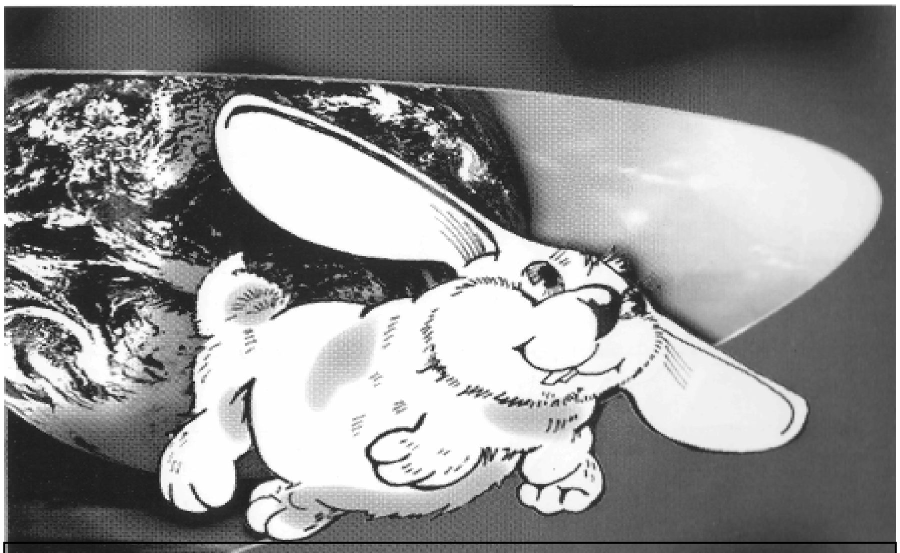
МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
КУРГАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
КАФЕДРА «ИНФОРМАТИКА»



# Многомерная оптимизация

Методические указания
для выполнения лабораторной работы
по курсу «Методы вычислений» и
«Теория системного анализа и принятия решений»
для студентов специальностей 010101, 280101

Курган 2010

**Кафедра**: «Информатика»

**Дисциплины**: «Методы вычислений», «Теория системного анализа и принятия решений» (специальности: 010101, 280101)

**Составили:** старший преподаватель Л.Г. Сысолятина;
старший преподаватель В.Я. Котликова

Утверждены на заседании кафедры «3» марта 2010 г.

Рекомендованы методическим советом университета «10» марта 2010 г.

СОДЕРЖАНИЕ

ЧИСЛЕННЫЕ МЕТОДЫ ПОИСКА БЕЗУСЛОВНОГО ЭКСТРЕМУМА

## *1. Общие положения*

Постановка задачи

Пусть дана функция $f(\boldsymbol{x})$, ограниченная снизу на множестве $R^n$ и имеющая непрерывные частные производные во всех его точках.

Требуется найти локальный минимум функции на множестве допустимых значений $\boldsymbol{X} \in R^n$, т.е. найти такую точку $\boldsymbol{x}^* \in X$, что

$$f(\boldsymbol{x}^*) = min\, f(\boldsymbol{x}),\ где\ \boldsymbol{x} \in X\ .$$

Применение необходимых и достаточных условий безусловного экстремума невозможно в тех случаях, когда целевая функция задана неявно либо не имеет непрерывных производных до второго порядка включительно. Кроме того, использование необходимого условия экстремума первого порядка (частные производные первого порядка равны нулю), т.е. градиент

$$\nabla f(\boldsymbol{x}) = (f'_{x1}, f'_{x2}, \ldots f_{xn}) = 0\ , \tag{1}$$

где $\boldsymbol{x} = (\,x_1, x_2, \ldots, x_n\,)$ - вектор $n$-мерного пространства $\boldsymbol{X}$, связано с решением системы $n$ в общем случае нелинейных алгебраических уравнений. Как правило, такую задачу приходится решать численно, причем трудоемкость решения сопоставима с трудоемкостью численного решения задачи оптимизации.

Численные методы оптимизации обеспечивают нахождение глобального экстремума только для выпуклых (вогнутых) функций, а в общем случае позволяют идентифицировать локальные экстремумы.

В данной работе рассмотрены некоторые численные методы поиска безусловного локального минимума функции

$$f(\boldsymbol{x}^*) = \min_{x \in R^n} f(\boldsymbol{x})\,. \tag{2}$$

Большинство численных методов оптимизации относится к классу итерационных. В соответствии с определенным правилом строится последовательность точек $\boldsymbol{x}^0, \boldsymbol{x}^1, \boldsymbol{x}^2, \boldsymbol{x}^3, \ldots$, обладающих свойством:

$$f(\boldsymbol{x}^{k+1}) < f(\boldsymbol{x}^k)\ ,\ k = 0, 1, 2\ldots \tag{3}$$

Преобразование точки $\boldsymbol{x}^k$ в точку $\boldsymbol{x}^{k+1}$ представляет собой итерацию. Очередная точка последовательности $\{\boldsymbol{x}^k\}$ определяется по формуле:

$$\boldsymbol{x}^{k+1} = \boldsymbol{x}^k + t_k \boldsymbol{d}^k \quad k = 0, 1, \ldots, \tag{4}$$

где $\boldsymbol{d}^k$ – вектор, определяющий направление перехода из точки $\boldsymbol{x}^k$ в точку $\boldsymbol{x}^{k+1}$, $t_k$ – величина «шага» (коэффициент изменения направления перехода).

Равенство (4) является общим, каждый конкретный алгоритм отличается правилами выбора векторов $\boldsymbol{d}^k$ и чисел $t_k$. Кроме того, для реализации итерационного процесса необходимо задать точку $\boldsymbol{x}^0$ (начальное приближение) и условие окончания алгоритма.

Начальная точка поиска $\boldsymbol{x}^0$ задается исходя из физического содержания решаемой задачи, в качестве критерия окончания счета, как правило, выбирается одно из условий:

$$\left\| x^{k+1} - x^k \right\| \le \varepsilon_1 \quad , \tag{5}$$

$$\left| \ f(\boldsymbol{x}^{k+1}) - f\left(\boldsymbol{x}^k\right) \right| \le \varepsilon_2 \ , \tag{6}$$

$$\left\| \nabla f\left(\boldsymbol{x}^k\right) \right\| \le \varepsilon_3 \quad , \tag{7}$$

где $\varepsilon_i$ , $i=\overline{1,3}$ – малое положительное число.

Иногда используются критерии, состоящие в одновременном выполнении двух из условий (5- 7) или всех трех условий.

Важной характеристикой численных методов оптимизации является сходимость. Метод $\left(\boldsymbol{x}^{k+1} = \boldsymbol{x}^k + t_k \boldsymbol{d}^k\right)$ сходится, если $\boldsymbol{x}^k \to \boldsymbol{x}^*$ при $k \to \infty$, где $\boldsymbol{x}^*$ - решение задачи (2).

Сходимость последовательности $\{\boldsymbol{x}^k\}$ зависит от характера функции $f(\boldsymbol{x})$ и от выбора начальной точки $\boldsymbol{x}^0$.

В зависимости от наивысшего порядка частных производных функции $f(\boldsymbol{x})$, используемых в алгоритмах решения задачи оптимизации, численные методы можно разделить на три группы.

1. Методы нулевого порядка (методы прямого поиска), использующие только информацию о значениях функции $f(\boldsymbol{x})$.

2. Методы первого порядка (градиентные методы), использующие информацию о первых производных функции $f(\boldsymbol{x})$.

3. Методы второго порядка, в которых требуется знание вторых производных функции $f(\boldsymbol{x})$.

Для каждой конкретной задачи вопрос о том, каким методом оптимизации воспользоваться, решается в зависимости от свойств минимизируемой функции (например, для минимизации недифференцируемой функции нельзя воспользоваться алгоритмом, предусматривающим возможность вычисления в произвольной точке градиента функции).

В специальной литературе представлены достаточно полные обзоры наиболее эффективных методов численного решения задач безусловной оптимизации.

Ниже рассмотрены лишь некоторые алгоритмы многомерной минимизации.

## *2. Численные методы многомерной минимизации*

### *2.1. Методы нулевого прядка или методы прямого поиска*

***1. Метод покоординатного спуска*** относится к методам прямого поиска и является простейшим методом поиска минимума функции многих переменных.

Пусть задана функция $F(x_1, x_2, \ldots, x_n)$, зависящая от $n$ переменных. Выберем начальное приближение $x_i^0\left(i=\overline{1,n}\right)$.

Зафиксируем значение всех координат кроме одной, например, первой $x_1^0$, тогда функция будет зависеть только от одной переменной. Используя какой-либо из методов одномерной минимизации (например, метод дихотомии, золотого сечения и т.д.), определим минимум функции одной переменной $f(x_1)$ и обозначим его через $x_1^1$.

Затем из новой точки делаем спуск по параллельному оси $x_2$ направлению, т.е. рассмотрим $f(x_2) = F\left(x_1^1, x_2, x_3^0, .., x_n^0\right)$. Далее будем оптимизировать исходную функцию $F$ поочередно по каждой координате. После того как завершится исследование по последней координате $x_n$, происходит возврат к первой и цикл повторяется.

Если в области минимума функция $F(\mathbf{x})$ достаточно гладкая, то процесс спуска по координатам будет линейно сходиться к минимуму. В сходящемся процессе с приближением к минимуму функции $F(\mathbf{x})$ расстояния между последовательными точками однокоординатных минимумов будут стремиться к нулю.

В качестве критериев окончания итерационного процесса координатного спуска выбираются условия:

$\left|x_i^k - x_i^{k+1}\right| < \varepsilon_i$, где $\varepsilon_i$ $i = \overline{1,n}$ - заданная допустимая абсолютная погрешность определения местоположения минимума по $i$-й координате.

Рассмотрим геометрическую трактовку спуска по координатам. На рис. 1 показаны линии уровня функции двух переменных $F(x_1, x_2)$, минимум которой находится в точке $\left(x_1^*, x_2^*\right)$.

Выберем нулевое приближение $x_1^0, x_2^0$ (точка А на рис. 1). Из точки А произведем поиск минимума вдоль направления оси $x_1$, зафиксировав значение координаты $x_2 = x_2^0$. Тогда функция будет зависеть только от одной переменной $x_1$. Используя какой-либо метод одномерной минимизации, определим минимум функции одной переменной (точка В ($x_1^1, x_2^0$) на рис. 1).

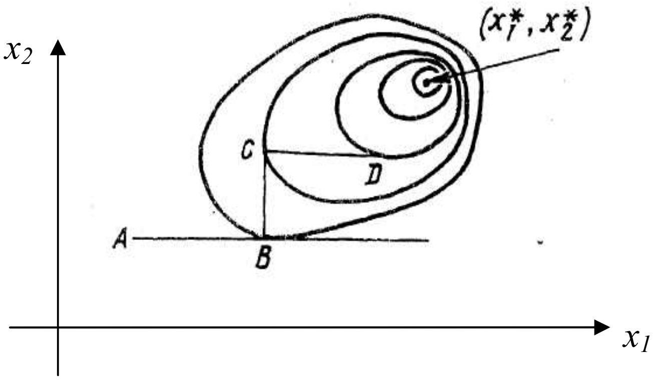



Рис. 1. Определение минимума функции одной переменной

Мы сделали шаг из точки A($x_1^0, x_2^0$) в точку B($x_1^1, x_2^0$) по направлению, параллельному оси $x_1$; на этом шаге значение функции уменьшилось. Далее из точки B сделаем спуск по направлению, параллельному оси $x_2$, зафиксировав значение координаты $x_1$: $x_1 = x_1^1$.

Рассмотрим функцию $f(x_2) = F(x_1^1, x_2)$, найдем ее минимум, обозначив его через $x_2^1$. Таким образом, второй шаг приводит нас в точку C($x_1^1, x_2^1$) и завершает цикл спусков. Повторяя циклы, мы приближаемся к оптимальной точке ($x_1^*, x_2^*$).

***2. Метод конфигураций*** (метод Хука-Дживса), также относящийся к методам прямого поиска, представляет собой комбинацию *исследующего поиска* вокруг базисной точки и *ускоряющего поиска по образцу*.

*Исследующий поиск* начинается в начальной точке $\boldsymbol{x}^0 = \left(x_1^0, x_2^0, \ldots, x_n^0\right)$, называемой старым базисом. В качестве направлений поиска выберем координатные направления, для каждого координатного направления зададим начальную длину шага $h_i\left(i = \overline{1,n}\right)$. Зафиксируем первое координатное направление и увеличим на выбранный шаг значение соответствующей переменной. Если значение функции в пробной точке меньше значения функции в исходной точке, шаг считается удачным, иначе необходимо значение соответствующей переменной уменьшить на выбранный шаг с последующей проверкой поведения функции. После перебора всех координат *исследующий поиск* завершается. Полученная точка называется *новым базисом.* В процессе поиска шаги $h_i$ будут уменьшаться до величин $\varepsilon_i$. Поиск завершается, когда величины шагов становятся достаточно малыми, условие $h_i < \varepsilon_i$ $i = \overline{1,n}$ используется для остановки алгоритма (завершения вычислений).

Рассмотрим алгоритм подбробнее.

Вычислим значение функции в начальной базисной точке $y^0 = f\left(x_1^0, x_2^0, \ldots, x_n^0\right)$. Зафиксировав первое координатное направление, сделаем шаг в сторону увеличения переменной $x_1^0$ : $x_1 = x_1^0 + h_1$ и определим $y^1 = f\left(x_1, x_2^0, .., x_n^0\right)$. Если $y^1 < y^0$, то шаг сделан в правильном направлении и координата $x_1^0$ заменяется на координату $x_1$, в противном случае делаем шаг в противоположном направлении $x_1 = x_1^0 - h_1$ и снова вычисляем $y^1 = f\left(x_1, x_2^0, .., x_n^0\right)$. Если $y^1 < y^0$, то фиксируем новое значение первой координаты. При $y^1 \geq y^0$ В обоих случаях значение первой координаты остается неизменным: $x_1 = x_1^0$.

