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).

Процесс поиска экстремального значения иногда называют оптимизацией.

$f(\vec{x})$ называют оптимизируемой или целевой функцией

$\vec{x}$ называют вектором параметров оптимизации

В области определения функции может быть несколько экстремумов.

Все существующие методы многомерной оптимизации позволяют определить лишь один из экстремумов. Какой именно экстремум будет определён, зависит от выбора

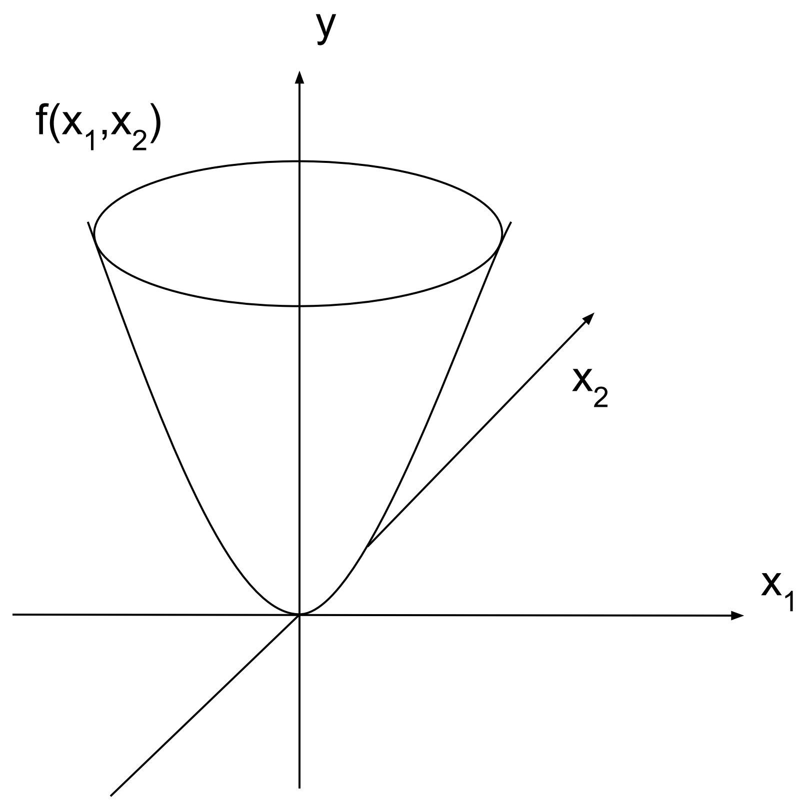
начального приближения: $\vec{x}^{(0)}$

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.

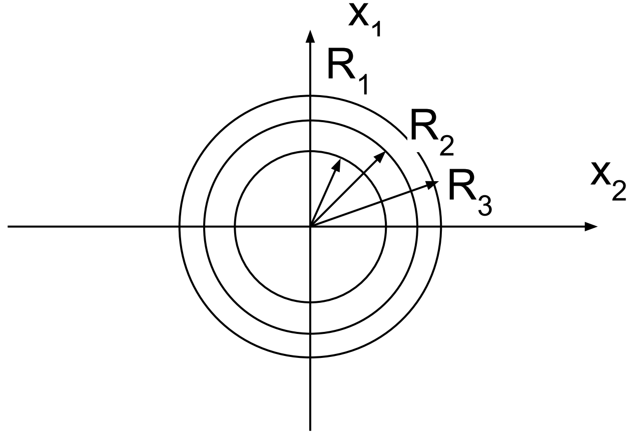
Рассечем заданную поверхность функции тремя плоскостями по уровням.

