Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Кемеровский государственный университет»
Новокузнецкий институт (филиал)

Факультет информатики, математики и экономики
Кафедра математики, физики и математического моделирования

Е.А. Вячкина, Е. С. Вячкин

МЕТОДЫ ОПТИМИЗАЦИИ

Методические указания к выполнению внеаудиторной самостоятельной работы по дисциплине для обучающихся очной формы по направлениям подготовки 01.03.02 Прикладная математика и информатика, профиль «Математическое моделирование и информационные технологии»

Новокузнецк

2020

Вячкина Е. А.

Методы оптимизации: методические указания к выполнению внеаудиторной самостоятельной работы по дисциплине для обучающихся очной формы по направлениям подготовки 01.03.02 Прикладная математика и информатика, профиль «Математическое моделирование и информационные технологии» / Е.А. Вячкина; Новокузнецкий ин-т (фил.) Кемеров. гос. ун-та. – Новокузнецк : НФИ КемГУ, 2020 – 49 с.

Методические указания содержат теоретические основы для решения задач, шесть контрольных работ с подробным описанием решения, вопросы к экзамену и список рекомендуемой литературы.

Методические указания предназначены для студентов очной формы обучения направлений 01.03.02 Прикладная математика и информатика, профиль «Математическое моделирование и информационные технологии».

Рекомендовано на заседании
кафедры математики, физики и
математического моделирования
Протокол № 5 от 10.12.2020

Заведующий каф. МФММ

______________ / Е.В.Решетникова

Текст представлен в авторской редакции

СОДЕРЖАНИЕ

ВВЕДЕНИЕ

Настоящие методические указания адресованы студентам очной формы обучения, получающим квалификацию бакалавр по направлению подготовки: 01.03.02 Прикладная математика и информатика, профиль «Математическое моделирование и информационные технологии» и направлены на оказание помощи студентам в выполнении внеаудиторной самостоятельной работы по дисциплине «Методы оптимизации».

Дисциплина «Методы оптимизации» является обязательной дисциплиной при подготовке бакалавров в области математического моделирования и математического обеспечения информационных технологий.

В рамках дисциплины изучается четыре основных раздела: Однако домашние контрольные работы выполняются только по двум разделам: «Методы решения задач безусловной одномерной оптимизации» и «Методы безусловной многомерной оптимизации». По данным разделам в методических указаниях представлены теоретический материал и домашние контрольные работы с подробным описанием решения.

В методические рекомендации также включены варианты контрольных работ, вопросы к экзамену и список рекомендуемой литературы.

Раздел Методы решения задач безусловной одномерной оптимизации

1. ПОСТАНОВКА ЗАДАЧИ

Одномерная оптимизация заключается в нахождении точки x^*, в которой целевая функция $f(x^*)$ принимает максимальное или минимальное значение. Часто в постановках задачи может быть задан отрезок $[a, b]$, в котором находится оптимальное значение.

Функция $f(x)$ имеет локальный минимум в точке x^*, если при $\varepsilon > 0$ существует окрестность $[x^* - \varepsilon, x^* + \varepsilon]$ такая, что для всех значений x в этой окрестности $f(x)$ больше $f(x^*)$. Функция $f(x)$ имеет глобальный минимум в точке x^*, если для всех x справедливо неравенство $f(x) > f(x^*)$.

2. АНАЛИТИЧЕСКИЙ АНАЛИЗ ФУНКЦИИ

2.1 Необходимые и достаточные условия экстремума

Классический подход к задаче нахождения экстремумов функции состоит в поиске условий, которым они должны удовлетворять. ***Необходимым условием*** экстремума в точке x^* является равенство нулю первой производной (теорема Ферма), т.е. требуется решить уравнение

$$f'(x) = 0. \tag{1}$$

Данному уравнению удовлетворяют как локальные и глобальные экстремумы, так и точки перегиба функции, поэтому приведенное условие является только необходимым, но недостаточным.