Далее изменяем переменную $x_2^0$ на величину шага $h_2$ в прямом $x_2 = x_2^0 + h_2$ и, если требуется, в обратном направлении $x_2 = x_2^0 - h_2$. Аналогично рассматриваем и преобразуем все $n$ переменных (координат).

Таким образом, в результате *исследующего поиска* будет найдена некоторая точка $\boldsymbol{x}=\left(x_1, x_2, \ldots, x_n\right)$ - новая базисная точка.

Если *исследующий поиск* с данными величинами шагов не привел к уменьшению функции ($\boldsymbol{x}^0 = \boldsymbol{x}$), следует произвести уменьшение шагов ($h_i=h_i/\alpha$, $i=\overline{1,n}$, $\alpha$=const) и повторить процедуру *исследующего поиска*. Поиск заканчивается в том случае, когда $h_i \le \varepsilon_i$ , $i=\overline{1,n}$ .

Если *исследующий поиск* привел к уменьшению функции ($\boldsymbol{x}^0 \ne \boldsymbol{x}$), то следует перейти к следующему этапу – *поиску по образцу*.

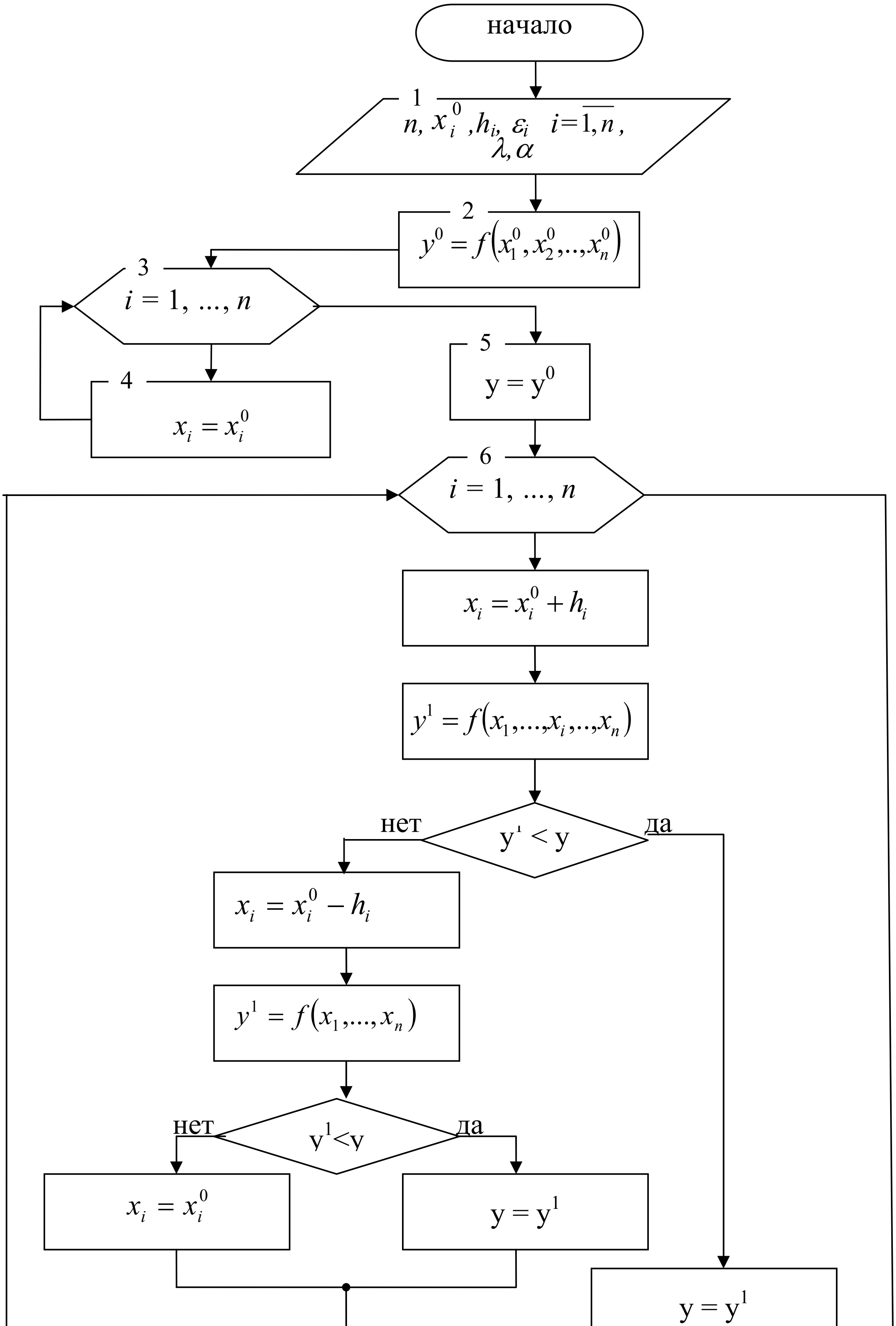
*Поиск по образцу* заключается в движении из полученной базисной точки вдоль прямой, соединяющей эту точку с предыдущей базисной точкой, т.е. от старого базиса к новому. Величина ускоряющего шага задается ускоряющим множителем $\lambda$.

Координаты новой точки $\boldsymbol{x^*}$ определяются формулами $x_i^* = x_i + \lambda\left(x_i - x_i^0\right),\ \ i=\overline{1,n}$ .

Если результат $y^* = f\left(x_1^*, x_2^*, .., x_n^*\right) < y = f\left(x_1, x_2, .., x_n\right)$, то ускоряющий шаг удачен и следует сделать еще один или несколько таких шагов.

В случае $y^* \ge y$ перейдем к *исследующему поиску* в точке $\boldsymbol{x}$ (аналогично тому, как это было сделано в точке $\boldsymbol{x^0}$).

На рис. 2 приведен вариант алгоритма описанного метода.


начало
1
$n, x_i^0, h_i, \varepsilon_i \quad i=\overline{1,n},$
$\lambda, \alpha$
2
$y^0 = f\left(x_1^0, x_2^0, \ldots, x_n^0\right)$
3
$i = 1, \ldots, n$
4
$x_i = x_i^0$
5
$y = y^0$
6
$i = 1, \ldots, n$
$x_i = x_i^0 + h_i$
$y^1 = f\left(x_1, \ldots, x_i, \ldots, x_n\right)$
нет
$y^1 < y$
да
$x_i = x_i^0 - h_i$
$y^1 = f\left(x_1, \ldots, x_n\right)$
нет
$y^1 < y$
да
$x_i = x_i^0$
$y = y^1$
$y = y^1$
16
C10

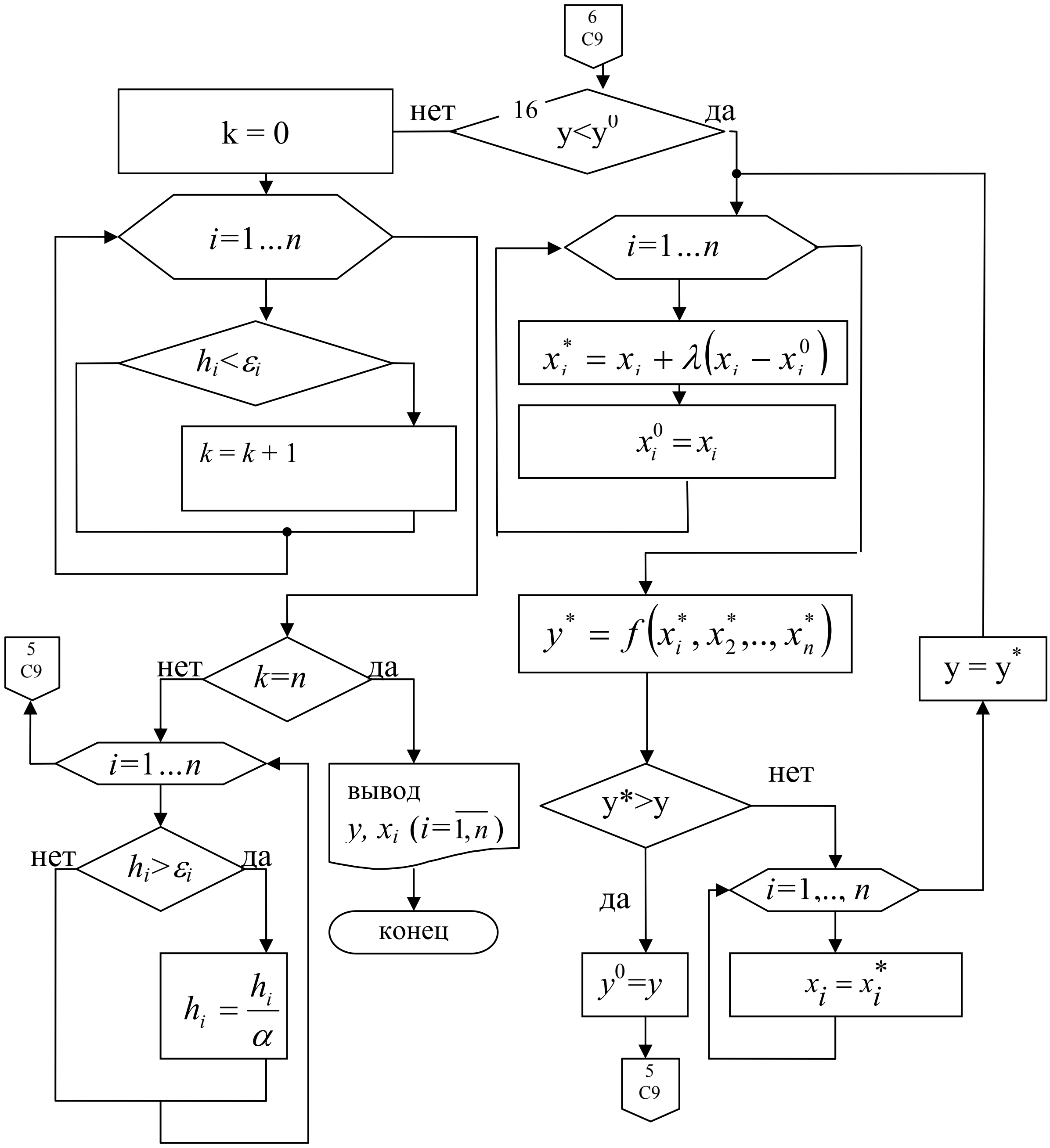



Рис. 2. Блок-схема алгоритма метода конфигураций

*Пример 1*. Найти минимум функции $f(\mathbf{x}) = 16(x_1-2)^2 + (x_2-3)^2$, *где* $\mathbf{x}=(x_1 , x_2)$.

Рассмотрим численное решение задачи методом Хука – Дживса.

Анализ условия задачи.

Определим линии уровня заданной функции. Если $f(\mathbf{x}) = const = c$, то уравнение примет вид: $16(x_1-2)^2 + (x_2-3)^2 = c,\ (c>0)$. Разделив обе части равенства на $c$, получим уравнение $\frac{(x_1-2)^2}{c/16} + \frac{(x_2-3)^2}{c} = 1$, которое описывает семейство эллипсов с полуосями $a = \sqrt{c/16}$, $b = \sqrt{c}$. При $c=16$ получим $a = 1$, $b=4$. Центр эллипса имеет координаты $x_1=2,\ x_2=3$. Это и есть точка локального минимума. Покажем это, используя метод Хука–Дживса.

Для этого зададим начальную точку минимума $\mathbf{x^0}=(x_1^0, x_2^0)$, число $\varepsilon > 0$ для остановки алгоритма, начальные значения шагов $h_1$, $h_2$ по направлениям $x_1$ и

$x_2$ соответственно (итерационный процесс закончится, если выполняются условия $h_1<\varepsilon$ и $h_2<\varepsilon$), ускоряющий множитель $\lambda > 0$, коэффициент уменьшения шага $\alpha > 1$.

Для решения данной задачи зададим старый базис $\boldsymbol{x}^0 =(3;5)$, $\varepsilon=0,4$, $h_1 = 1$, $h_2 = 3$, $\lambda =1$, $\alpha =2$ (хотя в алгоритме можно использовать и одинаковую величину шага, т.е. $h_1 = h_2$).

$f(\boldsymbol{x}) = 16(x_1-2)^2 +(x_2-3)^2$

$\boldsymbol{x}^0 =(3;5)$ $x_1^0 = 3$, $x_2^0 = 5$ $y_0 = f(\boldsymbol{x}^0) = f(3;5)=16\cdot1^2+2^2 = 16+4 = 20.$

Осуществим *исследующий поиск* по направлению оси $x_1$, т.е.

$x_1 = x_1^0 + h_1 = 3+1 = 4$

$y_1 =f(x_1 , x_2^0) = f(4; 5) = 16\cdot(4-2)^2+(5-3)^2 =64+4 = 68.$ Т.к. $y_1 > y_0$, то шаг выполнен неудачно. Рассмотрим противоположное направление, а именно,

$x_1 = x_1^0 - h_1 = 3-1 = 2$

$y_1 =f(x_1 , x_2^0) = f(2; 5) = 16\cdot(2-2)^2+(5-3)^2 =4.$

Т.к. $y_1 < y_0$, то шаг выбран удачно. Осуществим *исследующий поиск* по направлению оси $x_2$, учитывая, что $x_1^0 =2$ и $x_2^0 =5$, т.е.

$x_2 = x_2^0 + h_2 = 5+2= 7$

$y_2 =f( x_1^0 ,x_2) = f(2; 7) = 16\cdot(2-2)^2+(7-3)^2 =16.$

Т.к. $y_2 > y_1 = 4$, то шаг выполнен неудачно. Рассмотрим противоположное направление, а именно,

$x_2 = x_2^0 - h_2 = 5-2= 3$

$y_2 =f( x_1^0 ,x_2) = f(2; 3) = 16\cdot(2-2)^2+(3-3)^2 =0.$

Т.к. $y_2 < y_1 =4$, то шаг выбран удачно. Новый базис (новое приближение к минимуму) $\boldsymbol{x}^1 = (2; 3)$, $f(\boldsymbol{x}^1) = 0$.