$$y_1 = x_1^2 + x_2^2 = 1 \qquad x_1 = \sqrt{1 - x_2^2}$$

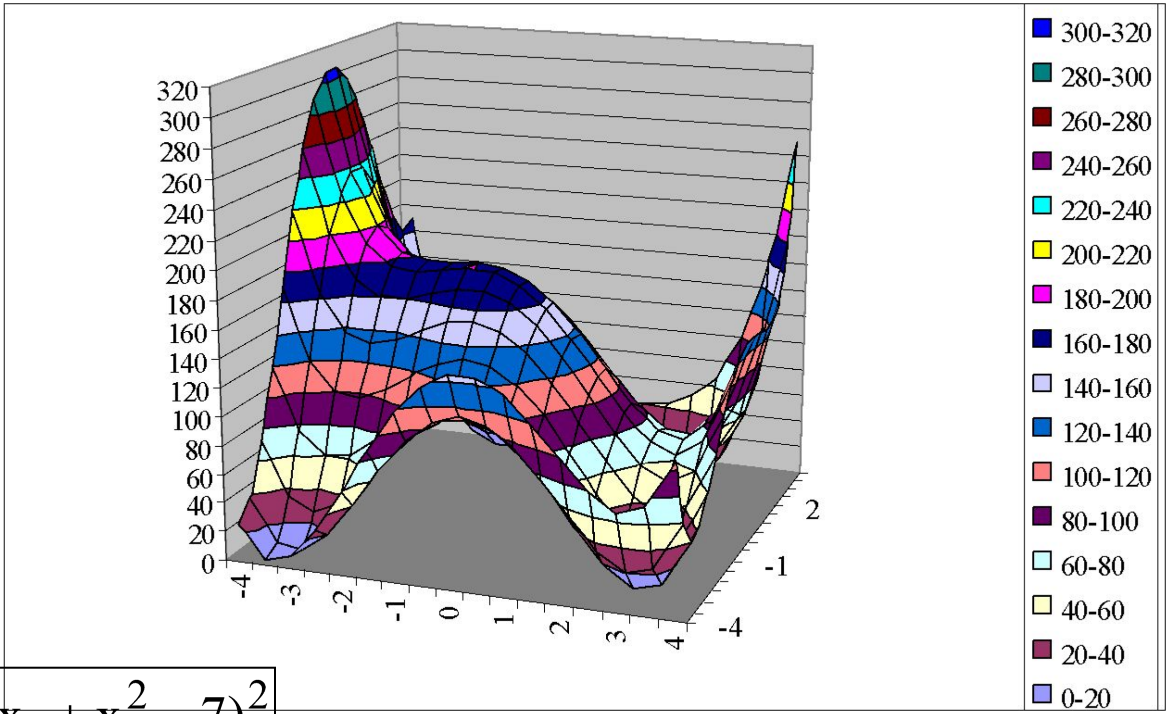
$$y_2 = x_1^2 + x_2^2 = 2 \qquad x_1 = \sqrt{2 - x_2^2}$$

$$y_3 = x_1^2 + x_2^2 = 3 \qquad x_1 = \sqrt{3 - x_2^2}$$

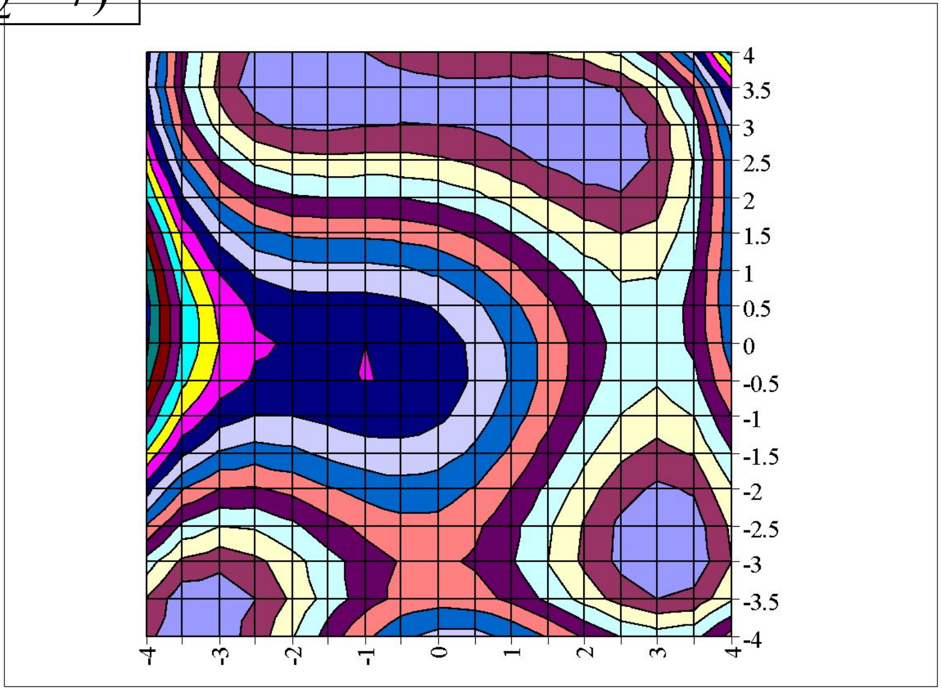
Определим зависимости x_1 от x_2 для соответствующих линий уровней.
Для заданной функции $f(x_1,x_2)$ линии уровней будут представлять окружности с соответствующими радиусами:
$R_1=\sqrt{1}=1$; $R_2=\sqrt{2}=1.41$; $R_3=\sqrt{3}=1.73$;



функция Химмельблау



$$f(x_1, x_2) = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2$$



Все методы многомерной оптимизации делятся на два класса

- Градиентные
- Безградиентные

Градиентный метод.