С целью получения ***достаточных условий*** требуется расчет значений вторых производных в точках, удовлетворяющих уравнению (1). При этом доказано, что минимуму функции соответствует положительное значение второй производной, т.е. $f''(x^*) > 0$, а максимуму – отрицательное, т.е. $f''(x^*) < 0$. Однако, если вторая производная равна нулю, ситуация остается неопределенной и необходимо исследовать высшие производные. При этом если первая высшая производная не равная 0 имеет четный порядок, то экстремум существует, в противном случае – нет.

2.2 Унимодальные функции

Дадим определение унимодальной функции при поиске минимума.

Определение. Непрерывная функция $f(x)$ называется унимодальной на отрезке $[a, b]$ если:

- точка x^* локального минимума функции принадлежит отрезку $[a, b]$;

- для любых двух точек отрезка x_1 и x_2 , взятых по одну сторону от точки минимума, точке x_1 более близкой к точке минимума соответствует меньшее значение функции, т.е. при $x^* < x_1 < x_2$ либо при $x_2 < x_1 < x^*$ справедливо неравенство $f(x_1) < f(x_2)$.

Достаточное условие унимодальности функции $f(x)$ на отрезке $[a, b]$ содержится в следующей теореме.

Теорема. Если функция $f(x)$ дважды дифференцируема на отрезке $[a, b]$ и $f''(x^*) > 0$ в любой точке этого отрезка, то $f(x)$ – унимодальная функция на $[a, b]$.

Заметим, что условие $f''(x^*) > 0$ определяет множество точек, на котором функция является выпуклой (вниз). Условие же $f''(x^*) < 0$ определяет вогнутую функцию, которая на отрезке $[a, b]$ имеет максимум и также является унимодальной.

3. ПОИСКОВЫЕ МЕТОДЫ

3.1 Методы точечного оценивания

3.1.1 Метод обратного переменного шага

Данный метод, называемый иногда ***методом сканирования***, не требует предварительного определения отрезка унимодальности.

Пусть функция $y=f(x)$ является унимодальной на некотором промежутке. Предположим, что произвольная точка x_0 этого промежутка является исходной для поиска точки x^* локального минимума и число ε – заданная точность нахождения x^*. Обозначим через Δ_0 произвольное приращение аргумента x и, сделав один шаг от точки x_0, получим новое значение аргумента $x_1 = x_0 + \Delta_0$.

Сравним значения функции $y_0 = f(x_0)$ и $y_1 = f(x_1) = f(x_0+\Delta_0)$. Возможны три различных продолжения в приближении к точке x^*.

I $y_1 < y_0$ - произошло уменьшение значения функции. Тогда примем в качестве нового стартового значения $x_0^{(1)} = x_1$, и сделаем шаг Δ_0 от этой точки $x_0^{(1)}$ к точке $x_1^{(1)}$, т.е. $x_1^{(1)} = x_0^{(1)} + \Delta_0$. Если окажется $y_1^{(1)} < y_0^{(1)}$, то снова сделаем шаг Δ_0 от новой стартовой точки $x_0^{(2)} = x_1^{(1)}$ и т.д. На некотором k-м шаге произойдет увеличение значения функции, т.е. $y_1^{(k)} > y_0^{(k)}$, и если при этом $|\Delta_0| < \varepsilon$, то принимаем $x^* \approx x_0^{(k)}$. В противном случае полагаем, что точка $x_0^{(k)}$ является исходной для продолжения вычислений по следующей схеме II.

II $y_1 > y_0$ - значение функции возросло. В этом случае полагаем, что начальной точкой вычислений является точка $x_0 = x_1$, а меньшим шагом в продолжении счета – величина $\Delta_1 = -\beta\,\Delta_0$, где β – некоторое положительное число, $\beta < 1$. Далее производим вычисления по схеме I или II, вплоть до достижения заданной точности.