Осуществим *поиск по образцу*, двигаясь по направлению от старого базиса к новому, т.е. из точки $\boldsymbol{x}^0$ через точку $\boldsymbol{x}^1$ к минимуму. Величина ускоряющего шага задается ускоряющим множителем $\lambda =1$, поиск новой точки осуществляем по формуле

$\boldsymbol{x}_1^* = \boldsymbol{x}^1 +\lambda\cdot(\boldsymbol{x}^1 - \boldsymbol{x}^0) =(2;3)+1\cdot((2;3) - (3;5)) = (2;3)+(-1;-1) = (1;1).$

$f(\boldsymbol{x}_1^*)=f(1;1)=16\cdot1^2+2^2 = 20 > f(\boldsymbol{x}^1) = 0,$

т.е. *поиск по образцу* неудачен, поэтому возвращаемся к новому базису, продолжаем *исследующий поиск* с уменьшенным шагом, т.к. точность ($h_1<\varepsilon$ и $h_2<\varepsilon$) еще не достигнута. Уменьшаем шаг вдвое $h_1=1/2 = 0,5$ и $h_2= 3/2=1,5$.

Осуществим *исследующий поиск* по направлению оси $x_1$ для базиса $\boldsymbol{x}^1 = (2; 3)$, где $x_1^0 = 2$, $x_2^0 = 3$, $f(\boldsymbol{x}^1) = 0$, т.е.

$x_1 = x_1^0 + h_1 = 2+0,5 = 2,5$

$y_1 =f( x_1 , x_2^0 ) = f(2,5; 3) = 16\cdot(2,5-2)^2+(3-3)^2 =4 > f(\boldsymbol{x}^1) = 0$, т.е. шаг неудачен.

Рассмотрим противоположное направление, а именно,

$x_1 = x_1^0 - h_1 = 2-0,5 = 1,5$

$y_1 = f(x_1, x_2^0) = f(1{,}5;\ 3) = 16\cdot(1{,}5-2)^2+(3-3)^2 = 4 > f(\boldsymbol{x}^1) = 0$, т.е. шаг неудачен.

Осуществим *исследующий поиск* по направлению оси $x_2$, учитывая, что $x_1^0 = 2$ и $x_2^0 = 3$, т.е.

$x_2 = x_2^0 + h_2 = 3+1{,}5 = 4{,}5$

$y_2 = f(x_1^0, x_2) = f(2;\ 4{,}5) = 16\cdot(2-2)^2+(4{,}5-3)^2 = 2{,}25 > f(\boldsymbol{x}^1) = 0,$

т.е. шаг выполнен неудачно. Рассмотрим противоположное направление, а именно,

$x_2 = x_2^0 - h_2 = 3-1{,}5 = 1{,}5$

$y_2 = f(x_1^0, x_2) = f(2;\ 1{,}5) = 16\cdot(2-2)^2+(1{,}5-3)^2 = 2{,}25 > f(\boldsymbol{x}^1) = 0,$ т.е. шаг неудачен.

Как видим, базис не изменился, т.е. $\boldsymbol{x}^2 = (2;\ 3)$ , где $x_1^0 = 2, x_2^0 = 3, f(\boldsymbol{x}^2) = 0.$

Осуществим *поиск по образцу*, двигаясь по направлению от старого базиса к новому, т.е. из точки $\boldsymbol{x}^1$ через точку $\boldsymbol{x}^2$ к минимуму. Величина ускоряющего шага задается ускоряющим множителем $\lambda = 1$, поиск новой точки осуществляем по формуле

$$\boldsymbol{x}_2^* = \boldsymbol{x}^2 + \lambda\cdot(\boldsymbol{x}^2 - \boldsymbol{x}^1) = (2;3)+1\cdot((2;3)-(2;3)) = (2;3).$$

$$f(\boldsymbol{x}_2^*) = f(2;3) = 16\cdot 0^2+0^2 = 0 = f(\boldsymbol{x}^2) = 0,$$

т.е. *поиск по образцу* неудачен, поэтому возвращаемся к новому базису, продолжаем *исследующий поиск* с уменьшенным шагом, т.к. точность ($h_1<\varepsilon$ и $h_2<\varepsilon$) еще не достигнута. Уменьшаем шаг еще раз вдвое $h_1 = 0{,}5/2 = 0{,}25$ и $h_2 = 1{,}5/2 = 0{,}75$.

Осуществим *исследующий поиск* по направлению оси $x_1$ для базиса $\boldsymbol{x}^2 = (2;\ 3)$ , где $x_1^0 = 2, x_2^0 = 3$, $f(\boldsymbol{x}^2) = 0,$ т.е.

$x_1 = x_1^0 + h_1 = 2+0{,}25 = 2{,}25$

$y_1 = f(x_1, x_2^0) = f(2{,}25;\ 3) = 16\cdot(2{,}25-2)^2+(3-3)^2 = 1 > f(\boldsymbol{x}^2) = 0,$ т.е. шаг неудачен.

Рассмотрим противоположное направление, а именно,

$x_1 = x_1^0 - h_1 = 2-0{,}25 = 1{,}75$

$y_1 = f(x_1, x_2^0) = f(1{,}75;3) = 16\cdot(1{,}75-2)^2+(3-3)^2 = 1 > f(\boldsymbol{x}^2) = 0,$ т.е. шаг неудачен.

Осуществим *исследующий поиск* по направлению оси $x_2$, учитывая, что $x_1^0 = 2$ и $x_2^0 = 3$, т.е.

$x_2 = x_2^0 + h_2 = 3+0{,}75 = 3{,}75$

$y_2 = f(x_1^0, x_2) = f(2;\ 3{,}75) = 16\cdot(2-2)^2+(3{,}75-3)^2 = 0{,}56 > f(\boldsymbol{x}^2) = 0,$

т.е. шаг выполнен неудачно. Рассмотрим противоположное направление, а именно,

$x_2 = x_2^0 - h_2 = 3-0{,}75 = 2{,}25$

$y_2=f(x_1^0, x_2)=f(2;2{,}25)=16\cdot(2-2)^2+(2{,}25-3)^2=0{,}56>f(\boldsymbol{x}^2)=0,$ т.е. шаг неудачен.

Как видим, базис не изменился, т.е. $\boldsymbol{x}^3=(2;\ 3)$, где $x_1^0=2, x_2^0=3$, $f(\boldsymbol{x}^3)=0.$

Осуществим *поиск по образцу*, двигаясь по направлению от старого базиса к новому, т.е. из точки $\boldsymbol{x}^2$ через точку $\boldsymbol{x}^3$ к минимуму. Величина ускоряющего шага задается ускоряющим множителем $\lambda=1$, поиск новой точки осуществляем по формуле

$$\boldsymbol{x}_3^*=\boldsymbol{x}^3+\lambda\cdot(\boldsymbol{x}^3-\boldsymbol{x}^2)=(2;3)+1\cdot((2;3)-(2;3))=(2;3).$$

$$f(\boldsymbol{x}_3^*)=f(2;3)=16\cdot0^2+0^2=0\ =\ f(\boldsymbol{x}^2)=0,$$

т.е. *поиск по образцу* неудачен, поэтому возвращаемся к новому базису, продолжаем *исследующий поиск* с уменьшенным шагом, т.к. точность ($h_1<\varepsilon$ и $h_2<\varepsilon$) еще не достигнута. Уменьшаем шаг ($h_1=0{,}25<0.4$) $h_2=0{,}75/2=0{,}375.$ Несложно убедиться в том, что результаты поиска не изменятся, т.е. базис останется прежним $\boldsymbol{x}^4=(2;\ 3)$ и точность поиска достигнута $h_1=0{,}25<\varepsilon=0.4$ и $h_2=0{,}75/2=0{,}375<\varepsilon=0.4.$

*Ответ*: локальный минимум достигнут в точке $\boldsymbol{x}=(2;\ 3)$, что и требовалось показать. Все результаты отображены в табл. 1 и на графике рис. 3, где удачные шаги обозначены сплошной линией.

Таблица 1

| График движения точек ($x_1$, $x_2$) к минимуму функции $f(\boldsymbol{x})=16(x_1-2)^2+(x_2-3)^2$ | | | | | | |
|---|---|---|---|---|---|---|
| $x_1$ | $x_2$ | n шага | | f(x) | $h_1$ | $h_2$ |
| 3 | 5 | 1 | | 20 | 1 | 3 |
| 4 | 5 | 2 | неудача | 68 | | |
| 2 | 5 | 3 | удача | 4 | | |
| 2 | 7 | 4 | неудача | 16 | | |
| ***2*** | ***3*** | ***5*** | ***удача*** | 0 | | |
| 1 | 1 | 6 | неудача | 20 | | |
| 2,5 | 3 | 7 | неудача | 4 | 0,5 | 1,5 |
| 1,5 | 3 | 8 | неудача | 4 | | |
| 2 | 4,5 | 9 | неудача | 2,25 | | |
| 2 | 1,5 | 10 | неудача | 2,25 | | |
| 2,25 | 3 | 11 | неудача | 1 | 0,25 | 0,75 |
| 1,75 | 3 | 12 | неудача | 1 | | |
| 2 | 3,75 | 13 | неудача | 0,56 | | |
| 2 | 2,25 | 14 | неудача | 0,56 | | |
| 2,25 | 3 | 15 | неудача | 1 | 0,25 | 0,375 |
| 1,75 | 3 | 16 | неудача | 1 | | |
| 2 | 3,375 | 17 | неудача | 0,0625 | | |
| 2 | 2,625 | 18 | неудача | 0,0625 | | |
| Точность достигнута | | | | | $h_1<0{,}4$ | $h_2<0{,}4$ |
| Локальный минимум $x_1=2$, $x_2=3$ | | | | | $f(x_1,x_2)=0$ | |

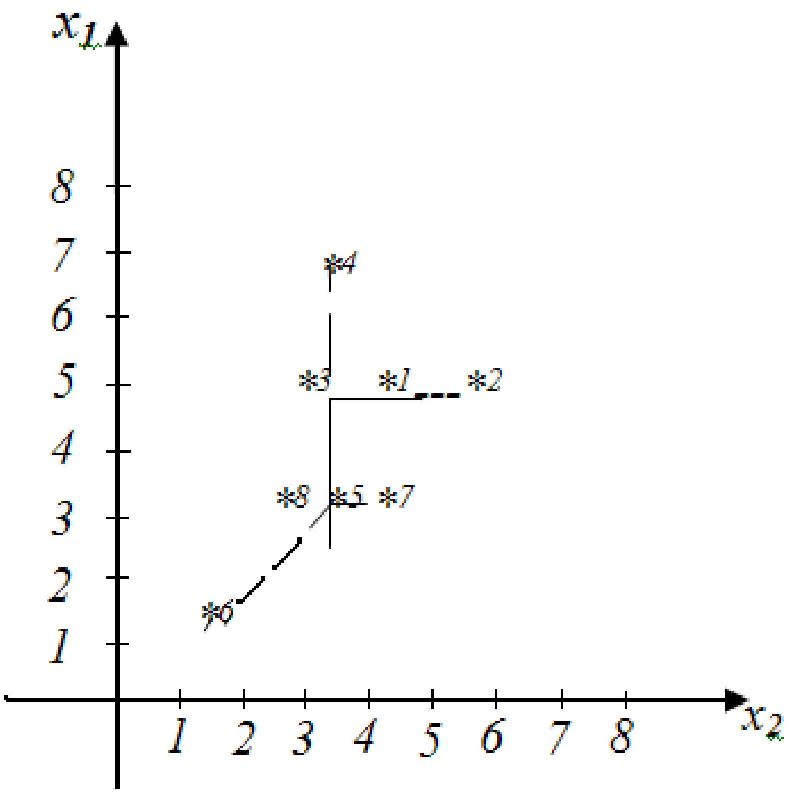



Рис. 3. График движения точек $(x_1,x_2)$ функции $f(\mathbf{x})$ к минимуму (здесь 1,2,3… - номера шагов поиска) по методу Хука-Дживса

## 2.2. Методы первого порядка или градиентные методы

Направление градиента является направлением наискорейшего возрастания функции, поэтому направлением наискорейшего убывания (спуска) является направление $-\nabla f(\boldsymbol{x})$.

Основой всех градиентных методов минимизации функции является итерационный процесс:

$$\boldsymbol{x}^{k+1} = \boldsymbol{x}^k - t_k \nabla f(\boldsymbol{x}^k), \qquad (8)$$

где $\boldsymbol{x}^k$ – текущее приближение к решению $\boldsymbol{x}^*$,

$\nabla f(\boldsymbol{x})$ – градиент функции $f(\boldsymbol{x})$ в $k$-ой точке $\boldsymbol{x}^k$,

$t_k$ – длина «шага» (коэффициент изменения направления перехода).

Начальное приближение $\boldsymbol{x}^0$ задается пользователем. Длину «шага» можно выбрать различными способами.

*1*. В ***методе градиентного спуска с постоянным шагом*** величина шага задается пользователем и остается постоянной до тех пор, пока функция убывает в точках последовательности. Если условие убывания $f(\boldsymbol{x}^{k+1}) < f(\boldsymbol{x}^k)$ не выполняется, следует уменьшать длину шага до тех пор, пока условие не будет выполнено.

Геометрическая интерпретация метода градиентного спуска с постоянным шагом для $n=2$ представлена на рис. 4.