Градиентом называется вектор равный сумме произведений частных производных на соответствующие орты.

Орта – единичный связанный вектор, направленный вдоль координатной оси.

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ . \\ 0 \end{bmatrix} \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ . \\ 0 \end{bmatrix} \quad \ldots \quad \vec{e}_n = \begin{bmatrix} 0 \\ 0 \\ . \\ 1 \end{bmatrix}$$

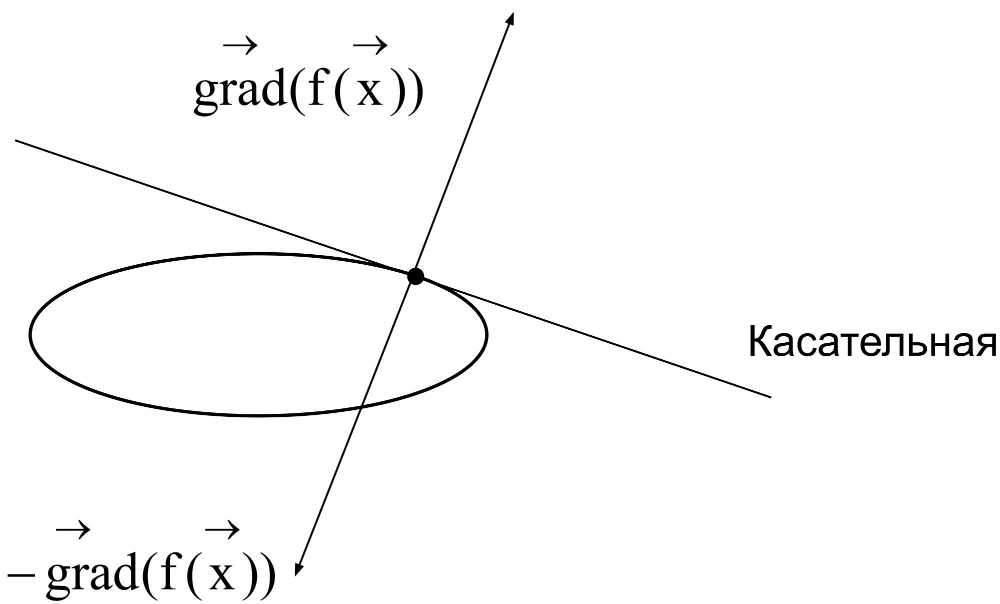
$$\vec{\operatorname{grad}}(f(\vec{x})) = \frac{\partial f}{\partial x_1} \cdot \vec{e}_1 + \frac{\partial f}{\partial x_2} \cdot \vec{e}_2 + \ldots\ldots\ldots + \frac{\partial f}{\partial x_n} \cdot \vec{e}_n = \begin{bmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \\ \vdots \\ \dfrac{\partial f}{\partial x_n} \end{bmatrix}$$

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.

3. Градиент перпендикулярен линии уровня.



Антиградиентом называется вектор, направленный в сторону противоположную градиенту.

Алгоритм

1. Дана функция n переменных $f(\vec{x})$, точность ε, параметр шага h, задаем начальное приближение $\vec{x} = \vec{x}^{(0)}$, вычисляем значение функции $FX = f(\vec{x})$

2. Вычисляем вектор градиента $\vec{G} = \operatorname{grad}(f(\vec{x}))$

И единичный вектор градиента $\vec{V} = \dfrac{1}{\|\vec{G}\|} \cdot \vec{G}$

3. Вычисляем новое приближение, делая шаг в направление антиградиента $\vec{z} = \vec{x} - h \cdot \vec{V}$