III $y_1 = y_0$. В этом практически маловероятном случае (опущенном при рассмотрении случаев I и II) естественно либо принять $x^* = (x_0 + x_1)/2$ при достижении заданной точности $|\Delta| < \varepsilon$, либо следовать схеме II.

Поиск минимума функции одной переменной указанным методом представляет собой колебательный процесс, совершающийся около точки x^* локального минимума функции $f(x)$ с непрерывно меняющейся амплитудой.
В некоторых модификациях данного метода при получении “удачного” шага его значение увеличивают, т.е. $\Delta_1 = \alpha\,\Delta_0$, $\alpha > 1$, однако это часто приводит к потере сходимости алгоритма, поэтому данную операцию можно считать целесообразной лишь на первых шагах алгоритма при большом удалении от точки оптимума.
Первые несколько шагов сканирования также можно использовать для поиска отрезка унимодальности для описанной ниже группы методов последовательного сокращения отрезков, имеющих хорошую сходимость.

3.1.2 Метод квадратичной аппроксимации

Основан на аппроксимации функции полиномом второго порядка в некоторой окрестности и расчета на его основе координаты точки оптимума.
Пусть известны значения функции в трех точках x_0, x_1, x_2, составляющие соответственно y_0, y_1, y_2. Тогда функцию $f(x)$ можно аппроксимировать полиномом

$$g(x)=a_0 + a_1(x- x_0) + a_2(x- x_0)\,(x- x_1)$$

с коэффициентами

$$a_0 = y_0 \quad ;$$

$$a_1 = \frac{y_1 - y_0}{x_1 - x_0} \quad ;$$

$$a_2 = \frac{1}{x_2 - x_1}\left(\frac{y_2 - y_0}{x_2 - x_0} - \frac{y_1 - y_0}{x_1 - x_0}\right) \quad .$$

Оптимальное значение оценивается по формуле

$$x^* \approx \frac{x_1 + x_0}{2} - \frac{a_1}{2a_2} \ .$$

3.1.3 Метод Пауэлла

Основан на последовательном применении квадратичной аппроксимации. Рассмотрим алгоритм метода задав начальную точку – x_0 и шаг по оси x - Δx.

1. Вычислить $x_1 = x_0 + \Delta x$.
2. Вычислить $y_0 = f(x_0)$ и $y_1 = f(x_1)$.
3. Если $y_0 > y_1$, то вычислить $x_2 = x_0 + 2\Delta x$, иначе, т.е. если $y_0 \le y_1$, то $x_2 = x_0 - \Delta x$.
4. Вычислить $y_2 = f(x_2)$

5 Используя значения x_0 , x_1 , x_2 и y_0 , y_1 , y_2 вычислить x^* с помощью квадратичной аппроксимации.

6 Найти $y_{\min} = \min(y_0 , y_1 , y_2)$ и $x_{\min}$, соответствующую $y_{\min}$.

7 Проверить условие окончания поиска $| x_{\min} - x^*| \le \varepsilon$, где ε - заданная точность поиска. Если условие выполняется закончить поиск; в противном случае перейти к следующему шагу.

8 Выбрать "наилучшую" точку ($x_{\min}$ или x^*) и две точки по обе стороны от нее и перейти к шагу 5). Если выбранная точка является "крайней", то отбрасывается точка с наибольшим значением целевой функции.

3.2 Методы последовательного сокращения отрезка унимодальности

Основой многих одномерных численных методов является сокращение отрезка унимодальности, а именно: построение последовательности отрезков $[a_k, b_k]$, стягивающихся к точке x^* – минимуму функции на исходном отрезке. Методы оптимизации отличаются друг от друга лишь различным выбором точек на начальном отрезке унимодальности.

Общая последовательность реализации методов:

- выбор точек на начальном отрезке унимодальности;
- вычисление значений функции в этих точках и сравнение этих значений;
- определение нового отрезка;
- проверка критерия останова.

