Методы многомерной оптимизации.

Дана некоторая функция многих переменных $f(x_1,x_2,x_3,\ldots,x_n)$ или $f(\vec{x})$ надо найти такое значение $\vec{x}^*$, при котором функция $f(\vec{x}^*)$ принимает экстремальное значение (минимальное или максимальное).

Некоторая функция $f(\vec{x})$ в точке $\vec{x}^*$ имеет минимум, если в любой сколь угодно малой окрестности точки $\vec{x}^*$ выполняется условие $f(\vec{x}) > f(\vec{x}^*)$ и максимум, если выполняется условие $f(\vec{x}) < f(\vec{x}^*)$.

Процесс поиска экстремального значения иногда называют оптимизацией. Функцию $f(\vec{x})$ называют оптимизируемой или целевой функцией, а вектор $\vec{x}$ оптимизируемые параметры или параметры оптимизации. В области определения функции может быть несколько экстремумов. В этом случае все экстремальные значения называются локальными экстремумами, а наибольшие или наименьшие значения из всех локальных называются глобальными экстремумами. Перед использованием метода оптимизации необходимо чтобы было известно:

1) вид функции $f(\vec{x})$;
2) границы поиска по каждой переменной $[a_1;b_1]$, $[a_2;b_2]$,…..,$[a_n;b_n]$ либо, вместо границ, может быть задана начальная точка поиска экстремума

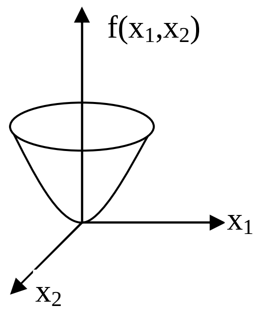
$$x^{(0)} = \begin{bmatrix} x_1^{(0)} \\ x_2^{(0)} \\ \ldots\ldots \\ x_n^{(0)} \end{bmatrix};$$

3) точность поиска экстремума.

Все существующие методы многомерной оптимизации позволяют определить лишь один из локальных экстремумов.

Какой именно локальный экстремум будет определён, зависит от выбора начального приближения.

Все методы поиска экстремума будем рассматривать относительно случая поиска минимума, для функции двух переменных. Для удобства графической иллюстрации методов определим представление функции в виде линий уровня.



Дана целевая функция $f(\vec{x}) = x_1^2 + x_2^2$, которая графически представляет собой поверхность параболоида вращения.

Проведем сечения поверхности равно отстоящими плоскостями, которые параллельны плоскости изменения переменных x_1 и x_2.

Линии этих сечений проецируем на плоскость изменения переменных. Получим концентрические окружности. Эти линии называются линиями уровня или линиями постоянных значений. Основная характеристика любой из линий это то, что в любой точке этой линии значение функции постоянно.

Рассечем заданную поверхность функции тремя плоскостями по уровням.

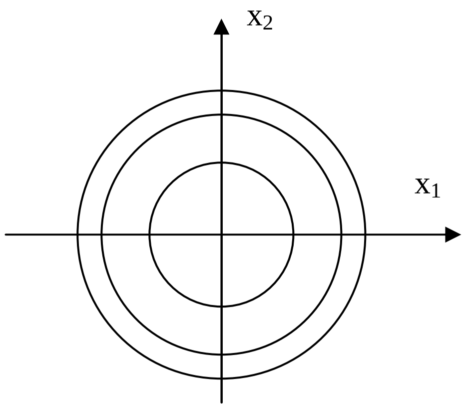
$$y_1 = x_1^2 + x_2^2 = 1 \qquad y_2 = x_1^2 + x_2^2 = 2 \qquad y_3 = x_1^2 + x_2^2 = 3$$

Тогда линии уровня будут представлены уравнениями

$$x_1 = \sqrt{1 - x_2^2} \qquad x_1 = \sqrt{2 - x_2^2} \qquad x_1 = \sqrt{3 - x_2^2}$$

или окружностями с соответствующими радиусами

$R_1=\sqrt{1}=1$ $R_2=\sqrt{2}=1.4$ $R_3=\sqrt{3}=1.73$



Все методы многомерной оптимизации делятся на два класса:

1) Градиентные
2) Безградиентные

Градиентом называется вектор равный сумме произведений частных производных на соответствующие орты. Орта – единичный связанный вектор, направленный вдоль координатной оси. $\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ . \\ 0 \end{bmatrix}$

$\vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ . \\ 0 \end{bmatrix}$ ……….. $\vec{e}_n = \begin{bmatrix} 0 \\ 0 \\ . \\ 1 \end{bmatrix}$ $\quad \vec{\mathrm{grad}}(f(\vec{x})) = \frac{\partial f}{\partial x_1} \cdot \vec{e}_1 + \frac{\partial f}{\partial x_2} \cdot \vec{e}_2 + \ldots\ldots\ldots + \frac{\partial f}{\partial x_n} \cdot \vec{e}_n$

Для случая функции двух переменных $\vec{\mathrm{grad}}(f(\vec{x})) = \frac{\partial f}{\partial x_1} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \frac{\partial f}{\partial x_2} \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix}$

Основные свойства градиента:

1) Норма градиента определяет скорость изменения функции в направление градиента.
2) Градиент всегда направлен в сторону наиболее быстрого возрастания функции, т.е. в этом направлении норма вектора градиента максимальна.

