Методы оптимизации

Лекция №1

От 10.02.2022

Постановка задачи оптимизации

Постановка задачи оптимизации подразумевает:

1. Формулировку цели, ради которой ставится задача.
2. Определение критерия отбора путей достижения цели.
3. Задание множества путей достижения цели: *множества допустимых решений* (МДР).

Задача оптимизации формулируется следующим образом: найти среди всех путей, ведущих к цели, наилучший, тот, для которого критерий принимает оптимальное значение.

По типу формирования критерия различают:

1. Задачи математического программирования – требуется найти оптимальное значение скалярной функции векторного аргумента.
2. Задачи вариационного исчисления – требуется найти оптимальное значение функционала.

ЧАСТЬ I. ЗАДАЧИ МАТЕМАТИЧЕСКОГО ПРОГРАММИРОВАНИЯ

Основные понятия и определения

Заглавными буквами будем обозначать вектора, строчными – переменные.

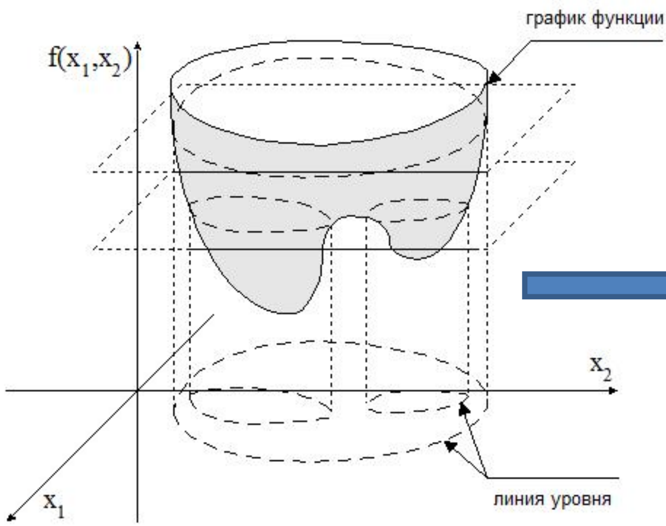
Рассматривается $f(X)$ – функция многих переменных (ФМП), здесь $X = (x_1,...x_n)^T$ – вектор переменных.

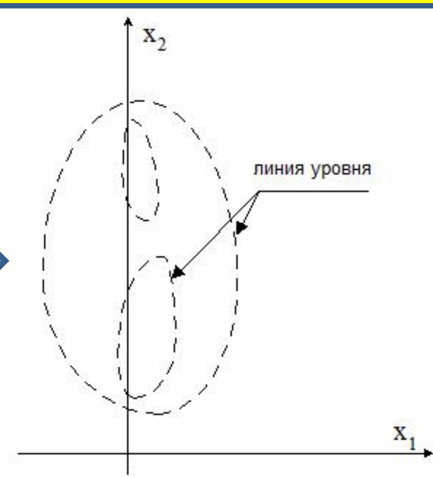
Определение 1. *<u>Поверхностью уровня</u>* функции $f(X)$ называют геометрическое место точек, такое что $f(X) = C - \text{const}$. В случае 2-х переменных поверхность уровня называют линией уровня.

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

Линии уровня позволяют получить представление о поведении функции без построения трехмерного графика.

Примером использования линий уровня являются физические географические карты.





Для построения линии уровня функции $f(X)$ через заданную точку $X^0 = (x_1^0, x_2^0)$ необходимо:

1. Вычислить значение функции в точке $f(X^0) = C_0$.
2. Записать уравнение линии уровня – уравнение плоской кривой в неявном виде $f(X) = C_0$.
3. Построить соответствующий график.

Пример 1. Построить линию уровня функции $f(X) = x_1^2 + x_2^2 - 2x_1 + 4x_2 + 5$, проходящую через точку $X^0 = (4;\ 2)$.

Решение:

1. Вычислим значение функции в точке $f(X^0) = 4^2 + 2^2 - 8 + 8 + 5 = 25$.

2. Запишем уравнение линии уровня: $x_1^2 + x_2^2 - 2x_1 + 4x_2 + 5 = 25$.

