если ее собственные числа	положительны $\lambda_j > 0$
	отрицательно определенной $A < 0$		отрицательны $\lambda_j < 0$
	положительно полуопределенной $A \ge 0$		неотрицательны $\lambda_j \ge 0$
	отрицательно полуопределенной $A \le 0$		неположительны $\lambda_j \le 0$
	знаконеопределенной $A <> 0$		разного знака

Критерий Сильвестра (критерий знакоопределенности матрицы)

Матрица $A = \begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ ... & ... & ... & ... \\ a_{n1} & a_{n2} & ... & a_{nn} \end{pmatrix}$ является положительно определенной, если все ее диагональные миноры положительны; отрицательно определенной, если они чередуют знак, начиная с **«-»**.

Здесь диагональные миноры:

$$\Delta_1 = a_{11} \qquad \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \qquad ... \qquad \Delta_n = \begin{vmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ ... & ... & ... & ... \\ a_{n1} & a_{n2} & ... & a_{nn} \end{vmatrix}$$

Квадратичная функция двух переменных

Квадратичная функция 2-х переменных имеет вид:

$$f(X) = a_1x_1^2 + a_2x_1x_2 + a_3x_2^2 + a_4x_1 + a_5x_2 + a_6$$

Уравнение линии уровня квадратичной функции имеет вид:

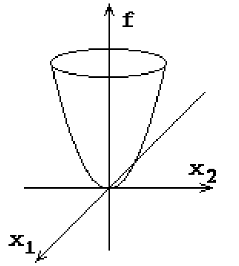
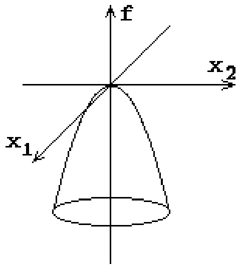
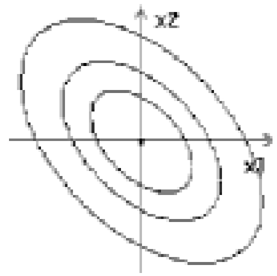
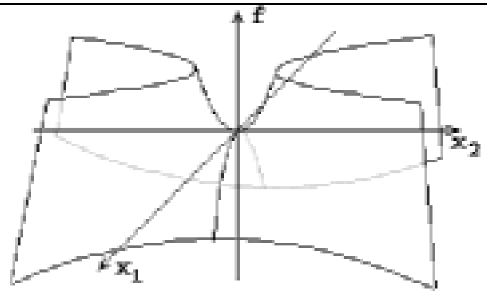
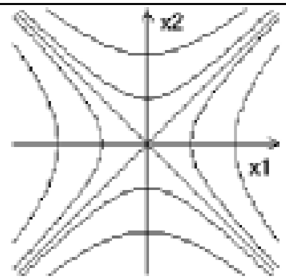
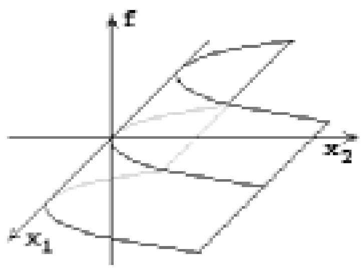
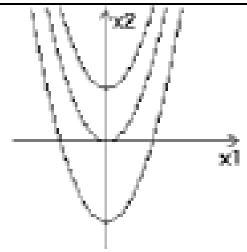
$$a_1x_1^2 + a_2x_1x_2 + a_3x_2^2 + a_4x_1 + a_5x_2 + a_6 = C$$

Инварианты для определения типа линии уровня квадратичной функции:

$$I = a_1 + a_2$$

$$D = \det\begin{pmatrix} a_1 & \frac{a_2}{2} \\ \frac{a_2}{2} & a_3 \end{pmatrix}$$ - определяющий

$$A = \det\begin{pmatrix} a_1 & \frac{a_2}{2} & \frac{a_4}{2} \\ \frac{a_2}{2} & a_3 & \frac{a_5}{2} \\ \frac{a_4}{2} & \frac{a_5}{2} & a_6 - C \end{pmatrix}$$

Инвариант	Функция	Линии уровня	Каноническое уравнение линии уровня
$D>0$ $\frac{A}{I}<0$ $I>0$ $D>0$ $\frac{A}{I}<0$ $I<0$	***минимум***  ***максимум*** 		***Эллипс (окружность)*** $\frac{(x_1-a)^2}{A^2}+\frac{(x_2-b)^2}{B^2}=1$ или $\alpha^2(x_1-a)^2+\beta^2(x_2-b)^2=R^2$ (если $A=B$ или $\alpha=\beta$ - то линия уровня окружность, в противном случае эллипс)
$D<0$ $A\neq 0$	***седло*** (нет ни минимума, ни максимума) 		***Гипербола*** $-\frac{(x_1-a)^2}{A^2}+\frac{(x_2-b)^2}{B^2}=1$ или $\frac{(x_1-a)^2}{A^2}-\frac{(x_2-b)^2}{B^2}=1$
$D=0$ $A\neq 0$	нет ни минимума, ни максимума 		***Парабола*** $x_2=Ax_1^2+Bx_1+C$ или $x_1=Ax_2^2+Bx_2+C$