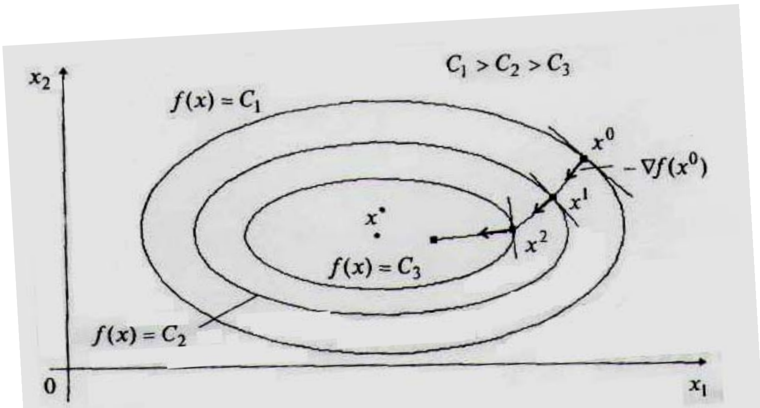



Рис. 4. Геометрическая интерпретация метода градиентного спуска с постоянным шагом

*2*. В ***методе наискорейшего градиентного спуска*** значение параметра $t_k$ выбирается таким образом, чтобы минимизировать функцию

$$h(t) = f\left(\boldsymbol{x}^k - t\nabla f\left(x^k\right)\right). \qquad (9)$$

Градиентный метод поиска экстремума (8) с выбором шага по способу (9) называется методом наискорейшего спуска. Такой метод требует наименьшего количества итераций, но на каждом шаге приходится решать задачу поиска минимума функции одной переменной *h(t)* (9).

Для случая функции двух переменных метод наискорейшего спуска имеет следующую геометрическую интерпретацию (рис. 5): луч, идущий от точки $\boldsymbol{x}^k$ к точке $\boldsymbol{x}^{k+1}$, перпендикулярен к линии уровня функции $f(x_1, x_2)$, проходящей через точку $\boldsymbol{x}^k$, и касается линии уровня, проходящей через точку $\boldsymbol{x}^{k+1}$. Таким образом, на плоскости наискорейший спуск осуществляется по двум взаимно перпендикулярным направлениям.

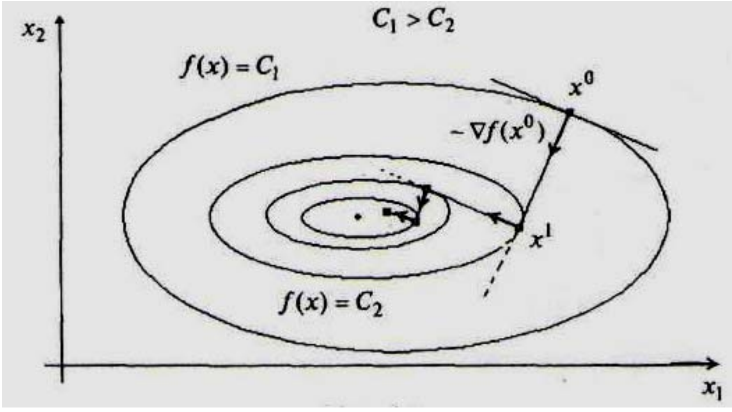



Рис.5. Метод наискорейшего спуска двух переменных

Вычисления в соответствии с градиентным методом завершаются при нахождении точки, в которой градиент становится нулевым вектором, т.е. $\left\|\nabla f\left(\boldsymbol{x}^k\right)\right\| \le \varepsilon_1$, где $\varepsilon_1$ – заданное малое положительное число, или $k < M$, где $M$ – предельное число итераций, или при *двукратном одновременном* выполнении двух неравенств $\left\|\boldsymbol{x}^{k+1} - \boldsymbol{x}^k\right\| \le \varepsilon_2$, $\left|f\left(\boldsymbol{x}^{k+1}\right) - f\left(\boldsymbol{x}^k\right)\right| \le \varepsilon_2$, где $\varepsilon_2$ – малое положительное число.

Процесс (8) обеспечивает сходимость последовательности $\{\boldsymbol{x}^k\}$ к стационарной точке $\boldsymbol{x}^*$, где $\nabla f(\boldsymbol{x}^*) = 0$.

Если функция $f(\boldsymbol{x})$ выпукла, то найденная стационарная точка будет решением задачи минимизации. В противном случае необходимо провести дополнительное исследование целевой функции в окрестности найденной стационарной точки.

На рис. 6 приведена укрупненная блок-схема алгоритма метода градиентного спуска с постоянным шагом. (Напомним, что $\boldsymbol{x}=(x_1, x_2, \ldots, x_n)$ – это вектор). Для реализации алгоритма, необходимо предварительно найти градиент минимизируемой функции в произвольной точке $\nabla f(x) = \left(\dfrac{\partial f(x)}{\partial x_1}, \ldots, \dfrac{\partial f(x)}{\partial x_n}\right)$.

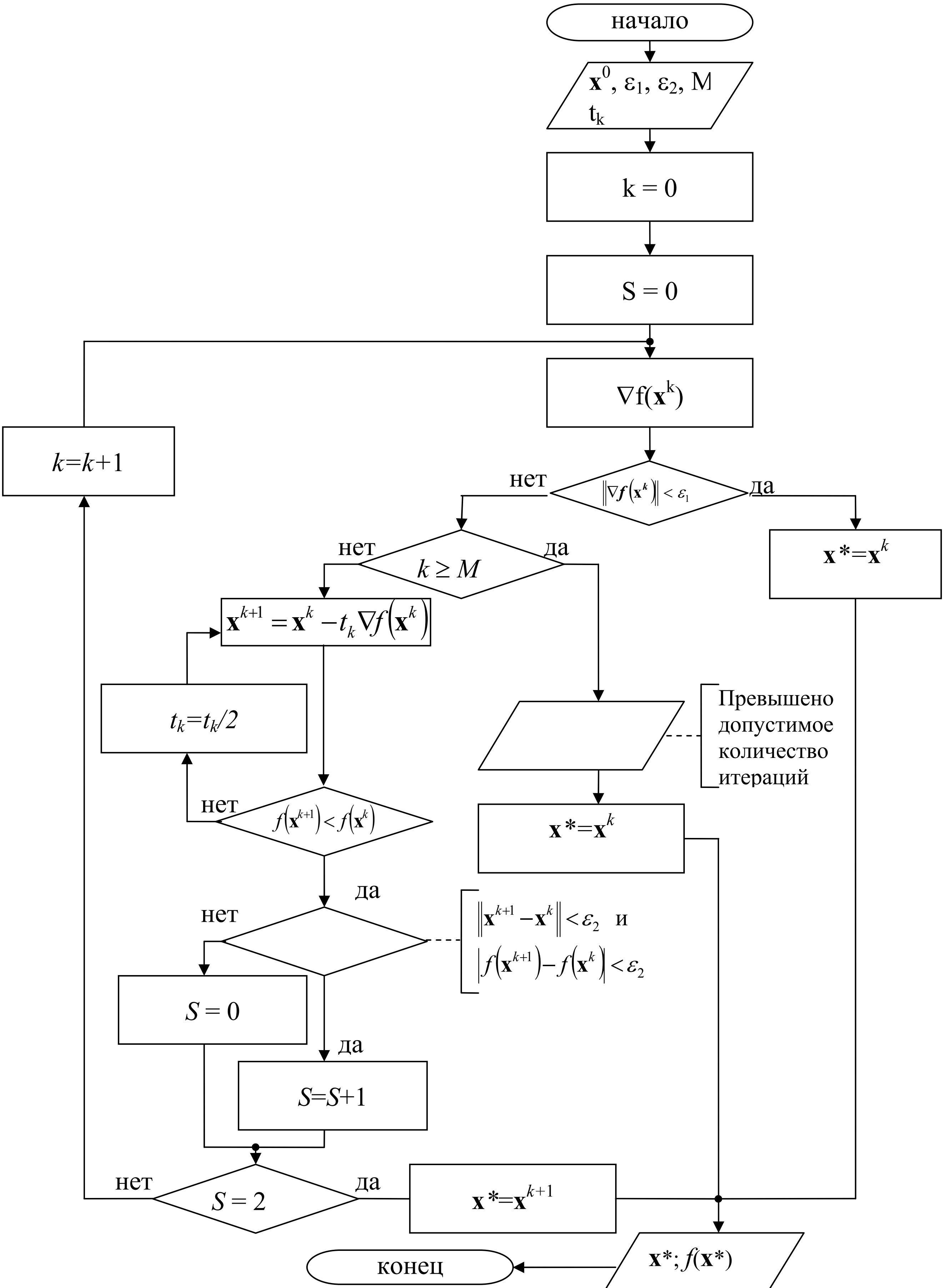



Рис.6. Блок-схема алгоритма метода градиентного спуска с постоянным шагом (s=2 - это *двукратное одновременное* выполнение условия окончания расчета)

*Пример 2.* Найти локальный минимум функции $f(\mathbf{x}) = 3x^2 + 2xy + 4y^2$, где $\mathbf{x}=(x, y)$.

Рассмотрим численное решение задачи методом градиентного спуска с постоянным шагом.

Введем обозначения:

$к$ – счетчик количества итераций, начальное значение $к=0$;

$M$ – предельное число итераций, пусть $M=10$;

$\mathbf{x}^к=(x_к, y_к)$ – вектор $к$-го приближения к минимуму функции;

$\mathbf{d}^к$ – вектор направления спуска для $к$–й итерации, $\mathbf{d}^к = \nabla f(\mathbf{x}^к)$;

$t_к$ – коэффициент изменения направления перехода;

$\varepsilon_1$, $\varepsilon_2$ – заданные малые положительные числа, влияющие на процесс окончания построения итерационной последовательности $\{ \mathbf{x}^к \}$, а именно,

$\| \nabla f(\mathbf{x}^к) \| < \varepsilon_1$ или *двукратное одновременное* выполнение двух неравенств $\| \mathbf{x}^{к+1} - \mathbf{x}^к \| < \varepsilon_2$ и $|f(\mathbf{x}^{к+1}) - f(\mathbf{x}^к)| < \varepsilon_2$ ,

где $\nabla f(\mathbf{x}) = (f_x', f_y')$ – градиент функции $f(\mathbf{x})$, т.е. $f_x' = 6x+2y$; $f_y' = 2x+8y$. Матрица Гессе $H(\mathbf{x}^0) = \begin{pmatrix} f''_{xx} & f''_{xy} \\ f''_{yx} & f''_{yy} \end{pmatrix} = \begin{pmatrix} 6 & 2 \\ 2 & 8 \end{pmatrix}$. $H(\mathbf{x}^0)>0$, т.к. $\Delta_1 = 6>0$ и

$\Delta_2 = \begin{vmatrix} 6 & 2 \\ 2 & 8 \end{vmatrix} = 44 > 0$

*Первая итерация*

$f(\mathbf{x}) = 3x^2 + 2xy + 4y^2$. Зададим $\mathbf{x}^0 = (0,5;\ 1)$; $\varepsilon_1 = 0,15$ $\varepsilon_2=0,4$; $M=10$; $f(\mathbf{x}^0) = 3\cdot 0,5^2 + 2\cdot 0,5\cdot 1 + 4\cdot 1^2 = 5,75$. Найдем градиент в произвольной точке $\mathbf{x}=(x, y)$, т.е. $\nabla f(\mathbf{x}) = (f_x', f_y') = (6x+2y;\ 2x+8y)$.

Пусть $к = 0$.

1. $\nabla f(\mathbf{x}^0) = (6\cdot 0,5+2\cdot 1;\ 2\cdot 0,5+8\cdot 1) = (5;\ 9)$.

2. Вычислим норму градиента по формуле $\| \nabla f(\mathbf{x}^к) \| = \sqrt{(f_x'(x^0))^2 + (f_y'(x^0))^2} = \sqrt{(5)^2 + (9)^2} = \sqrt{106} = 10.3 > \varepsilon_1 = 0.15$

3. $к = 0 < M = 10$ – итерационный процесс не окончен.

4. Пусть $t_0 = 0,5$.

5. Вычислим $\mathbf{x}^1 = \mathbf{x}^0 - t_0 \cdot \nabla f(\mathbf{x}^0) = (0,5;1) - 0,5\cdot(5;9) = (-2;-3,5)$;

$f(\mathbf{x}^1) = 3\cdot(-2)^2 + 2\cdot(-2)\cdot(-3,5) + 4\cdot(-3,5)^2 = 12+14+12,25 = 38,25$. Видим, что $f(\mathbf{x}^1) > f(\mathbf{x}^0)$. Вывод: условие $f(\mathbf{x}^{к+1}) < f(\mathbf{x}^к)$ для $к = 0$ не выполнилось, поэтому уменьшим $t_0$ вдвое, уменьшая $t_к$ несколько раз, добьемся выполнения этого условия. Таким образом, $t_0=0,25$. Найдем новое значение $\mathbf{x}^1 = \mathbf{x}^0 - t_0 \cdot \nabla f(\mathbf{x}^0) = (0,5;1) - 0,25\cdot(5;9) = (-0,75;-1,25)$.

Тогда $f(\mathbf{x}^1) = 3\cdot(-0,75)^2 + 2\cdot(-0,75)\cdot(-1,25) + +4\cdot(-1,25)^2 = 1,69+1,88+6,25 = 9,82$. Сравнив результаты, видим, что $f(\mathbf{x}^1) > f(\mathbf{x}^0)$, т.е. условие $f(\mathbf{x}^{к+1}) < f(\mathbf{x}^к)$ для $к = 0$ не выполнилось, уменьшаем $t_0$ вдвое: $t_0 = 0,125$, тогда $\mathbf{x}^1 = \mathbf{x}^0 - t_0 \cdot \nabla f(\mathbf{x}^0) = (0,5;1) - 0,125\cdot(5;9) = (-0,125;-0,125)$;

$f(\mathbf{x}^1) = 3\cdot(-0,125)^2 + 2\cdot(-0,125)\cdot(-0,125) + 4\cdot(-0,125)^2 = 0,141$. Видим, что $f(\mathbf{x}^1) < f(\mathbf{x}^0)$.