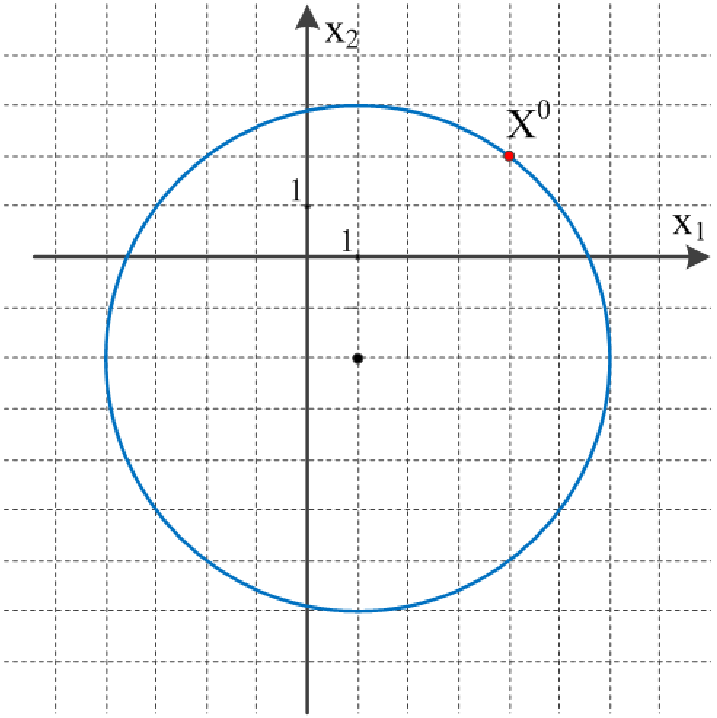
3. Для построения чертежа выделим полные квадраты в левой части уравнения:

$$\underbrace{x_1^2 - 2x_1 + 1} - 1 + \underbrace{x_2^2 + 4x_2 + 4} - 4 + 5 = 25$$

$$(x_1 - 1)^2 + (x_2 + 2)^2 = 25$$

Это уравнение окружности с центром в точке $(1;\ -2)$ и радиусом 5.

Построим чертеж линии уровня.



Важно! Заданная точка должна принадлежать построенной линии уровня.

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

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

В данном курсе используется символ ∇ для обозначения градиента.
Другие обозначения не допускаются.

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.

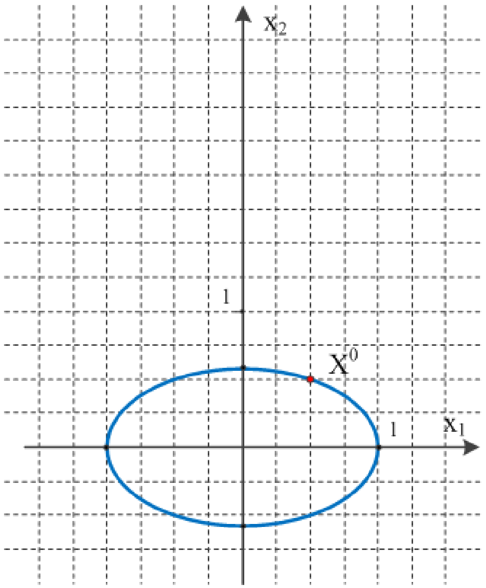
Для построения градиента функции двух переменных $f(X)$ в заданной точке $X^0=(x_1^0,x_2^0)$ необходимо:

1. Найти частные производные функции $f(X)$ и записать градиент функции $\nabla f(X)=\left(\frac{\partial f}{\partial x_1},\frac{\partial f}{\partial x_2}\right)^T$.
2. Вычислить значения частных производных функции в точке X^0 и записать полученный числовой вектор $\nabla f(X^0)=\left(\left.\frac{\partial f}{\partial x_1}\right|_{X=X^0},\left.\frac{\partial f}{\partial x_2}\right|_{X=X^0}\right)^T$.
3. Построить полученный числовой вектор на координатной плоскости из точки $(0,0)$.
4. Перенести его в заданную точку X^0, используя параллельный перенос.

Пример 2. Построить линию уровня и градиент функции $f(X) = x_1^2 + 3x_2^2$ в точке $X^0 = (0{,}5;\ 0{,}5)$.

Решение:

I. Построим линию уровня функции в точке $X^0 = (0{,}5;\ 0{,}5)$.



1. Найдем значение функции в заданной точке: $f(X^0) = 0{,}5^2 + 3 \cdot 0{,}5^2 = 1$.

2. Запишем уравнение линии уровня: $x_1^2 + 3x_2^2 = 1$ – это уравнение эллипса с центром в точке $(0;\ 0)$.

3. Для построения линии уровня используем точки пересечения эллипса с главными осями.

Пусть $x_2 = 0$, тогда $x_1^2 = 1 \Rightarrow x_1 = \pm 1$.

Пусть $x_1 = 0$, тогда $3x_2^2 = 1 \Rightarrow x_2 = \pm\sqrt{1/3} \approx \pm 0{,}5773$.