и вычисляем значение функции $FZ = f(\vec{z})$

4. Проверяем условие FZ < FX

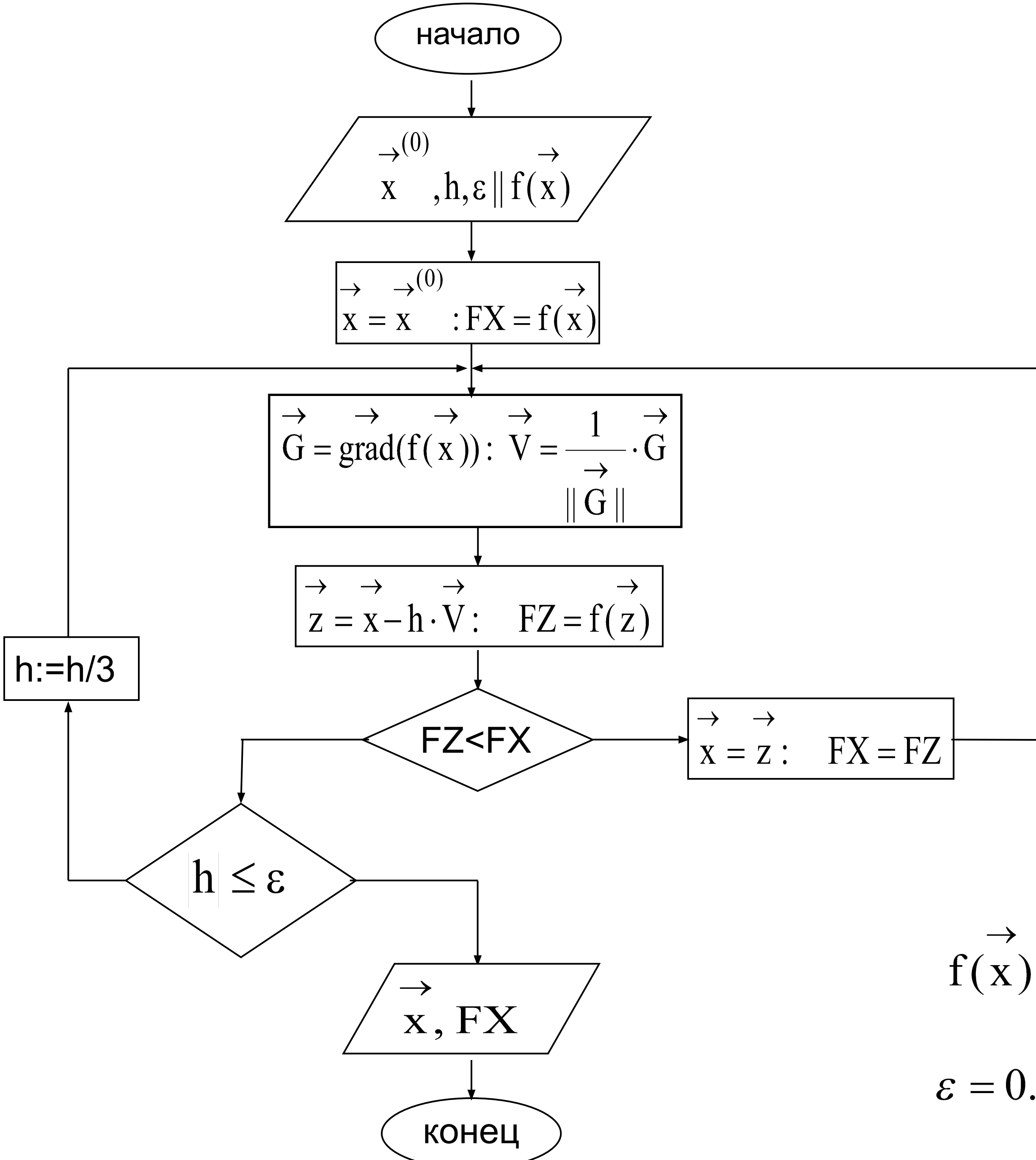
5. Если условие выполняется, то за начальное приближение принимаем $\vec{z}$ т.е.