6. Проверим один из критериев окончания итерационных процессов, а именно, *одновременное* выполнение двух неравенств $\| \mathbf{x}^{к+1} - \mathbf{x}^к \| < \varepsilon_2$ и

$|f(\boldsymbol{x}^{к+1}) - f(\boldsymbol{x}^{к})| < \varepsilon_2$. Получим $\boldsymbol{x}^1 - \boldsymbol{x}^0 = (-0,125;-0,125)-(0,5;1)=(-0,625;-1,125)$;
$\| \boldsymbol{x}^1 - \boldsymbol{x}^0 \| = \sqrt{(-0.625)^2 + (-1.125)^2} = \sqrt{0.391 + 1.266} = \sqrt{1.657} = 1.29 > \varepsilon_2 = 0.4$.

$|f(\boldsymbol{x}^1) - f(\boldsymbol{x}^0)| = |0,14-5,75| = 5,61 > \varepsilon_2$. Вывод: итерационный процесс не закончен.

*Вторая итерация*

$к=1$, начальное приближение сейчас $\boldsymbol{x}^1 = (-0,125;-0,125)$.

1. $\nabla f(\boldsymbol{x}^1) = (6\cdot(-0,125)+2\cdot(-0,125);\ 2\cdot(-0,125)+8\cdot(-0,125)) = (-1;-1,25)$;
2. $\| \nabla f(\boldsymbol{x}^1) \| = \sqrt{(f_x'(x^1))^2 + (f_y'(x^1))^2} = \sqrt{(-1)^2 + (-1.25)^2} = \sqrt{2.56} = 1.6 > \varepsilon_1 = 0.15$
3. $к=1<M=10$, значит, расчеты можно продолжить.
4. Зададим $t_1 = 0,125$, которое найдено последним.
5. Вычислим $\boldsymbol{x}^2 = \boldsymbol{x}^1 - t_1 \cdot \nabla f(\boldsymbol{x}^1) = (-0,125;-0,125) - 0,125\cdot(-1;-1,25) =$ $=(0;0,031)$; $f(\boldsymbol{x}^2) = 3\cdot(-0)^2+2\cdot(-0)\cdot(0,031)+4\cdot(0,031)^2=0,004$. Видим, что $f(\boldsymbol{x}^2) < f(\boldsymbol{x}^1)$.
6. Проверим критерий окончания итерационных процессов: $\boldsymbol{x}^2 - \boldsymbol{x}^1 =$ $=(0;0,031)-(-0,125;-0,125)=(0,125;0,156)$;

$\| \boldsymbol{x}^2 - \boldsymbol{x}^1 \| = \sqrt{(0.125)^2 + (0.156)^2} = \sqrt{0.04} = 0.2 < \varepsilon_2 = 0.4$.

$|f(\boldsymbol{x}^2) - f(\boldsymbol{x}^1)| = |0,004-0.14| = 0.136 < \varepsilon_2 = 0.4$. Вывод: надо проверить *двукратное одновременное* выполнение двух неравенств $\| \boldsymbol{x}^{к+1} - \boldsymbol{x}^{к} \| < \varepsilon_2$ и $|f(\boldsymbol{x}^{к+1}) - f(\boldsymbol{x}^{к})| < \varepsilon_2$, поэтому продолжим расчеты.

*Третья итерация*

$к=2$, начальное приближение сейчас $\boldsymbol{x}^2 = (0;\ 0,031)$; $f(\boldsymbol{x}^2) = 0,004$.

1. $\nabla f(\boldsymbol{x}^2) = (6\cdot(0)+2\cdot(0,031);\ 2\cdot(0)+8\cdot(0,031)) = (0,062;0,248)$;
2. $\| \nabla f(\boldsymbol{x}^2) \| =$
$\sqrt{(f_x'(x^2))^2 + (f_y'(x^2))^2} = \sqrt{(0.062)^2 + (0.248)^2} = \sqrt{0.065} = 0.25 > \varepsilon_1 = 0.15$
3. $к=2<M=10$, значит, расчеты можно продолжить.
4. Зададим $t_2 = 0,125$.
5. Вычислим $\boldsymbol{x}^3 = \boldsymbol{x}^2 - t_2 \cdot \nabla f(\boldsymbol{x}^2) = (0;0,031) - 0,125\cdot(0,062;0,248) =$ $=(-0,007;0)$; $f(\boldsymbol{x}^3) = 3\cdot(-0,007)^2+2\cdot(-0,007)\cdot(0)+4\cdot(0)^2=0,0001$.
Видим, что $f(\boldsymbol{x}^3) < f(\boldsymbol{x}^2)$.
6. Проверим критерий окончания итерационных процессов: $\boldsymbol{x}^3 - \boldsymbol{x}^2 =$ $=(-0,007;0)-(0;0,031)=(-0,007;-0,031)$;

$\| \boldsymbol{x}^3 - \boldsymbol{x}^2 \| = \sqrt{(-0.007)^2 + (-0.031)^2} = \sqrt{0.0059} = 0.076 < \varepsilon_2 = 0.4$. $|f(\boldsymbol{x}^3) - f(\boldsymbol{x}^2)| =$ $= |0,0001-0.004| = 0,0039 < \varepsilon_2 = 0.4$. Вывод: наблюдаем *двукратное одновременное* выполнение двух неравенств $\| \boldsymbol{x}^{к+1} - \boldsymbol{x}^{к} \| < \varepsilon_2$ и $|f(\boldsymbol{x}^{к+1}) - f(\boldsymbol{x}^{к})| < \varepsilon_2$, поэтому расчет окончен. Найдена точка $\boldsymbol{x}^3 = =(-0,007;0)$; $f(\boldsymbol{x}^3) = 0,0001$.

Проведем *анализ* точки $\boldsymbol{x}^3$.

Функция $f(\boldsymbol{x}) = 3x^2 + 2xy + 4y^2$ является дважды дифференцируемой. Проведем проверку достаточных условий (необходимые выполнены, а именно, $\nabla f(\boldsymbol{x}^3) = (6\cdot(-0,007)+2\cdot(0);\ 2\cdot(-0,007)+8\cdot(0)) = (-0,0042;-0,0014)$; $\| \nabla f(\boldsymbol{x}^2) \| =$ $= \sqrt{(-0.0042)^2 + (-0.0014)^2} = 0.0047 < \varepsilon_1 = 0.15$). Для этого проанализируем матрицу

Гессе $H(\boldsymbol{x}^к)$ = $\begin{pmatrix} f''_{xx} & f''_{xy} \\ f''_{yx} & f''_{yy} \end{pmatrix} = \begin{pmatrix} 6 & 2 \\ 2 & 8 \end{pmatrix}$. Она постоянна и является положительно определенной для всех стационарных точек, в том числе и для точки $\boldsymbol{x}^3=(-0,007;0)$, т.е. $H(\boldsymbol{x}^к)>0$, т.к. оба ее угловых минора $\Delta_1 = 6>0$ и $\Delta_2=\begin{vmatrix} 6 & 2 \\ 2 & 8 \end{vmatrix} = 44 > 0$. Следовательно, $\boldsymbol{x}^3= =(-0,007;0)$ – найденное приближение точки локального минимума, а значение функции $f(\boldsymbol{x}^3) =0,0001$ – найденное приближение к значению функции в точке локального минимума. Но так как $H(\boldsymbol{x}^к) >0$ всегда, то функция $f(\boldsymbol{x}) = 3x^2 +2xy+4y^2$ строго выпукла, значит, $\boldsymbol{x}^3$ и $f(\boldsymbol{x}^3)$ соответствуют приближению точки глобального минимума и функции ее наименьшего значения на плоскости *ОХУ*.

Блок-схема алгоритма метода наискорейшего градиентного спуска представлена на рис. 7.

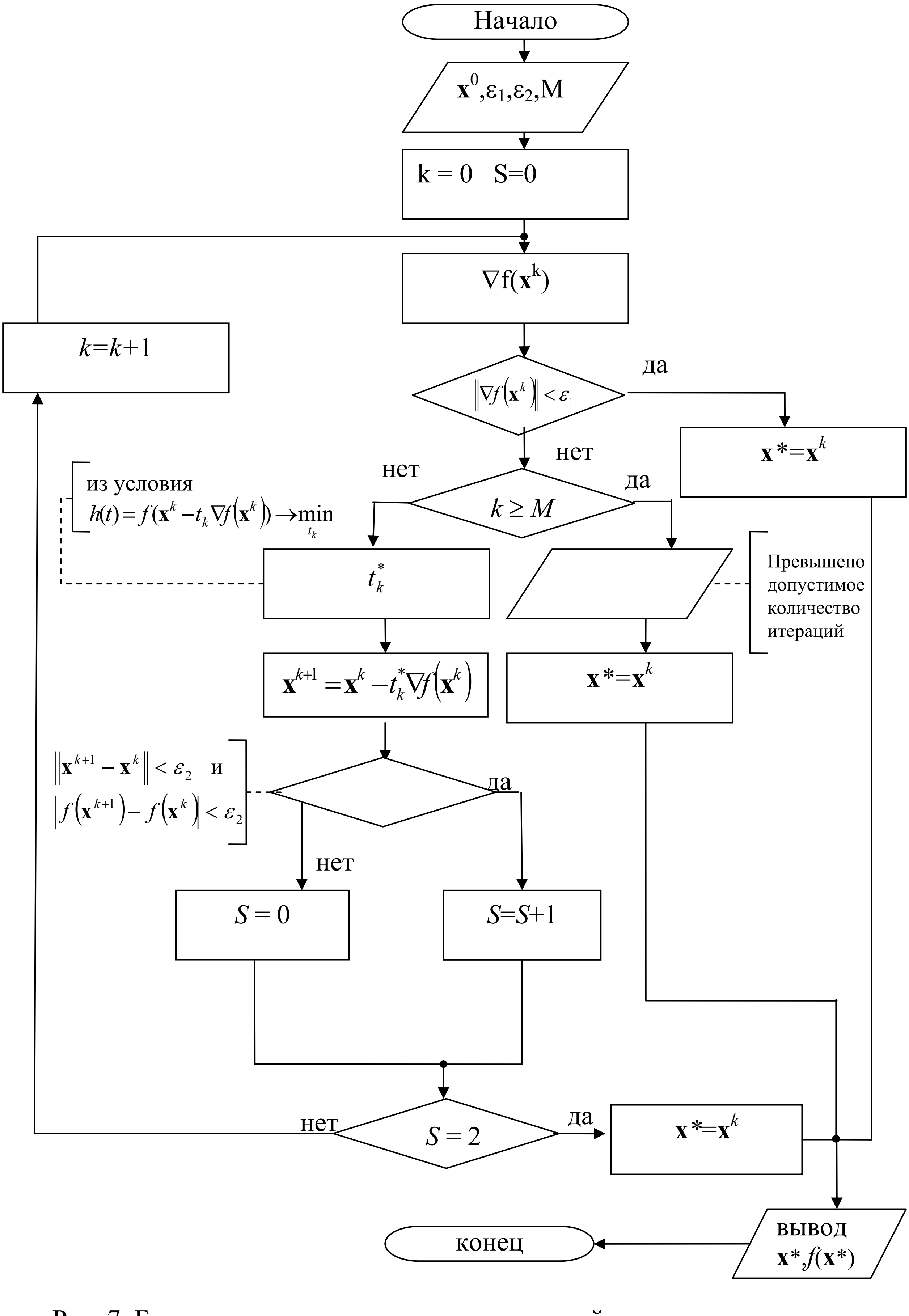



Рис. 7. Блок-схема алгоритма метода наискорейшего градиентного спуска

*Пример 3.*

Рассмотрим особенности поиска локального минимума для функции $f(\mathbf{x}) = 3x^2 + 2xy + 4y^2$, где $\mathbf{x} = (x, y)$ методом наискорейшего градиентного спуска.

Определим точку $\mathbf{x}^к$, в которой выполняется, по крайней мере, один из критериев окончания расчетов. Обозначения оставим такими же, что и в *примере 2*.

Итак, $к$ – счетчик количества итераций, начальное значение $к=0$;
$M$ – предельное число итераций, пусть $M=10$;
$\mathbf{x}^к=(x_к, y_к)$ – вектор $к$-го приближения к минимуму функции;
$\mathbf{d}^к$ – вектор направления спуска для $к$–й итерации, $\mathbf{d}^к = \nabla f(\mathbf{x}^к)$;
$t_к$ – коэффициент изменения направления перехода;
$\varepsilon_1$, $\varepsilon_2$- заданные малые положительные числа, влияющие на процесс окончания построения итерационной последовательности $\{ \mathbf{x}^к \}$, а именно,
$\| \nabla f(\mathbf{x}^к) \| < \varepsilon_1$ или *двукратное одновременное* выполнение двух неравенств $\| \mathbf{x}^{к+1} - \mathbf{x}^к \| < \varepsilon_2$ и $| f(\mathbf{x}^{к+1}) - f(\mathbf{x}^к)| < \varepsilon_2$, где $\nabla f(\mathbf{x}) = (f_x', f_y')$– градиент функции $f(\mathbf{x})$, т.е. $f_x' = 6x+2y$; $f_y' = 2x+8y$. Матрица Гессе $H(\mathbf{x}^0) = \begin{pmatrix} f''_{xx} & f''_{xy} \\ f''_{yx} & f''_{yy} \end{pmatrix} = \begin{pmatrix} 6 & 2 \\ 2 & 8 \end{pmatrix}$.

$H(\mathbf{x}^0)>0$, т.к. $\Delta_1 = 6>0$ и $\Delta_2 = \begin{vmatrix} 6 & 2 \\ 2 & 8 \end{vmatrix} = 44 > 0$

Дано $f(\mathbf{x}) = 3x^2 + 2xy + 4y^2$. Зададим $\mathbf{x}^0 = (0,5;\ 1)$; $\varepsilon_1 = 0,15$; $\varepsilon_2 = 0,4$; $M=10$; $f(\mathbf{x}^0) = 5,75$. Градиент в произвольной точке $\mathbf{x}=(x, y)$, т.е. $\nabla f(\mathbf{x}) = (6x+2y; 2x+8y)$.