Для определения начального отрезка $[a_0, b_0]$, на котором находится один экстремум (промежутка унимодальности) могут быть использованы как аналитические так и графические методы анализа функции (если это возможно), а также, как отмечалось ранее, такие методы, как метод сканирования или методы аппроксимации функций.

3.2.1 Равномерный поиск

Равномерный поиск является примером одномерного поиска, когда точки, в которых вычисляется значение функции, выбираются заранее. Начальный отрезок $[a_0, b_0]$ делится на равные отрезки длиной d сеткой из n точек $a_0 + d \cdot k$ для k=1, ..., n и тогда $b_0 = a_0+(n+1)\cdot d$. Функция $f(x)$ вычисляется в каждом из n узлов полученной сетки, и выбирается точка x, в которой она имеет минимальное значение. Этот метод обычно используется для начальной оценки отрезка $[x - d, x + d]$, которому принадлежит минимум. Для достижения высокой точности в этом методе необходимо осуществить большое число вычислений функции. Однако его преимуществом является возможность поиска глобальных экстремумов функции, когда нет уверенности в правильном определении начального отрезка унимодальности.

3.2.2 Метод локализации оптимума

С целью повышения точности и уменьшения числа расчетов $f(x)$ можно усовершенствовать стратегию равномерного поиска. После обычного равномерного поиска в новом отрезке $[x - d, x + d]$ вновь производится разбиение на то же количество частей, определяется новый отрезок и т.д. Поиск производится до тех пор пока длина нового отрезка не станет меньше заданной точности. Доказано, что данный метод работает наиболее эффективно если текущий отрезок унимодальности делится на четыре части части, т.е. $n = 3$. При этом значение целевой функции в середине нового отрезка уже известно и на каждой последующей итерации требуется вычислить только два значения функции. В результате каждый раз отрезок унимодальности делится надвое, поэтому метод получил также название ***метода деления отрезка пополам***.

3.2.3 Общая схема сужения промежутка унимодальности

В методах, рассматриваемых далее, для дальнейшего сужения промежутка унимодальности используют следующую идею.

Возьмем две точки x_1 и x_2 , принадлежащие начальному отрезку $[a_0 , b_0]$ такие, что $x_1 < x_2$. В каждом из трех следующих очевидных случаев можно указать отрезок меньших размеров $[a_1, b_1]$, содержащий точку минимума x^* и принадлежащий первоначальному отрезку (рисунок 1):