Антиградиентом называется вектор, направленный в противоположную сторону.

Пример:

дана функция $f(\vec{x}) = x_1^2 + x_2^2$. Вычислить вектор градиент в точке $\vec{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

$$\vec{\mathrm{grad}}(f(\vec{x})) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

Градиентный метод.

1. Дана функция n переменных $f(\vec{x})$. Также заданы точность ε, величина параметра шага h и точка начального приближения $\vec{x}^{(0)}$ расположенная в области поиска экстремума.
2. За текущее значение вектора приближения принимаем начальное приближение $\vec{x} = \vec{x}^{(0)}$ и вычисляем значение функции в этой точке FX=$f(\vec{x})$

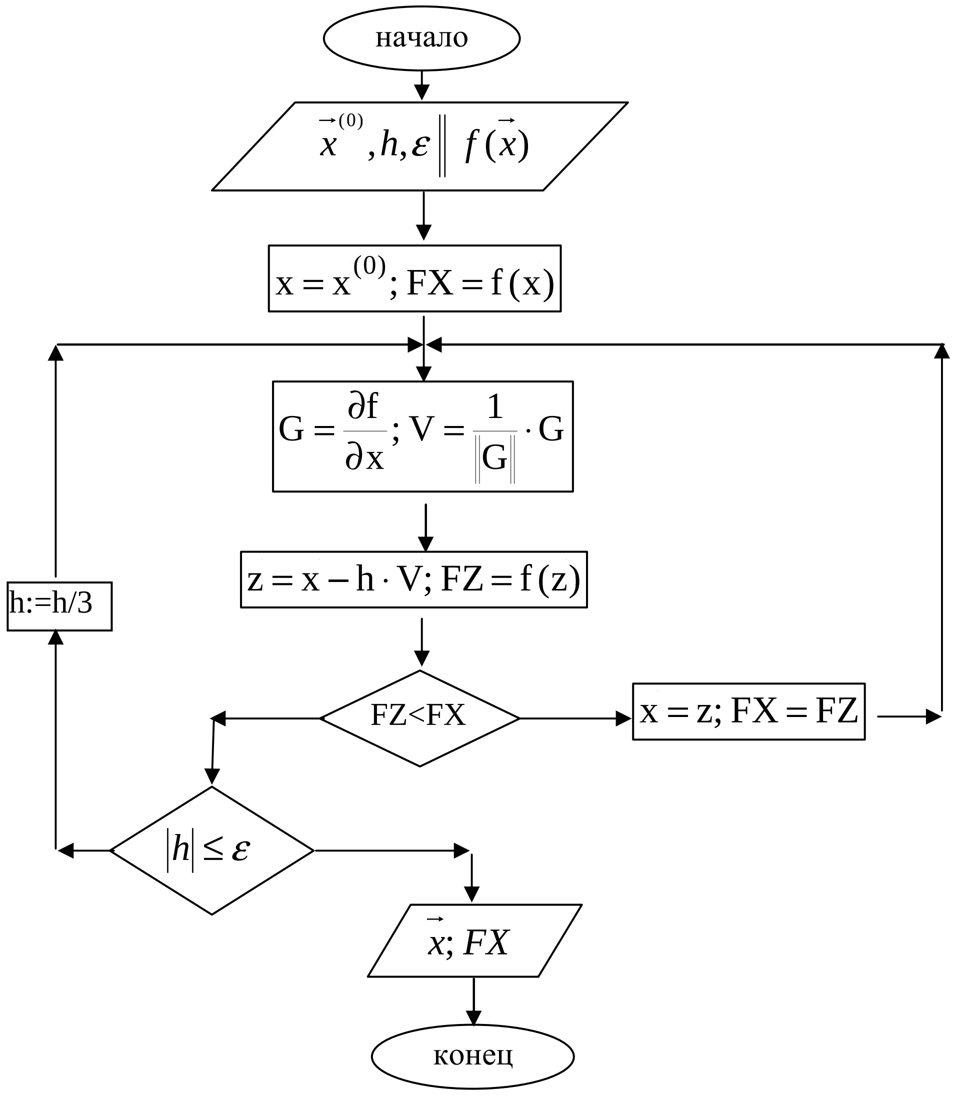
3. Вычисляем вектор градиента $\vec{G} = \frac{\partial f}{\partial \vec{x}} = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \dots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$ единичный вектор градиента $\vec{V} = \frac{1}{\|\vec{G}\|} \cdot \vec{G}$