*Первая итерация*

1. $к = 0$; $\nabla f(\mathbf{x}^0) = (5;\ 9)$;
2. $\| \nabla f(\mathbf{x}^0) \| = 10,9 > \varepsilon_1 = 0,15$;
3. $к=0<M=10$. Найдем координаты точки $\mathbf{x}^1=\mathbf{x}^0 - t_0 \cdot \nabla f(\mathbf{x}^0)=(0,5;1)- t_0 \cdot (5;9)= =(0,5- -5 \cdot t_0 ; 1 -9 \cdot t_0 )$. Определим значение $t_0$ как точку минимума функции $f(\mathbf{x}^1) =\varphi(t_0)$, для которого выполняются необходимое и достаточное условия существования безусловного экстремума: $\frac{d(\varphi(t_0))}{d(t_0)} = 0$ и $\frac{d^2(\varphi(t_0))}{d(t_0)^2} > 0$. Итак,

$\varphi(t_0)=f(\mathbf{x}^1)=3 \cdot (0,5-5 \cdot t_0)^2+2 \cdot (0,5-5 \cdot t_0) \cdot (1-9 \cdot t_0)+4 \cdot (1-9 \cdot t_0)^2=3 \cdot (0,25-5 \cdot t_0+25 \cdot t_0^2)+ +2 \cdot (0,5--9,5 \cdot t_0 +45 \cdot t_0^2)+4 \cdot (1-18 \cdot t_0 +36 \cdot t_0^2)= 5,75-106 \cdot t_0 +306 \cdot t_0^2$.

$\frac{d(\varphi(t_0))}{d(t_0)} = (5,75-106 \cdot t_0 +306 \cdot t_0^2) = -106+612\ t_0=0$; $t_0=0,173$; $\frac{d^2(\varphi(t_0))}{d(t_0)^2} = 612 > 0$.

Вывод: найденное значение $t_0 = 0,173$ обеспечивает минимум функции $\varphi(t_0)=f(\mathbf{x}^1)$; таким образом, $\mathbf{x}^1=\mathbf{x}^0- t_0 \cdot \nabla f(\mathbf{x}^0)=(0,5;-1)-0,173(0,5-5 \cdot 0,173;1-9 \cdot 0,173)=(0,865;1,557)$; $f(\mathbf{x}^1) =3 \cdot 0,865^2+2 \cdot 0,865 \cdot 1,557+4 \cdot 1,557^2=14,636$.

4. $| f(\mathbf{x}^1) - f(\mathbf{x}^0)| = | 14,636-5,75| = 8,886 > \varepsilon_2 = 0,4$; $\mathbf{x}^1 - \mathbf{x}^0 = =(0.865;1.557)-(0.5;1)=(0.365;0.557)$;

$\| \mathbf{x}^1 - \mathbf{x}^0 \| = \sqrt{(0.365)^2 + (0.557)^2} = \sqrt{0.443} = 0.665 > \varepsilon_2 = 0.4$. Вывод: итерационный процесс надо продолжить.

*Вторая итерация*

Полагаем $к=1 < M=10$; $\mathbf{x}^1=(0,865;1,557)$- начальное приближение к локальному минимуму функции $f(\mathbf{x}) = 3x^2 +2xy+4y^2$;

1. $\nabla f(\mathbf{x}^1) = (6x+2y;2x+8y)= (6\cdot 0,865+2\cdot 1,557;2\cdot 0,865+8\cdot 1,557)=(8,304;14,186)$.

2. $\| \nabla f(\mathbf{x}^1) \| = \sqrt{8.304^2 + 14.186^2} = 16.44 > \varepsilon_1 = 0.15$

3. Чтобы найти координаты точки $\mathbf{x}^2 = \mathbf{x}^1 - t_1\cdot \nabla f(\mathbf{x}^1)= (0,865;1,557) - t_1\cdot(8,304;14,186)=(0,865-8,304\cdot t_1;\ 1,557\ -14,186\cdot t_1)$. Определим значение $t_1$ как точку минимума функции $f(\mathbf{x}^2) =\varphi\ (t_1)$, для которого выполняются необходимое и достаточное условия существования безусловного экстремума: $\frac{d(\varphi(t_1))}{d(t_1)} = 0$ и $\frac{d^2(\varphi(t_1))}{d(t_1)^2} > 0$.

Итак, $\varphi(t_1)=f(\mathbf{x}^2)=3\cdot(0,865-8,304\cdot t_1)^2+2\cdot(0,865-8,304\cdot t_1)\cdot(1,557-14,186\cdot t_1)+4\cdot(1,557-14,186\cdot t_1)^2=3\cdot(0,748-14,366\cdot t_1+68,956\cdot t^2)+2\cdot(1,347-12,929\cdot t_1 -12,271\cdot t_1 +117,800\cdot t_1^{\ 2})+4\cdot(2,424-44,175\cdot t_1 +201,242\cdot t_1^{\ 2})=14,634-270,198\cdot t_1+1265,436\cdot t_1^2$. Причем $\frac{d(\varphi(t_1))}{d(t_1)} = (14,634-270,198\cdot t_1+1265,436\cdot t_1^2)' = -270,198+2530,872t_1=0;\quad t_1=0,107;$

$\frac{d^2(\varphi(t_1))}{d(t_1)^2} = 2530.872 > 0$. Значит, $t_1$ – точка минимума $f(\mathbf{x}^2)$ и координаты

$\mathbf{x}^2=\mathbf{x}^1-t_1\cdot \nabla f(\mathbf{x}^1)=(0,865;1,557)-0,107\cdot(8,304;14,186)=(0,865-8,304\cdot 0,107;1,557-$
$-14,186\cdot 0,107)=(-0,023;0,039)$;

$f(\mathbf{x}^2)=3\cdot(-0,023)^2+2\cdot(-0,023)\cdot 0,039+4\cdot 0,039^2=0,006$.

4. $|f(\mathbf{x}^2) - f(\mathbf{x}^1)| = | 0,006-14,636| = 14,63 > \varepsilon_2=0,4$;

$\mathbf{x}^2 - \mathbf{x}^1 =(-0,023;0,039)-(0,865;1,557)=(-0,888;-1,518)$;

$\| \mathbf{x}^1 - \mathbf{x}^0 \| = \sqrt{(-0.888)^2 + (-1.518)^2} = \sqrt{3.093} = 1.759 > \varepsilon_2=0.4$. Вывод: итерационный процесс надо продолжить.

*Третья итерация*

$к=2 < M=10$; $\mathbf{x}^2=(-0,023;0,039)$- начальное приближение к локальному минимуму функции $f(\mathbf{x}) = 3x^2 +2xy+4y^2$;

1. $\nabla f(\mathbf{x}^2) = (6x+2y;2x+8y)= (6\cdot(-0,023)+2\cdot 0,039;2\cdot(-0,023)+8\cdot 0,039)= =(-0,06;0,266)$.

2. $\| \nabla f(\mathbf{x}^2) \| = \sqrt{0.06^2 + 0.266^2} = 0.27 > \varepsilon_1 = 0.15$

3. Чтобы найти координаты точки $\mathbf{x}^3 = \mathbf{x}^2 - t_2\cdot \nabla f(\mathbf{x}^2)=$ $(-0,023;0,039) - t_2\cdot(-0,06;0,266)=(-0,023+0,06\cdot t_2;\ 0,039\ -0,266\cdot t_2)$. Определим значение $t_2$ как точку минимума функции $f(\mathbf{x}^3)=\varphi(t_2)$, для которого выполняются необходимое и достаточное условия существования безусловного экстремума: $\frac{d(\varphi(t_2))}{d(t_2)} = 0$ и $\frac{d^2(\varphi(t_2))}{d(t_2)^2} > 0$.

Итак, $\varphi(t_2)=f(\mathbf{x}^3)=3\cdot(-0,023+0,06\cdot t_2)^2+2\cdot\cdot(-0,023+0,06\cdot t_2)\cdot(0,039-0,266\cdot t_2)+4\cdot(0,039-0,266\cdot t_2)^2=3\cdot(0,0005-0,0028\cdot t_2 +0,0036\cdot t_2^2)+2\cdot(0,0009-0,0023\cdot t_2 -0,006\cdot t_2 +0,0159\cdot t_2^2)+4\cdot(0,015-0,0207\, t_2 +201,242\cdot t_2^2)= 0,0093-0,1072\cdot t_2+0,3258\cdot t_2^2$. Причем

$\frac{d(\varphi(t_2))}{d(t_2)} =$ $(0,0093-0,1072\cdot t_2+0,3258\cdot t_2^2)' = -0,1072 + 0,6516\cdot t_2=0$ ; $t_2=0,164$;

$\frac{d^2(\varphi(t_2))}{d(t_2)^2} = 0.6516 > 0$. Значит, $t_2$ – точка минимума $f(\boldsymbol{x}^3)$ и координаты

$\boldsymbol{x}^3 = \boldsymbol{x}^2 - t_2\cdot\nabla f(\boldsymbol{x}^2) = (-0,023;0,039) - 0,164\cdot(-0,06;0,266) = (-0,013; -0,005)$.

$f(\boldsymbol{x}^3) = 3\cdot(-0,013)^2+2\cdot(-0,013)\cdot(-0,005)+4\cdot(-0,005)^2=0,0007$.

4. $|f(\boldsymbol{x}^3) - f(\boldsymbol{x}^2)| = |0,0007-0,006| = 0,0053 < \varepsilon_2=0,4$;

$\boldsymbol{x}^3 - \boldsymbol{x}^2 = (-0,013;-0,005)-(-0.023;0.039)=(0.01; 0.044)$;

$\|\boldsymbol{x}^3-\boldsymbol{x}^2\| = \sqrt{(0.01)^2 + (0.044)^2} = 0.035 < \varepsilon_2=0.4$.

5. $\nabla f(\boldsymbol{x}^3) = (6x+2y;2x+8y) = (6\cdot(-0,013)+2\cdot(-0,005);2\cdot(-0,013)+8\cdot(-0,005)) =$ $(-0,088;-0,066)$.

$\|\nabla f(\boldsymbol{x}^3)\| = \sqrt{0.088^2 + 0.066^2} = 0.11 < \varepsilon_1 = 0.15$

Расчет окончен. Найдена точка $\boldsymbol{x}^3 = (-0,013;-0,005)$, $f(\boldsymbol{x}^3)=0,0007$. Причем несложно убедиться в том, что эта точка удовлетворяет условиям как локального, так и глобального минимума.

### *2.3. Методы второго порядка*

В градиентных методах производится линеаризация, связанная с заменой минимизируемой целевой функции $f(\boldsymbol{x})$ линейным членом разложения ее в ряд Тейлора. Но если функция дважды непрерывно дифференцируема, то можно использовать для ее аппроксимации два члена ряда Тейлора. Использование такой аппроксимации в итерационном процессе позволяет повысить скорость сходимости.

Наиболее известным из методов второго порядка является ***метод Ньютона***.

Разложим целевую функцию $f(\boldsymbol{x})$ в ряд Тейлора

$$f(\boldsymbol{x}) = f(\boldsymbol{x}^k) + \nabla f(\boldsymbol{x}^k)^{\mathrm{T}} \Delta\boldsymbol{x} + \frac{1}{2}\Delta\boldsymbol{x}^{\mathrm{T}}\nabla^2 f(\boldsymbol{x}^k)\Delta\boldsymbol{x} + O(\Delta\boldsymbol{x}^3).$$

Отбрасывая все члены разложения третьего порядка и выше, получим квадратичную аппроксимацию $f(\boldsymbol{x})$:

$$F_k = f(\boldsymbol{x}^k) + \nabla f(\boldsymbol{x}^k)^{\mathrm{T}} \Delta\boldsymbol{x} + \frac{1}{2}\Delta\boldsymbol{x}^{\mathrm{T}}\nabla^2 f(\boldsymbol{x}^k)\Delta\boldsymbol{x}.$$

Используя квадратичную аппроксимацию функции $f(\boldsymbol{x})$, сформируем последовательность итераций таким образом, чтобы в точке $\boldsymbol{x}^{k+1}$ градиент аппроксимирующей функции обращался в ноль (в соответствии с необходимым условием экстремума первого порядка):

$$\nabla F_k = \nabla f(\boldsymbol{x}^k) + \nabla^2 f(\boldsymbol{x}^k)\Delta\boldsymbol{x} = 0.$$

Следовательно, $\Delta\boldsymbol{x} = -(\nabla^2 f(\boldsymbol{x}^k))^{-1}\nabla f(\boldsymbol{x}^k)$.

Последовательно применяя квадратичную аппроксимацию, построим последовательность $\{\boldsymbol{x}^k\}$ в соответствии с правилом

$$\boldsymbol{x}^{k+1} = \boldsymbol{x}^k - (\nabla^2 f(\boldsymbol{x}^k))^{-1}\nabla f(\boldsymbol{x}^k).$$

При выполнении требования

$H\left(\boldsymbol{x}^k\right)=\nabla^2 f\left(\boldsymbol{x}^k\right)>0$ последовательность $\{\boldsymbol{x}^k\}$ является последовательностью точек минимумов квадратичных функций $F_k$, k=0, 1,… (рис. 8).