$\vec{x} = \vec{z}$ и $FX = FZ$ переходим на пункт 2

6. Иначе, проверяем условие окончания h < ε

7. Если условие выполняется, то выводим $\vec{x}, f(\vec{x})$. Конец алгоритма

8. Если не выполняется, то уменьшаем параметр шага h=h/3 и переходим на пункт 2



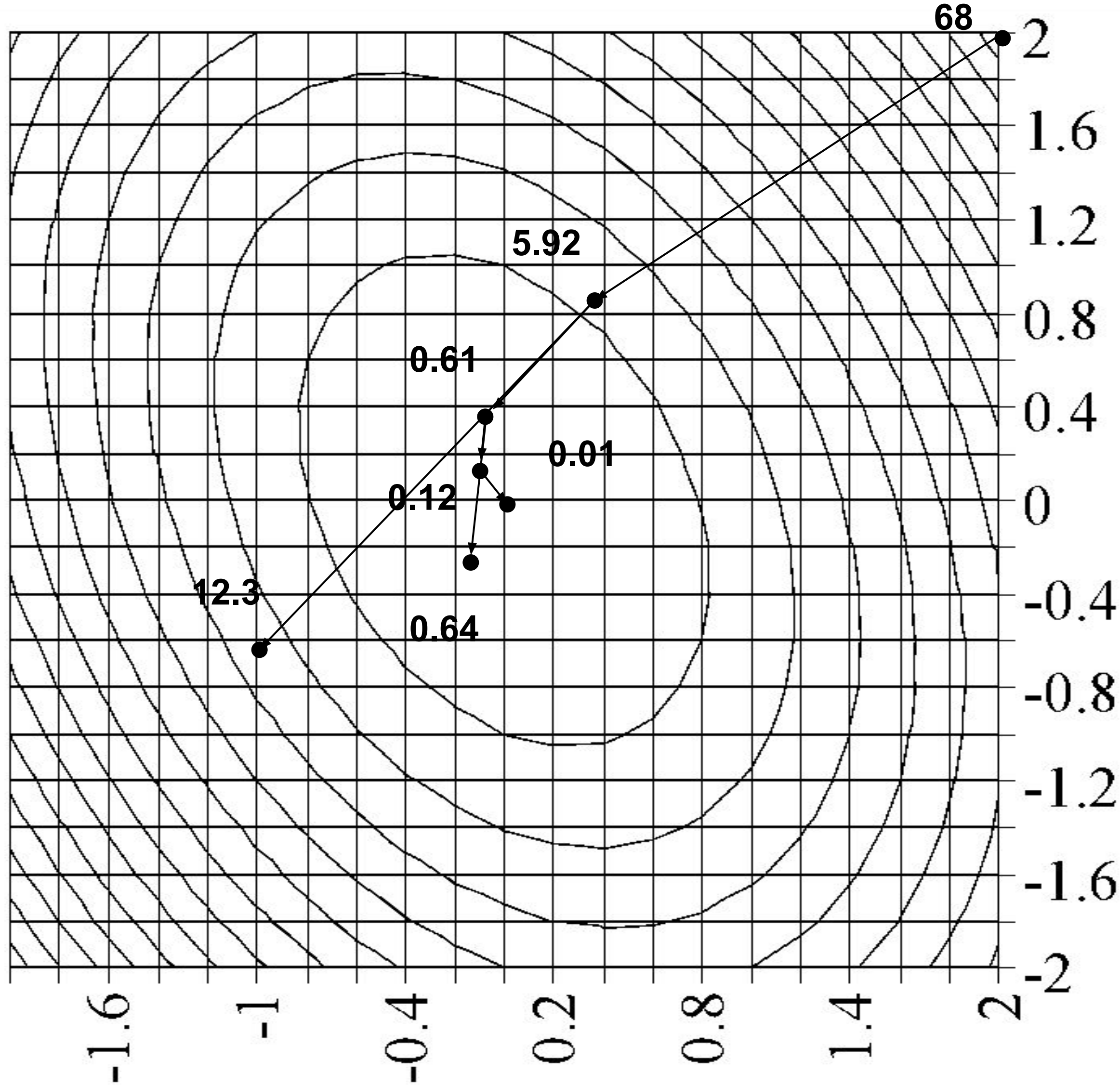
Пример:

$$f(\vec{x}) = 8x_1^2 + 4x_1x_2 + 5x_2^2$$

$$\varepsilon = 0.2 \quad h = 2 \quad \vec{x}^{(0)} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

$\vec{x}$	FX	$\vec{G}$	$\lVert \vec{G} \rVert$	$\vec{V}$	h	$\vec{z}$	FZ	Примечание
2 2	68.00	40 28	48.8	0.82 0.57	2	0.36 0.85	5.918	да
0.36 0.85	5.918	9.2 9.98	13.6	0.68 0.74	2	-0.99 -0.62	12.27	нет
0.36 0.85	5.918	9.2 9.98	13.6	0.68 0.74	0.66	-0.09 0.37	0.6090	да
-0.09 0.37	0.6090	0.1 3.33	3.34	0.03 1	0.66	-0.11 -0.29	0.6380	нет
-0.09 0.37	0.6090	0.1 3.33	3.34	0.03 1	0.22	-0.09 0.15	0.1230	да
-0.09 0.15	0.123	-0.89 1.11	1.42	-0.62 0.78	0.22	0.04 -0.02	0.0150	да
0.04 -0.02	0.0150	0.62 -0.06	0.62	1 -0.1	0.22	-0.17 0	0.2440	нет
0.04 -0.02	0.0150	0.62 -0.06	0.62	1 -0.1	0.07	-0.02 -0.02	0.0080	да
-0.02 -0.02	0.0080	-0.47 -0.27	0.54	-0.86 -0.5	0.07	0.04 0.02	0.0140	нет