Также, воспользуемся тем, что точка $X^0 = (0{,}5;\ 0{,}5)$ принадлежит линии уровня.

x_1	x_2
± 1	0
0	$\pm\sqrt{1/3} \approx \pm 0{,}5773$
0,5	0,5

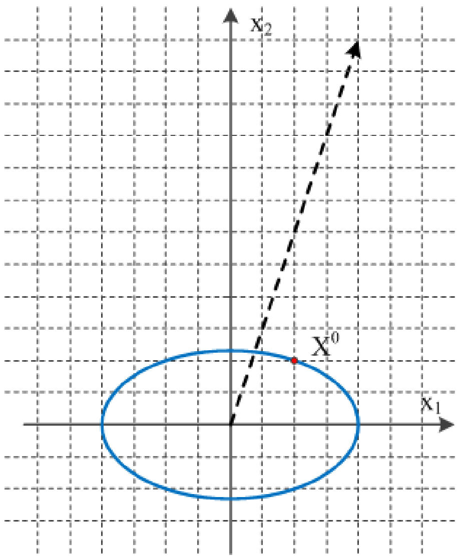
II. Построим градиент функции в точке $X^0 = (0{,}5;\ 0{,}5)$.

1. Найдём частные производные функции 1-го порядка: $\dfrac{\partial f}{\partial x_1} = 2x_1$; $\dfrac{\partial f}{\partial x_2} = 6x_2$. Следовательно $\nabla f(X) = (2x_1,\ \ 6x_2)^T$.

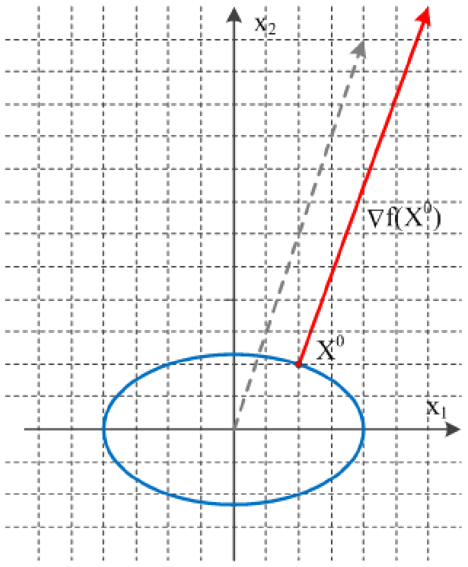
2. Вычислим значения частных производных функции в точке $X^0 = (0{,}5;\ 0{,}5)$: $\left.\dfrac{\partial f}{\partial x_1}\right|_{X^0} = 2 \cdot 0{,}5 = 1$, $\left.\dfrac{\partial f}{\partial x_2}\right|_{X^0} = 6 \cdot 0{,}5 = 3$.

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

3. Построим полученный вектор $\nabla f(X^0) = (1;\ 3)^T$ на координатной плоскости из точки $(0;\ 0)$.



4. Перенесем построенный вектор в точку $X^0 = (0{,}5;\ 0{,}5)$, используя параллельный перенос.



Важно!

1. При параллельном переносе сохраняется длина вектора.

2. Построенный градиент должен быть перпендикулярен касательной к линии уровня в заданной точке.

Определение 4. *Матрицей Гессе* называется квадратная матрица, составленная из вторых частных производных функции $f(X)$ по всем переменным:

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

> **Людвиг Отто Гессé —**
> **немецкий математик, член Баварской АН, профессор Политехнической школы в Мюнхене.** Основные работы относятся к геометрии, линейной алгебре, вариационному исчислению; ввёл понятие гессиана.

Матрица Гессе – это квадратная матрица размерности $(n \times n)$, где n – число переменных функции, симметричная относительно главной диагонали.

Определение 5. *Собственные значения (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 $A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$ **является** положительно определенной, если все ее *угловые миноры* положительны.

Матрица $A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$ **является** отрицательно определенной, если ее *угловые миноры* чередуют знак, начиная со знака минус.

Угловые миноры матрицы $A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$ имеют вид: $\Delta_1 = a_{11}$ $\quad \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$ $\quad \dots \quad$ $\Delta_n = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$.

Пример 3. Найти матрицу Гессе для функции $f(X) = x_1^2 - x_1 x_2 + 2x_2^2$.

Исследовать знакоопределенность матрицы по критерию Сильвестра и на основании информации о собственных значениях.

Решение:

1. Найдём сначала частные производные функции 1-го порядка:

$$\frac{\partial f}{\partial x_1} = 2x_1 - x_2;$$

$$\frac{\partial f}{\partial x_2} = -x_1 + 4x_2.$$

Теперь найдем частные производные функции 2-го порядка:

$$\frac{\partial^2 f}{\partial x_1 \partial x_1} = 2; \qquad \frac{\partial^2 f}{\partial x_1 \partial x_2} = -1;$$

$$\frac{\partial^2 f}{\partial x_2 \partial x_1} = -1; \qquad \frac{\partial^2 f}{\partial x_1 \partial x_2} = 4.$$

Следовательно, матрица Гессе имеет вид: $H(X) = \begin{pmatrix} 2 & -1 \\ -1 & 4 \end{pmatrix}$. Она является **постоянной**.