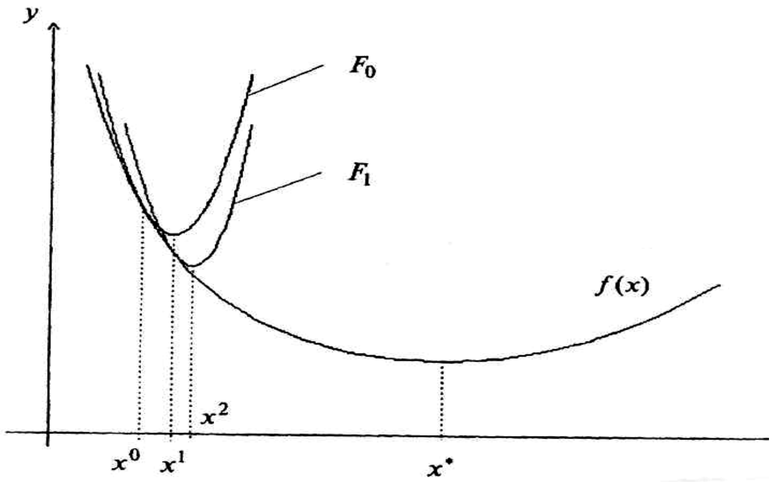



Рис.8. Последовательность точек минимумов квадратичных функций $F_k$

Чтобы обеспечить выполнение требования $f\left(\boldsymbol{x}^{k+1}\right)<f\left(\boldsymbol{x}^k\right)$, k = 0,1,… для тех случаев, когда матрица Гессе (например, для функции *f(**x**)=f(x,y.z),* имеющая вид: $H(\boldsymbol{x}^k)=\begin{pmatrix} f''_{xx} & f''_{xy} & f''_{xz} \\ f''_{yx} & f''_{yy} & f''_{yz} \\ f''_{zx} & f''_{zy} & f''_{zz} \end{pmatrix}$) не окажется положительно определенной, рекомендуется для соответствующих значений $k$ вычислить точку $\boldsymbol{x}^{k+1}$ по методу градиентного спуска $\boldsymbol{x}^{k+1}=\boldsymbol{x}^k-t_k\ \nabla f\left(\boldsymbol{x}^k\right)$ с выбором величины шага $t_k$ из условия $f\left(\boldsymbol{x}^k-t_k\nabla f\left(\boldsymbol{x}^k\right)\right)<f\left(\boldsymbol{x}^k\right)$.

Построение последовательности $\{\boldsymbol{x}^k\}$ заканчивается в точке $\boldsymbol{x}^k$, для которой $\left\|\nabla f\left(\boldsymbol{x}^k\right)\right\|<\varepsilon_1$, где $\varepsilon_1$ – заданное малое положительное число, или $k<M$, где $M$ – предельное число итераций, или при *двукратном одновременном* выполнении двух неравенств $\left\|\boldsymbol{x}^{k+1}-\boldsymbol{x}^k\right\|\le\varepsilon_2$, $\left|f\left(\boldsymbol{x}^{k+1}\right)-f\left(\boldsymbol{x}^k\right)\right|\le\varepsilon_2$, где $\varepsilon_2$ – малое положительное число.

Блок-схема алгоритма Ньютона представлена на рис.9.

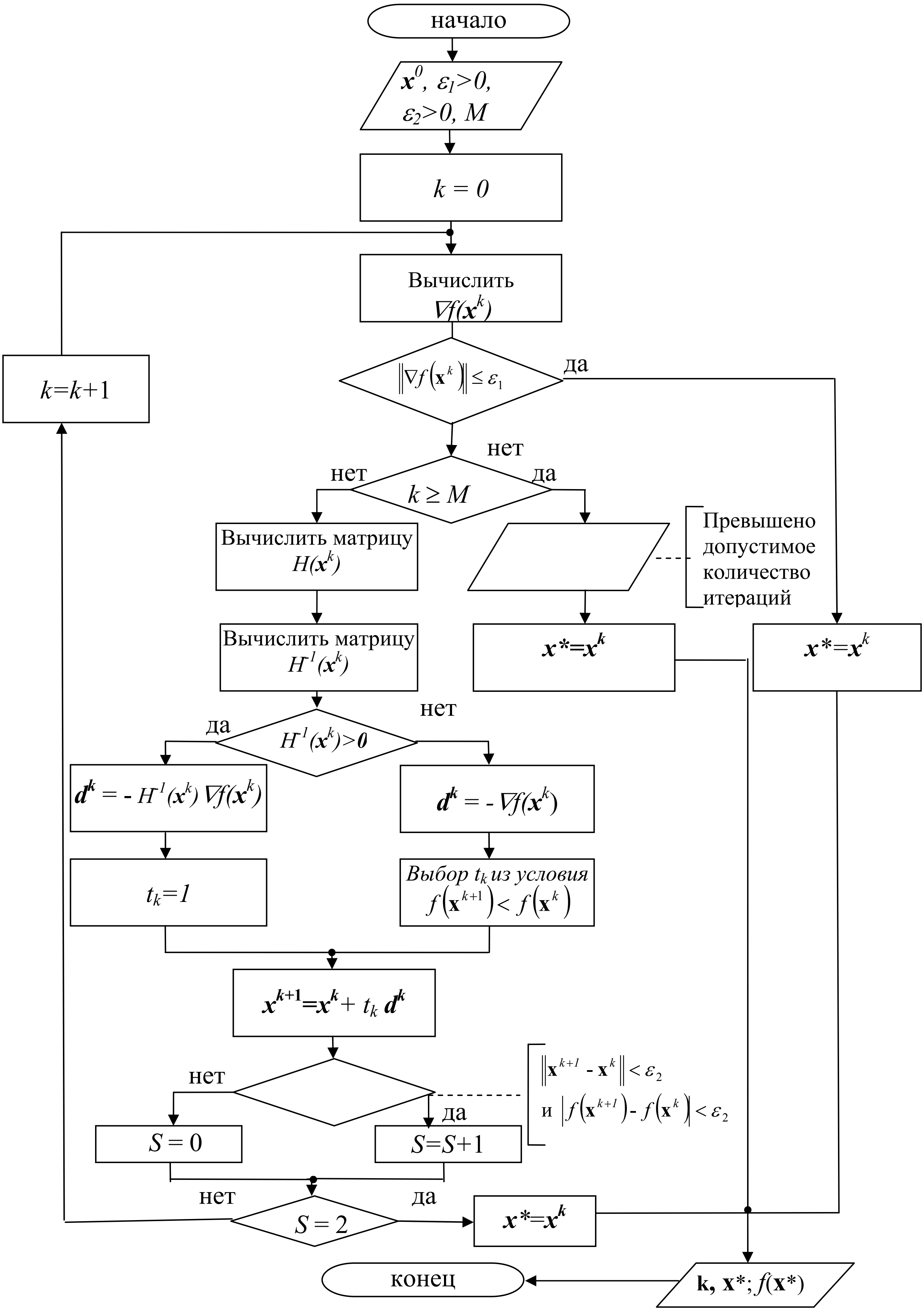



Рис.9. Блок-схема алгоритма метода Ньютона

*Пример 4*. Найти локальный минимум функции $f(\mathbf{x}) = 5x^2 + 2xy + 6y^2$, где $\mathbf{x}=(x, y)$.

Рассмотрим численное решение задачи методом Ньютона.

Введем обозначения:

$к$ – счетчик количества итераций, начальное значение $к=0$;

$M$ – предельное число итераций, пусть $M=10$;

$\mathbf{x}^к=(x_к, y_к)$ – вектор $к$-го приближения к минимуму функции;

$\mathbf{d}^к$ – вектор направления спуска для $к$–й итерации, $\mathbf{d}^к = -H^{-1}(\mathbf{x}^k)\,\nabla f(\mathbf{x}^k)^T$;

$\varepsilon_1$, $\varepsilon_2$ – заданные малые положительные числа, влияющие на процесс окончания построения итерационной последовательности $\{\mathbf{x}^к\}$, а именно,

$\|\nabla f(\mathbf{x}^к)\| < \varepsilon_1$ или *двукратное одновременное* выполнение двух неравенств $\|\mathbf{x}^{к+1} - \mathbf{x}^к\| < \varepsilon_2$ и $|f(\mathbf{x}^{к+1}) - f(\mathbf{x}^к)| < \varepsilon_2$,

где $\nabla f(\mathbf{x}) = (f_x', f_y') = (10x+2y;\ 2x+12y)$ – градиент функции $f(\mathbf{x})$, т.е.

$f_x' = 10x+2y$; $f_y' = 2x+12y$. Матрица Гессе $H(\mathbf{x}^0) = \begin{pmatrix} f''_{xx} & f''_{xy} \\ f''_{yx} & f''_{yy} \end{pmatrix} = \begin{pmatrix} 10 & 2 \\ 2 & 12 \end{pmatrix}$.

Видим, что $H(\mathbf{x}^к)$ положительно определена для всех векторов $\mathbf{x}^к$, т.к. $\Delta_1 = 10>0$ и $\Delta_2 = \begin{vmatrix} 10 & 2 \\ 2 & 12 \end{vmatrix} = 120 - 4 = 116 > 0$ и не зависит от вектора $\mathbf{x}^к=(x_к, y_к)$.

Итак, пусть при $к = 0$ начальное приближение $\mathbf{x}^0 = (1;2)$, $\varepsilon_1, =0,1;$ $\varepsilon_2 = 0,15;$ т.е. $x_0 = 1$, $y_0 = 2$.

*Первая итерация*

1. $f(\mathbf{x}^0) = 5\cdot 1^2+2\cdot 1\cdot 2+6\cdot 2^2 = 33$.

2. Найдем значение градиента в точке $\mathbf{x}^0$, т.е. $\nabla f(\mathbf{x}^0) = (10\cdot 2+2\cdot 2;\ 2\cdot 1+12\cdot 2) = (14;\ 26)$.

3. Проверим условие окончания итерационных процессов, например, вычислив норму градиента в точке $\mathbf{x}^0$ по формуле $\|\nabla f(\mathbf{x}^0)\| = \sqrt{(f_x'(x^0))^2 + (f_y'(x^0))^2} = \sqrt{(10\cdot 1+2\cdot 2)^2+(2\cdot 1+12\cdot 2)^2} = \sqrt{14^2+26^2} = \sqrt{872} = 29.53 > \varepsilon_1 = 0.1$

Вывод: итерационный процесс не закончен.

4. Знаем, что матрица Гессе $H(\mathbf{x}^0) = \begin{pmatrix} f''_{xx} & f''_{xy} \\ f''_{yx} & f''_{yy} \end{pmatrix} = \begin{pmatrix} 10 & 2 \\ 2 & 12 \end{pmatrix}$. Найдем обратную матрицу $H^{-1}(\mathbf{x}^0) = \frac{1}{\Delta H(x^0)}\cdot\begin{pmatrix} A_{11} & A_{21} \\ A_{12} & A_{22} \end{pmatrix}$, где $A_{ij} = (-1)^{i+j}\cdot M_{ij}$ - алгебраические дополнения элементов матрицы $H(\mathbf{x}^0)$. Таким образом, $\Delta H(\mathbf{x}^0) = \begin{vmatrix} 10 & 2 \\ 2 & 12 \end{vmatrix} = 116$;

$A_{11}=12$, $A_{12}=-2$, $A_{21}=-2$, $A_{22}=10$; $H^{-1}(\mathbf{x}^0) = \frac{1}{116}\cdot\begin{pmatrix} 10 & -2 \\ -2 & 12 \end{pmatrix} = \begin{pmatrix} \frac{5}{58} & \frac{-1}{58} \\ \frac{-1}{58} & \frac{3}{29} \end{pmatrix}$.

Здесь $\Delta_1 = \frac{5}{58} > 0$ , $\Delta_2 = \begin{vmatrix} \frac{5}{58} & \frac{-1}{58} \\ \frac{-1}{58} & \frac{3}{58} \end{vmatrix} = \frac{29}{58} > 0$. Значит, согласно критерию Сильвестра матрица $H^{-1}(\mathbf{x}^0)$ является положительно определенной, т.е. $H^{-1}(\mathbf{x}^0) > 0$.

*5.* Определим $\mathbf{d}^0$ – вектор изменения координат начальной точки $\mathbf{x}^0$ по формуле $\mathbf{d}^0 = -H^{-1}(\mathbf{x}^0)\cdot(\nabla f(\mathbf{x}^0))^T = -\begin{pmatrix} \frac{5}{58} & \frac{-1}{58} \\ \frac{-1}{58} & \frac{3}{29} \end{pmatrix}\cdot\begin{pmatrix} 14 \\ 26 \end{pmatrix} = -\begin{pmatrix} \frac{44}{58} \\ \frac{142}{58} \end{pmatrix} = \begin{pmatrix} \frac{-44}{58} \\ \frac{-142}{58} \end{pmatrix}$.

*6.* Находим новое приближение к минимуму по формуле
$\mathbf{x}^1 = \mathbf{x}^0 + (\mathbf{d}^0)^T = (1;2) + (-44/58;-142/58) = (14/58;-26/58)$, т.е. $x_1 = 7/29$; $y_1 = -13/29$.

*7.* Проверим условие окончания итерационных процессов по второму критерию. Норма $\|\mathbf{x}^1 - \mathbf{x}^0\| < \varepsilon_2$ и $|f(\mathbf{x}^1) - f(\mathbf{x}^0)| < \varepsilon_2$, зная, что $\|\mathbf{x}^1 - \mathbf{x}^0\| = \sqrt{(x_0 - x_1)^2 + (y_0 - y_1)^2} = \sqrt{(1 - \frac{7}{29})^2 + (2 + \frac{13}{29})^2} = \sqrt{\frac{5525}{29^2}} = 2.56 > \varepsilon_2$. Учитывая условие задачи $f(\mathbf{x}) = 5x^2 + 2xy + 6y^2$, найденное и начальное приближения $\mathbf{x}^1 = (14/58;-26/58)$, $\mathbf{x}^0 = (1;2)$, вычислим $|f(\mathbf{x}^1) - f(\mathbf{x}^0)| = |5\cdot(7/29)^2 + 2\cdot 7/29\cdot(-13/29) + (-13/29)^2 - 33| = 32,7 > \varepsilon_2$. Таким образом, критерий не выполнен, значит, надо повторить все пункты *первой итерации*, взяв $\mathbf{x}^1 = (7/29;-13/29)$ в качестве начального приближения, т.к. количество итераций ($к=0 < M=10$) еще не исчерпано.