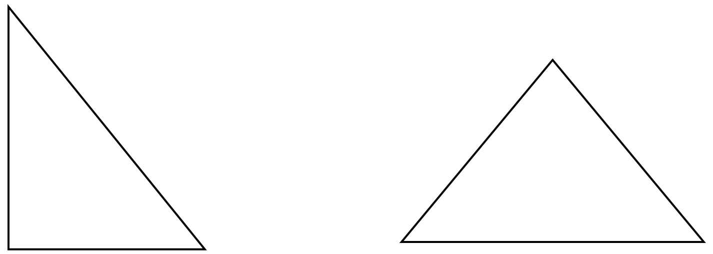
$$\vec{x}^{*} = \begin{bmatrix} -0.02 \\ -0.02 \end{bmatrix} \quad f(\vec{x}) = 0.0080$$


68
5.92
0.61
0.01
0.12
12.3
0.64
2
1.6
1.2
0.8
0.4
0
-0.4
-0.8
-1.2
-1.6
-2
-1.6
-1
-0.4
0.2
0.8
1.4
2

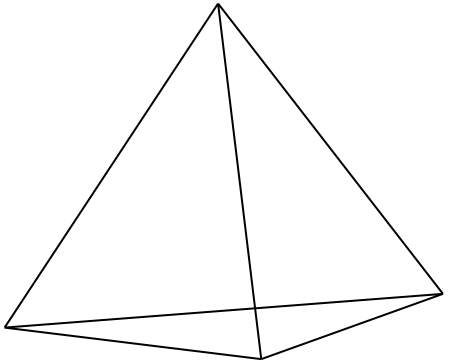
Симплексный метод

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

$n=2$ треугольник



$n=3$ тетраэдр



Алгоритм

1. Дана функция 2^x переменных $f(\vec{x})$, точность ε, параметр h, начальное приближение $\vec{x}^{(0)}$

2. Вычисляем координаты вершин симплекса

$$\vec{x}^{(1)} = \vec{x}^{(0)} \qquad \vec{x}^{(2)} = \vec{x}^{(0)} + \begin{bmatrix} h \\ 0 \end{bmatrix} \qquad \vec{x}^{(3)} = \vec{x}^{(0)} + \begin{bmatrix} 0 \\ h \end{bmatrix}$$

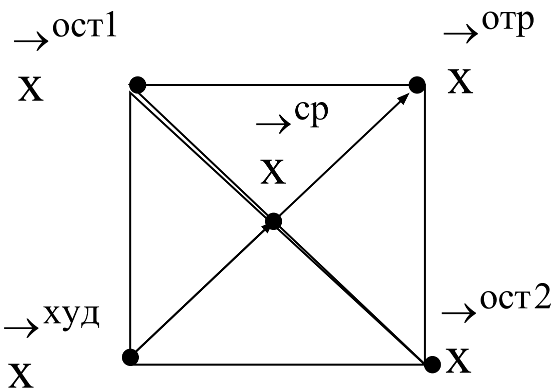
3. Вычисляем значения функции $F1 = f(\vec{x}^{(1)}) \quad F2 = f(\vec{x}^{(2)}) \quad F3 = f(\vec{x}^{(3)})$

4. Определяем худшую вершину $\vec{x}^{худ}, F^{худ}$ и координаты отраженной вершины, она будет лежать на прямой исходящей из худшей вершины и проходящей через середину противоположной грани

$$\vec{x}^{ср} = 0.5*(\vec{x}^{ост1} + \vec{x}^{ост2}); \quad \vec{V} = \vec{x}^{ср} - \vec{x}^{худ}$$

$$\vec{x}^{отр} = \vec{x}^{худ} + 2*\vec{V}; \quad \vec{x}^{отр} = \vec{x}^{ост1} + \vec{x}^{ост2} - \vec{x}^{худ}$$