2. Исследуем знакоопределенность матрицы по критерию Сильвестра:

$$\Delta_1 = 2 > 0$$

$$\Delta_2 = \begin{vmatrix} 2 & -1 \\ -1 & 4 \end{vmatrix} = 2 \cdot 4 - (-1) \cdot (-1) = 7 > 0$$

Т.к. все угловые миноры положительны, то матрица положительно определена $H(X) > 0$ в любой точке.

3. Исследуем знакоопределенность матрицы на основании информации о собственных значениях.

Для этого составим характеристическое уравнение:

$$\det(H - \lambda E) = \begin{vmatrix} 2-\lambda & -1 \\ -1 & 4-\lambda \end{vmatrix} = (2-\lambda) \cdot (4-\lambda) - (-1) \cdot (-1) = 0$$

$$\lambda^2 - 6\lambda + 7 = 0$$

$$\lambda_1 = \frac{6+\sqrt{8}}{2} > 0; \quad \lambda_2 = \frac{6-\sqrt{8}}{2} > 0.$$

Т.к. все собственные значения положительны, то матрица положительно определена $H(X) > 0$ в любой точке.

Квадратичная функция двух переменных

Квадратичная функция 2-х переменных имеет вид:

$$f(X) = a_1 x_1^2 + a_2 x_1 x_2 + a_3 x_2^2 + a_4 x_1 + a_5 x_2 + a_6 .$$

Уравнение линии уровня квадратичной функции имеет вид:

$$a_1 x_1^2 + a_2 x_1 x_2 + a_3 x_2^2 + a_4 x_1 + a_5 x_2 + a_6 = C .$$

Инварианты для определения типа линии уровня квадратичной функции:

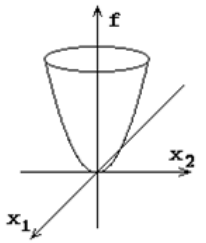
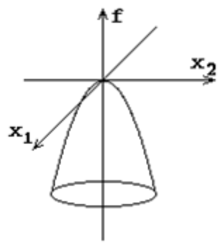
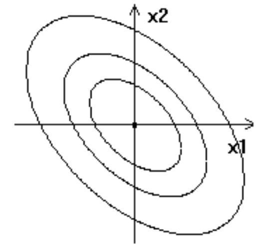
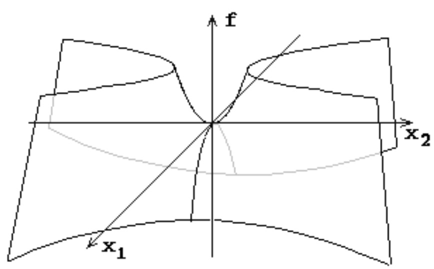
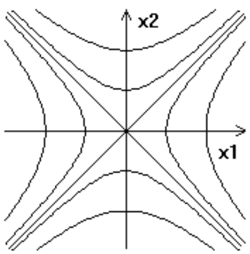
$$I = a_1 + a_2 .$$

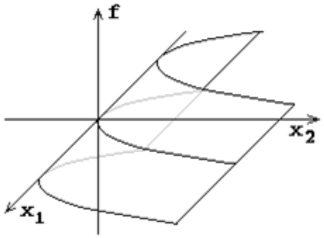
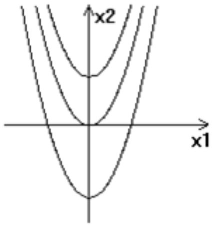
$$D = \det\begin{pmatrix} a_1 & \frac{a_2}{2} \\ \frac{a_2}{2} & a_3 \end{pmatrix}$$ – определяющий.

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

Инвариант – термин, обозначающий нечто неизменяемое. В данном случае инварианты являются характеристиками линии уровня.

Классификация невырожденных линий уровня квадратичной функции

Инвариант	Функция		Линии уровня	Уравнение линии уровня
$D > 0$ $\frac{A}{I} < 0$ $I > 0$ $D > 0$ $\frac{A}{I} < 0$ $I < 0$	***минимум*** ***максимум***	 	 ***Эллипс (окружность)***	***Эллипс (окружность)*** $A^2(x_1-a)^2 + B^2(x_2-b)^2 = R^2$ **или** $\frac{(x_1-a)^2}{\alpha^2} + \frac{(x_2-b)^2}{\beta^2} = 1$ **(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***	***Гипербола*** $-A^2(x_1-a)^2 + B^2(x_2-b)^2 = R^2$ или $A^2(x_1-a)^2 - B^2(x_2-b)^2 = R^2$

$D = 0$ $A \neq 0$	нет ни минимума, ни максимума 	 ***Парабола***	***Парабола*** $x_2 = Ax_1^2 + Bx_1 + C$ или $x_1 = Ax_2^2 + Bx_2 + C$