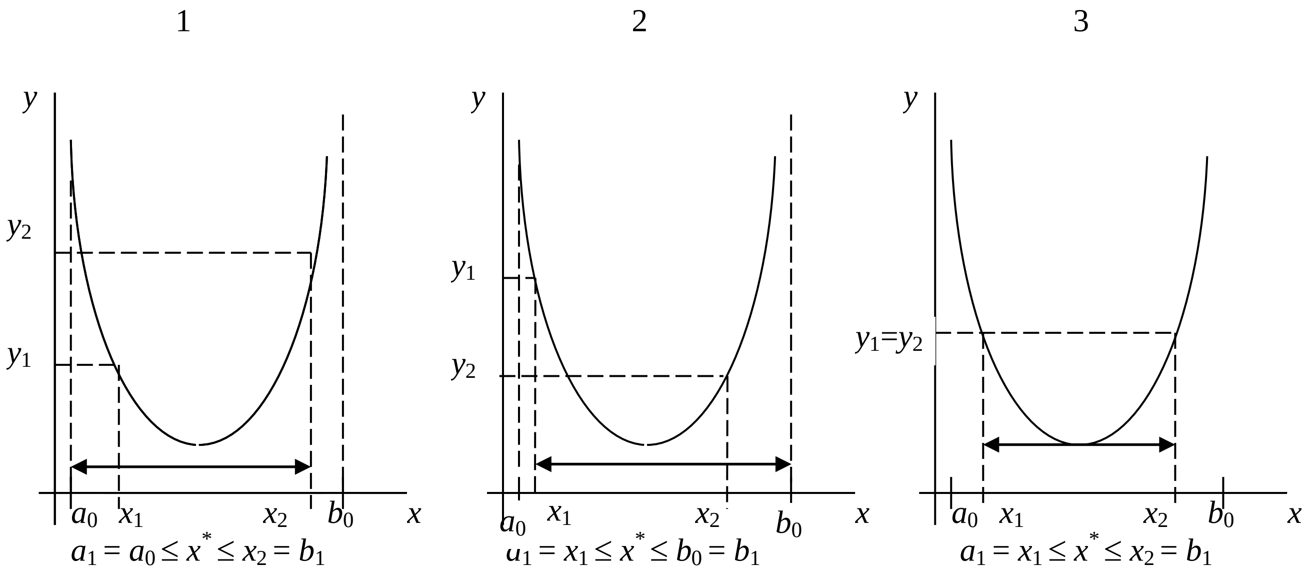


Рисунок 1 – Возможные ситуации при сужении отрезка

1 Если $y_1 = f(x_1) < y_2 = f(x_2)$, то положим $a_1 = a_0$ и $b_1 = x_2$ и получим меньший отрезок унимодальности $[a_1, b_1]$.
2 Если $y_1 = f(x_1) > y_2 = f(x_2)$, то естественно принять $a_1 = x_1$ и $b_1 = b_0$.
3 Если $y_1 = f(x_1) = y_2 = f(x_2)$, то $a_1 = x_1$ и $b_1 = x_2$.

3.2.4 Метод половинного деления

Метод половинного деления, называемый также ***методом дихотомии***, является процедурой последовательного поиска. Пусть определен отрезок $[a_0, b_0]$, которому принадлежит точка локального минимума x^*, и функция $f(x)$ является унимодальной на этом отрезке. Далее для сужения промежутка унимодальности используем две точки x_1 и x_2, расположенные симметрично на расстоянии $\delta > 0$ от середины отрезка:

$$x_1 = \frac{a_0 + b_0}{2} - \delta \ ;$$

$$x_2 = \frac{a_0 + b_0}{2} + \delta \ .$$

Константа δ должна быть меньше допустимой конечной длины отрезка, $\Delta_k = b_k - a_k > 0$.

Рассчитываем значение функции в этих точках $y_1=f(x_1)$ и $y_2=f(x_2)$ и в зависимости от их соотношения новые границы отрезка унимодальности $[a_1, b_1]$ будут следующие:

- $y_1 < y_2$, $a_1=a_0$ и $b_1=x_2$;
- $y_1 > y_2$, $a_1=x_1$ и $b_1=b_0$;
- $y_1 = y_2$, $a_1=x_1$ и $b_1= x_2$.

Название *метода половинного деления* мотивировано тем, что если величина ε достаточно мала, то длина отрезка унимодальности $(b - a)$ уменьшается почти вдвое.

В этом суженном промежутке $[a_1, b_1]$ вновь рассчитываются две точки $x_1^{(1)}$ и $x_2^{(1)}$, симметричные относительно его середины, и значение функции в этих точках. Процедура будет повторяться до тех пор, пока не будет выполняться условие $\Delta_k = b_k - a_k \le \varepsilon$, где ε – точность поиска, и тогда в качестве точки локального минимума можно приближенно принять середину отрезка $x^* \approx \frac{a_k + b_k}{2}$.

3.2.5 Метод золотого сечения

Термин “золотое сечение” ввел Леонардо да Винчи. Точка x_1 является золотым сечением отрезка $[a, b]$, если отношение длины $b-a$ всего отрезка к длине $b-x_1$ большей части равно отношению длины большей части к длине $x_1 - a$ меньшей части (рисунок 2), т.е. x_1 – золотое сечение, если справедливо соотношение $\frac{b-a}{b-x_1} = \frac{b-x_1}{x_1-a}$. Аналогично, точка x_2 симметричная точке x_1 относительно середины отрезка $[a, b]$, является вторым золотым сечением этого отрезка.