$$F^{отр} = f(\vec{x}^{отр})$$



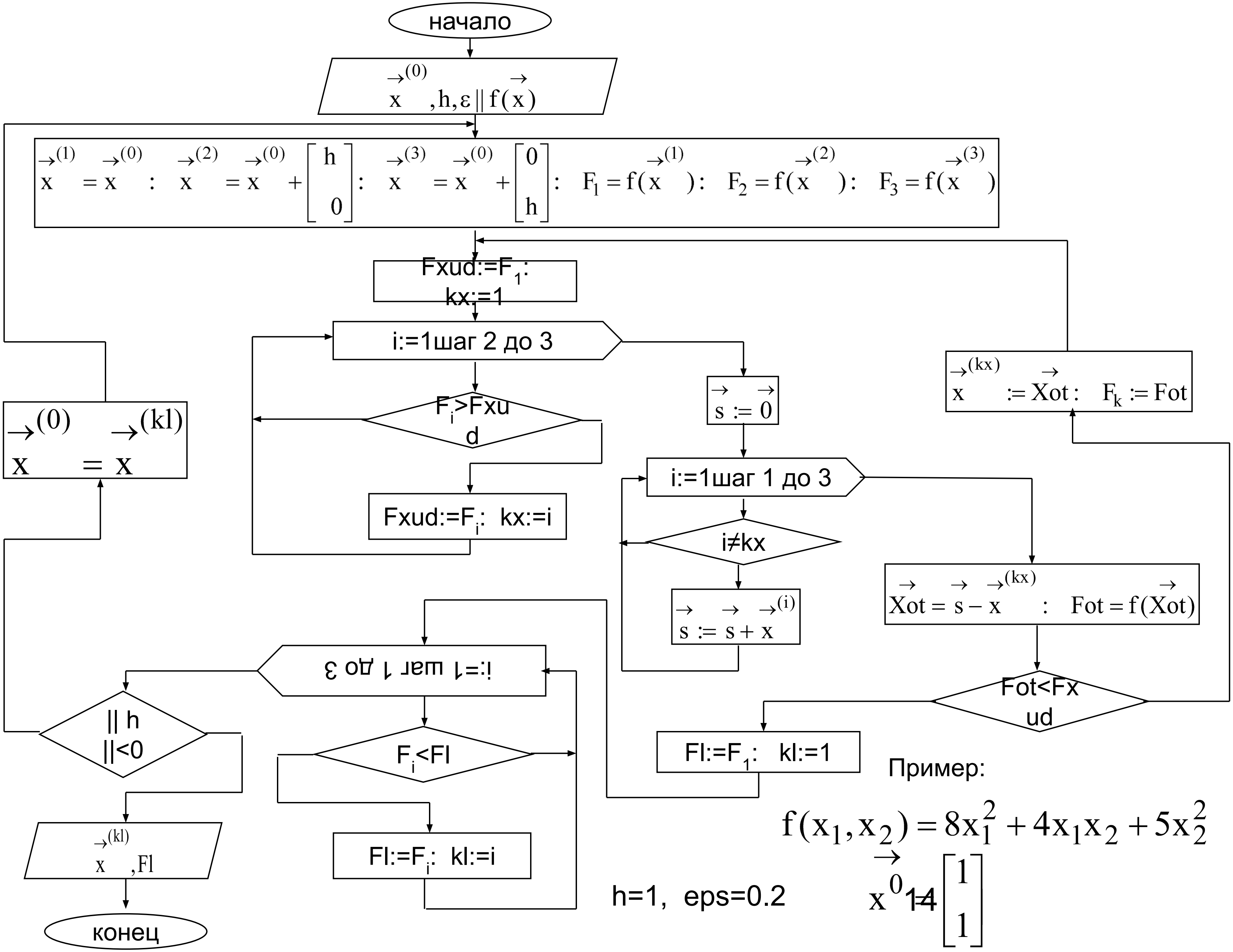
5. Сравниваем значения функции $F^{отр} < F^{худ}$

6. Условие выполняется. За новый симплекс принимаем симплекс с вершиной $\vec{x}^{отр}$ вместо $\vec{x}^{худ}$ и повторяем с пункта 3

7. Условие не выполняется. Проверяем условие окончания h< ε

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

9. Условие не выполняется. За $\vec{x}^{(0)}$ принимаем лучшую вершину последнего симплекса, уменьшаем длину грани h=h/3 и повторяем с пункта 2.



Пример:

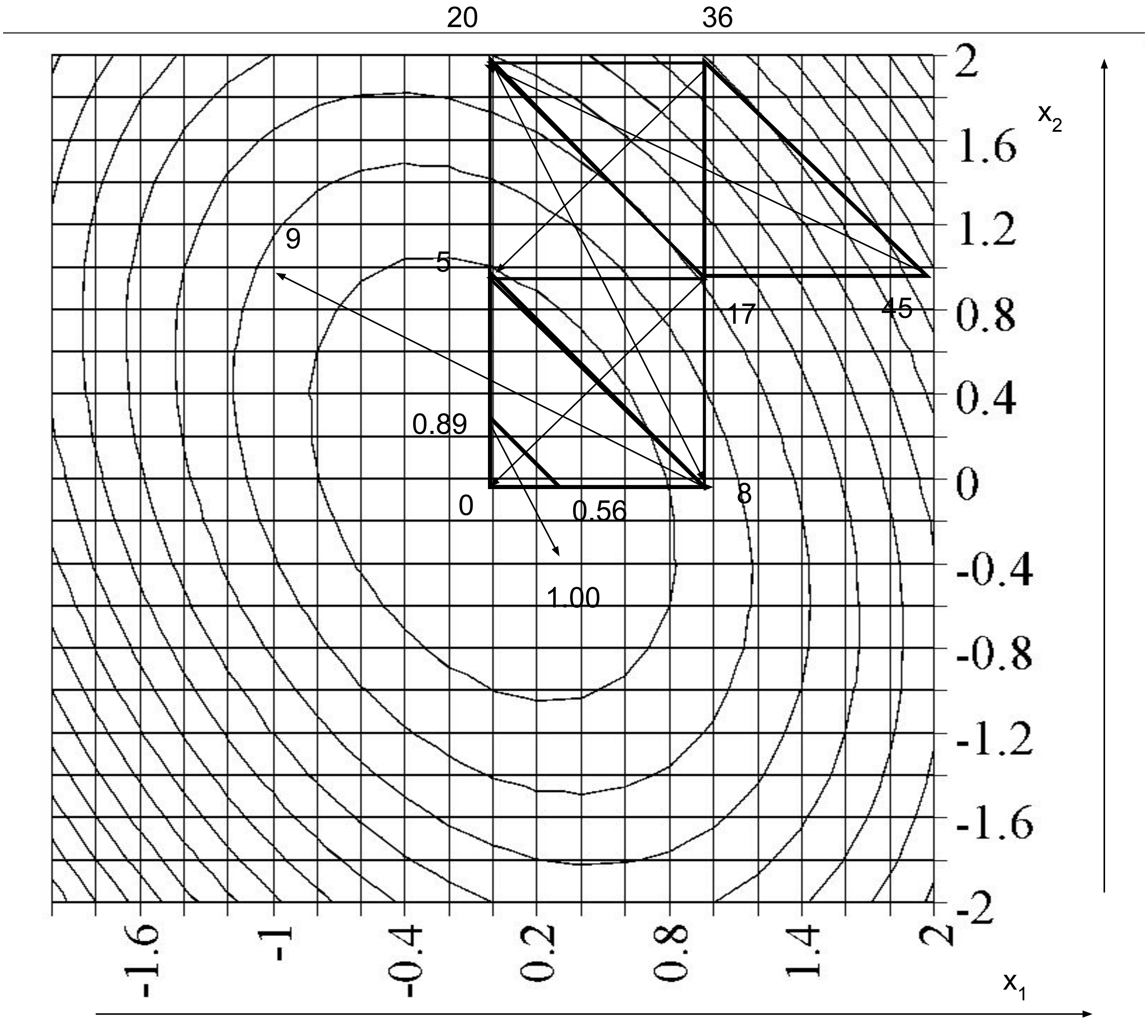
$$f(x_1, x_2) = 8x_1^2 + 4x_1x_2 + 5x_2^2$$

h=1, eps=0.2

$$\vec{x}^0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

№	x_1	x_2	$f(x_1,x_2)$	h	точки симплекса и примечания
1	1.00	1.00	17.0	1	симплекс т.1,2,3 т.2 худшая отражаем
2	2.00	1.00	45.0		
3	1.00	2.00	36.0		
4	0.00	2.00	20.0		удачно симплекс 1,3,4 т.3 худшая отражаем
5	0.00	1.00	5.00		удачно симплекс 1,4,5 т.4 худшая отражаем
6	1.00	0.00	8.00		удачно симплекс 1,5,6 т.1 худшая отражаем
7	0.00	0.00	0.00		удачно симплекс 5,6,7 т.6 худшая отражаем
8	-1.00	1.00	9.00		неудачно 1>0.2 h=h/3 = 0.33
1	0.00	0.00	0.00	0.33	симплекс т.1,2,3, т.2 худшая отражаем
2	0.333	0.00	0.889		
3	0.00	0.333	0.556		
4	-0.333	0.333	1.00		неудачно 0.33>0.2 h=h/3 = 0.11
1	0.00	0.00	0.00	0.11	симплекс т.1,2,3 т. 2 худшая отражаем
2	0.111	0.00	0.0988		
3	0.00	0.111	0.0617		

Ответ: $\vec{x} = \begin{bmatrix} 0.00 \\ 0.00 \end{bmatrix}$ $f(\vec{x}) = 0.000$


20
36
x_2
2
1.6
1.2
0.8
0.4
0
-0.4
-0.8
-1.2
-1.6
-2
9
5
17
45
0.89
0
8
0.56
1.00
-1.6
-1
-0.4
0.2
0.8
1.4
2
x_1