*Вторая итерация*

*1.* $к=1$, $\mathbf{x}^1 = (7/29;-13/29)$, т.е. $x_0 = 7/29$; $y_0 = -13/29$ , $f(\mathbf{x}^1) = 0,27$

*2.* Градиент функции $\nabla f(\mathbf{x}^1) = (10\cdot 7/29 + 2\cdot(-13/29);\ 2\cdot 7/29 + 12\cdot(-13/29)) = (44/29;-142/29)$.

*3.* Проверяем условие окончания итерационного процесса. Вычислим норму градиента $\|\nabla f(\mathbf{x}^1)\| = \sqrt{(f_x'(x^1))^2 + (f_y'(x^1))^2} = \sqrt{(\frac{44}{29})^2 + (-\frac{142}{29})^2} = 5.1 > \varepsilon_1 = 0.1$

Условие не выполнено.

*4.* Ранее было показано, что для всех точек $\mathbf{x}^к$ матрица Гессе $H(\mathbf{x}^к) = \begin{pmatrix} f''_{xx} & f''_{xy} \\ f''_{yx} & f''_{yy} \end{pmatrix} = \begin{pmatrix} 10 & 2 \\ 2 & 12 \end{pmatrix}$ и обратная ей для данной задачи положительно определены и постоянны, поэтому переходим к следующему пункту вычислений.

*5.* Определим

$$\mathbf{d}^1 = -H^{-1}(\mathbf{x}^1)\cdot(\nabla f(\mathbf{x}^1))^T = -\begin{pmatrix} \frac{5}{58} & \frac{-1}{58} \\ \frac{-1}{58} & \frac{3}{58} \end{pmatrix}\cdot\begin{pmatrix} \frac{44}{29} \\ \frac{-142}{29} \end{pmatrix} = \begin{pmatrix} \frac{-362}{58\cdot 29} \\ \frac{896}{58\cdot 29} \end{pmatrix} = \begin{pmatrix} \frac{-181}{29^2} \\ \frac{448}{29^2} \end{pmatrix} = \begin{pmatrix} -0.215 \\ 0.533 \end{pmatrix}.$$

*6.* Находим новое приближение $\mathbf{x}^2 = \mathbf{x}^1 + (\mathbf{d}^1)^T = (7/29;-13/29) + (-0,215;\ 0,533) = (0,026;\ 0,085)$, т.е. $x_2 = 0,026$; $y_2 = 0,085$.

*7.* Проверим условие окончания итерационного процесса по второму критерию. Норма $\|\mathbf{x}^2 - \mathbf{x}^1\| < \varepsilon_2$ и $|f(\mathbf{x}^2) - f(\mathbf{x}^1)| < \varepsilon_2$, зная, что $\|\mathbf{x}^2 - \mathbf{x}^1\| =$

$$\sqrt{(x_1-x_2)^2+(y_1-y_2)^2} = \sqrt{(\frac{7}{29}-0.026)^2+(-\frac{13}{29}-0.085)^2} = \sqrt{0.215^2+0.533^2} = 0.574 > \varepsilon_2 = 0.15$$

$|f(\mathbf{x}^2) - f(\mathbf{x}^1)| = |5\cdot 0{,}026^2+2\cdot 0{,}026\cdot 0{,}085+6\cdot 0{,}085^2-0{,}27| = |0{,}05-0{,}27| = 0{,}22>\varepsilon_2.$

Таким образом, критерий не выполнен, значит, надо повторить все пункты *первой итерации*, взяв $\mathbf{x}^2=(0{,}026;\ 0{,}085)$ в качестве начального приближения, т.к. количество итераций ($к=1 < M=10$) еще не исчерпано.

*Третья итерация*

1. $к=2$, $\mathbf{x}^2=(0{,}026;0{,}085)$, т.е. $x_2 = 0{,}026;\ y_2 = 0{,}085$ , $f(\mathbf{x}^2)= 0{,}05$
2. $\nabla f(\mathbf{x}^2) = (10\cdot 0{,}026+2\cdot 0{,}085;\ 2\cdot 0{,}026+12\cdot 0{,}085) = (0{,}43;1{,}072).$
3. $\|\nabla f(\mathbf{x}^2)\| = \sqrt{(f_x'(x^2))^2+(f_y'(x^2))^2} = \sqrt{(0.43)^2+(1.072)^2} = 1.155 > \varepsilon_1 = 0.1$
4. Без изменений.
5. $\mathbf{d}^2 = -H^{-1}(\mathbf{x}^2)\cdot(\nabla f(\mathbf{x}^2))^T = -\begin{pmatrix} \frac{5}{58} & \frac{-1}{58} \\ \frac{-1}{58} & \frac{3}{58} \end{pmatrix}\cdot\begin{pmatrix} 0.43 \\ 1.155 \end{pmatrix} = \begin{pmatrix} 0.017 \\ 0.052 \end{pmatrix}.$
6. $\mathbf{x}^3= \mathbf{x}^2+(\mathbf{d}^2)^T=(0{,}026\ ;0{,}085)+(0{,}017;\ 0{,}052)=(0{,}043;0{,}137)$
7. $\|\mathbf{x}^3-\mathbf{x}^2\| = \sqrt{(x_2-x_3)^2+(y_2-y_3)^2} == \sqrt{0.017^2+0.052^2} = 0.054 < \varepsilon_2 = 0.15$

$|f(\mathbf{x}^3)- f(\mathbf{x}^2)| = |5\cdot 0{,}044^2+2\cdot 0{,}044\cdot 0{,}133+6\cdot 0{,}133^2-0{,}27| = |0{,}127-0{,}05| = 0{,}077<\varepsilon_2.$

Вывод. Расчет закончен. Но так как выполнилось необходимое условие первого порядка, то точка $\mathbf{x}^3=(0{,}043;0{,}137)$ является стационарной точкой. Проведем *анализ* полученной стационарной точки.

Функция $f(\mathbf{x}) = 5x^2 +2xy+6y^2$ является строго выпуклой, т.к. ее матрица вторых производных $H(\mathbf{x}) = \begin{pmatrix} f''_{xx} & f''_{xy} \\ f''_{yx} & f''_{yy} \end{pmatrix} = \begin{pmatrix} 10 & 2 \\ 2 & 12 \end{pmatrix}$>0 в силу того, что $\Delta_1 = 10$ и $\Delta_2 = \begin{vmatrix} 10 & 2 \\ 2 & 12 \end{vmatrix} =120 - 4 = 116 > 0$ для всех критических точек. Таким образом, $\mathbf{x}^3=(0{,}043;0{,}137)$ – точка как локального, так и глобального минимума функции $f(\mathbf{x}) = 5x^2 +2xy+6y^2$. Итоги расчетов отображены в таблице и на графике рис. 10.

| Координаты точек $x^к$ | |
|---|---|
| **x** | **y** |
| 1 | 2 |
| 0,24 | -0,45 |
| 0,026 | 0,085 |
| 0,043 | 0,137 |


Рис.10. Движение точек к минимуму по методу Ньютона

Пример поиска минимума средствами MathCAD.

$$f(x, y) := 2\cdot(x - 5.07)^2 + (y - 10.03)^2 - 0.2\cdot(x - 5.07)^3$$

$x := 3$ $\qquad$ $y := 3$

Given

$0 < x < 15$

$0 < y < 20$

$$\mathrm{Minimize}(f, x, y) = \begin{pmatrix} 5.07 \\ 10.03 \end{pmatrix}$$

$$f(x, y) := 2\cdot(x - 5.07)^2 + (y - 10.03)^2 - 0.2\cdot(x - 5.07)^3$$

$x := 3$ $\qquad$ $y := 3$

Given

$f(x, y) = 0$

$0 < x < 15$

$0 < y < 15$

$$\mathrm{Minerr}(x, y) = \begin{pmatrix} 5.07 \\ 10.034 \end{pmatrix}$$

***Варианты заданий***

Найти точку локального минимума, используя указанные методы, причем расчеты по методу, который указан первым, выполнить вручную, записав все промежуточные расчеты в тетрадь. Для решения задачи вторым методом можно использовать программирование в среде PASCAL, MathCad, создавать электронные таблицы в среде EXCEL; результаты расчета и анализа предоставить в печатном виде, подклеив их в тетрадь. Отобразить геометрическую интерпретацию метода графически в соответствии с проведенными расчетами, указать количество итераций и свести результаты переходов в сводную таблицу. При защите лабораторной работы необходимо знать все методы.

Таблица 2

| ***№ варианта*** | ***Функция f(x)*** | ***Метод*** | $< \varepsilon_1$ | $< \varepsilon_2$ | ***M*** | ***Начальное приближение*** |
|---|---|---|---|---|---|---|
| 1 | $x^3+3xy^2-15x-12y$ | 1,2 | 0,1 | 0,15 | 10 | (1,5; 0,6) |
| 2 | $(x-1)^2+2y^2$ | 2,3 | 0,1 | 0,15 | 9 | (0,6;0,2) |
| 3 | $x^2+xy+y^2-2x-y$ | 3,4 | 0,1 | 0,15 | 8 | (0,7;0,3) |
| 4 | $x^4+y^4-2x^2+4xy-2y^2$ | 4,5 | 0,1 | 0,15 | 8 | (1,3;-1,1) |
| 5 | $x^4+y^4-2x^2+4xy-2y^2$ | 5,4 | 0,1 | 0,15 | 8 | (-1,2;1,1) |

| Продолжение табл.2 **№ варианта** | ***Функция f(x)*** | ***Метод*** | $< \varepsilon_1$ | $< \varepsilon_2$ | ***M*** | ***Начальное приближение*** |
|---|---|---|---|---|---|---|
| 6 | $x^3y^2(x+y-6)$ | 1,5 | 0,1 | 0,15 | 10 | (2,7;1,7) |
| 7 | $x^2+y^2-4y+4$ | 2,5 | 0,1 | 0,15 | 10 | (0,2;1,7) |
| 8 | $(x-3)^2+2y^2-6y$ | 3,5 | 0,1 | 0,15 | 9 | (2,7;1,7) |
| 9 | $(x+3)^2+2y^2-6y$ | 4,5 | 0,1 | 0,15 | 9 | (-3,3;1,6) |
| 10 | $2x^2+3y^2-x-7y$ | 1,4 | 0,1 | 0,15 | 10 | (0,1;0,8) |
| 11 | $1+x-2y+6x^2+y^2$ | 1,3 | 0,1 | 0,15 | 9 | (-0,08;0,7) |
| 12 | $4(y-x)+x^2+y^2$ | 2,4 | 0,1 | 0,15 | 9 | (1,97;-2,2) |
| 13 | $3x^2+2y^2-2xy+10$ | 3,4 | 0,1 | 0,15 | 10 | (0,4;0,3) |
| 14 | $x^3+3xy^2-15x-12y$ | 3,5 | 0,1 | 0,15 | 10 | (1,5; 0,6) |
| 15 | $(x-1)^2+2y^2$ | 4,5 | 0,1 | 0,15 | 9 | (0,6;0,2) |
| 15 | $x^2+xy+y^2-2x-y$ | 4,3 | 0,1 | 0,15 | 8 | (0,7;0,3) |
| 17 | $x^4+y^4-2x^2+4xy-2y^2$ | 5,4 | 0,1 | 0,15 | 8 | (1,3;-1,1) |
| 18 | $x^4+y^4-2x^2+4xy-2y^2$ | 5,2 | 0,1 | 0,15 | 8 | (-1,2;1,1) |
| 19 | $x^3y^2(x+y-6)$ | 4,1 | 0,1 | 0,15 | 10 | (2,7;1,7) |
| 20 | $x^2+y^2-4y+14$ | 3,1 | 0,1 | 0,15 | 9 | (0,2;1,7) |

В таблице заданий указаны следующие методы:

1 – Метод нулевого порядка: *метод покоординатного спуска;*

2 – Метод нулевого порядка: *метод конфигураций (метод Хука – Дживса);*

3 –Метод первого порядка: *метод градиентного спуска с постоянным шагом;*

4 – Метод первого порядка: *метод наискорейшего градиентного спуска;*

5 – Метод второго порядка: *метод Ньютона.*

## *Список литературы*


1. Вержбицкий В.М. Основы численных методов.- М.: Высш. школа, 2002.
2. Бахвалов Н.С., Жидков Н.П., Кобельков Г.М. Численные методы.- М.: Лаборатория базовых знаний, 2001.
3. Дьяконов В.П. Компьютерная математика. Теория и практика.- М.: Нолидж, 2001.
4. Лапчик М.П., Рагулина М.И., Е.К. Хеннер. Численные методы: Учеб. пособие для студ.вузов/ Под ред. М.П. Лапчика.- М.: Издательский центр «Академия», 2004.
5. Мудров А.Е. Численные методы для ПЭВМ на языках Бейсик, Фортран, Паскаль.- Томск: МП «РАСКО», 1991.
6. Пантелеев А.В., Летова Т.А. Методы оптимизации в примерах и задачах.- М.: Высшая школа, 2002.
7. Плис А.И., Сливина Н.А. Лабораторный практикум по высшей математике. – М.: Высшая школа, 1994.
8. Плис А.И., Сливина Н.А. MathCAD: математический практикум для экономистов и инженеров: Учебное пособие. – М.: Финансы и статистика, 1999.
9. Демидович Б.П., Марон И.А. Основы вычислительной математики. – М.: Наука, 1970.
10. Кирьянов Д.В. Самоучитель MathCAD 11. – СПб.: БХВ-Петербург, 2004.

Сысолятина Лидия Геннадьевна
Котликова Вера Яковлевна

# Многомерная оптимизация

Методические указания
для выполнения лабораторной работы
по курсу «Методы вычислений» и
«Теория системного анализа и принятия решений»
для студентов специальностей 010101, 280101

Редактор Н.Л. Попова

| | | |
|---|---|---|
| Подписано к печати | Формат 60*84 1/16 | Бумага тип. №1 |
| Печать трафаретная | Усл. печ. л. 2,0 | Уч. – изд. л. 2,0 |
| Заказ | Тираж 50 | Цена свободная |

Редакционно-издательский центр КГУ.
640669, г. Курган, ул. Гоголя, 25.
Курганский государственный университет.