4. Делаем шаг в направление антиградиента $\vec{z} = \vec{x} - h \cdot \vec{V}$ и вычисляем значение функции в точке $\vec{z}$ FZ=$f(\vec{z})$

5. Проверяем условие FZ < FX. Если условие выполняется, то за текущее приближение принимаем $\vec{z}$, $\vec{x} = \vec{z}$, FX=FZ и переходим на пункт 3.

6. Иначе, проверяем условие окончания h < ε, если оно выполняется, то выводим $\vec{x}$ и FX.

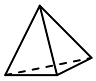
7. Если не выполняется, то уменьшаем параметр шага h=h/3 и переходим на пункт 3.



Симплексный метод

Симплексом в n-мерном пространстве называется выпуклый многоугольник с n+1 вершиной.

n=2  треугольник

n=3  тетраэдр

Алгоритм метода (n=2)

1. задаем функцию $f(\vec{x})$, начальное приближение $\vec{x}^0$, точность eps и длину грани h

2. вычисляем координаты вершин симплекса
$$\vec{x}^1 = \vec{x}^0 \quad \vec{x}^2 = \vec{x}^0 - \begin{bmatrix} h \\ 0 \end{bmatrix} \quad \vec{x}^3 = \vec{x}^0 + \begin{bmatrix} 0 \\ h \end{bmatrix}$$

3. вычисляем значения функции в вершинах $F1 = f(\vec{x}^1) \quad F2 = f(\vec{x}^2) \quad F3 = f(\vec{x}^3)$

4. определяем худшую вершину $F^{худ}$ и $\vec{x}^{худ}$

5. определяем координаты отраженной вершины. Она будет лежать на прямой исходящей из худшей вершины и проходящей через середину противоположной грани.
$\vec{x}^{ср} = 0.5*(\vec{x}^{ост1} + \vec{x}^{ост2}) \quad \vec{V} = \vec{x}^{ср} - \vec{x}^{худ} \quad \vec{x}^{отр} = \vec{x}^{худ} + 2*\vec{V}$, вычисляем значение функции в отраженной вершине $F^{отр} = f(\vec{x}^{отр})$

6. Сравниваем значения функции $F^{отр} < F^{худ}$ если да, то за новый симплекс принимаем симплекс с вершиной $\vec{x}^{отр}$ вместо $\vec{x}^{худ}$ и повторяем с пункта 3

7. Иначе проверяем условие окончания h<eps если условие не выполняется, то за $\vec{x}^0$ принимаем лучшую вершину последнего симплекса, уменьшаем длину грани
h=h/3 и повторяем с пункта 2.

8. Условие окончания выполняется. Выводим координаты и значение функции лучшей вершины.

пример:

$f(x_1,x_2)=15-4*x_1-13*x_2+x_1^2+3*x_2^2+x_1*x_2$,   h=1, eps=0.2, $\vec{x}^0 = \begin{bmatrix} 2.5 \\ 0 \end{bmatrix}$

№	x_1	x_2	$f(x_1,x_2)$	h	точки симплекса и примечания
1	2.5	0.0	11.25	1	симплекс т.1,2,3 т.1 худшая отражаем
2	1.5	0.0	11.25		
3	2.5	1.0	3.75		
4	1.5	1.0	2.75		удачно симплекс 2,3,4 т.2 худшая отражаем
5	2.5	2.0	2.25		удачно симплекс 3,4,5 т.3 худшая отражаем
6	1.5	2.0	0.25		удачно симплекс 4,5,6 т.4 худшая отражаем
7	2.5	3	6.75		неудачно 1>0.4 h=h/3 = 0.33
1	1.5	2.0	0.25	0.33	симплекс т.1,2,3, т.3 худшая отражаем
2	1.17	2.0	0.0289		
3	1.5	2.33	0.742		
4	1.17	1.67	0.299		удачно симплекс 1,2,4 т. 4 худшая
1	1.17	2.0	0.0289	0.11	симплекс т.1,2,3 т. 3 худшая отражаем
2	1.06	2.0	0.0036		
3	1.17	2.11	0.0839		
4	1.17	1.67	0.2995		неудача 0.11<0.2 ответ т. 2