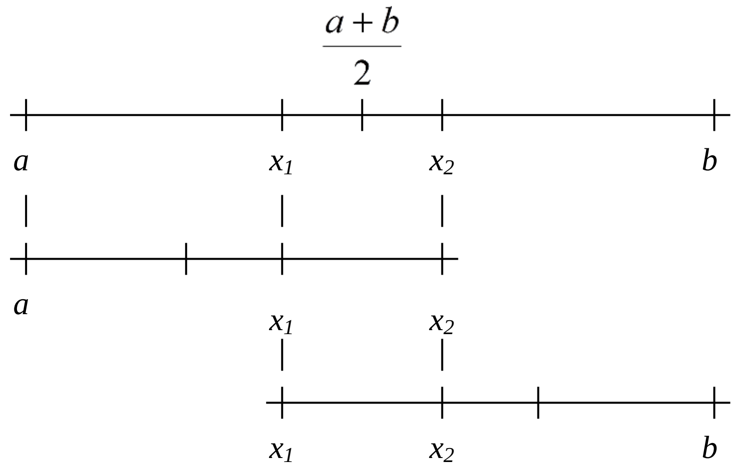


Рисунок 2 – Метод золотого сечения

Отметим свойство золотого сечения: точка x_1 одновременно является золотым сечением отрезка $[a, x_2]$, а другая точка x_2 - золотым сечением отрезка $[x_1, b]$.

Суть метода золотого сечения заключается в следующем. Сначала на исходном отрезке $[a_0, b_0]$ находятся точки x_1 и x_2 по следующим формулам:

$x_1= a_0 +(1-k)\cdot(b_0-a_0)$;

$x_2= a_0 +k\cdot(b_0-a_0)$;

где $k = \dfrac{\sqrt{5}-1}{2} = 0{,}618$ – коэффициент сжатия.

Затем вычисляются значения функции в точках x_1 и x_2 , т.е. $y_1=f(x_1)$ и $y_2=f(x_2)$. При этом возможны два случая:

1 $y_1 < y_2$, в этом случае новый отрезок будет равен: $a_1=a_0$ и $b_1=x_2$. В этом отрезке вновь выбираются две точки: $x_1^{(1)} = a_1+ (1- k)\cdot(b_1-a_1)$ и $x_2^{(1)}=x_1$.

2 $y_1 < y_2$, тогда новый отрезок будет составлять: $a_1=x_1$ и $b_1=b_0$. В новом отрезке также выбираются две точки: $x_1^{(1)} = x_2$ и $x_2^{(1)}= a_1 + k\cdot(b_1 - a_1)$.

И в первом и во втором случаях рассчитывается лишь одна новая точка (вторая известна). В новой точке рассчитывается значение функции и вновь производится сравнение в двух точках, и в зависимости от этого выбирается новый отрезок. Процедура повторяется до тех пор, пока не будет выполняться условие $(b_k - a_k) \le \varepsilon$, где ε – точность поиска.

3.2.6 Метод Фибоначчи

В методе Фибоначчи требуется, чтобы общее число n вычислений функции было выбрано заранее, так как точки, в которых производится вычисление, определяются по формулам:

$$x_1^{(k)} = a_k + \frac{F_{n-k-2}}{F_{n-k}}\cdot(b_k - a_k), \qquad k=0, ..., n-2 ;$$

$$x_2^{(k)} = a_k + \frac{F_{n-k-1}}{F_{n-k}} \cdot (b_k - a_k), \qquad k=0, ..., n-2 ;$$

где $F_i = F_{i-1}+F_{i-2}$, i=1, 2, ... $F_0 = F_1$= 1 - называется последовательностью чисел Фибоначчи (1, 1, 2, 3, 5, 8, 13, 21, 34, 55, ... , т.е. каждый член последовательности рассчитывается как сумма двух предыдущих членов).

Алгоритм поиска.

Выбирается начальный отрезок $[a_0, b_0]$ и число вычислений n таким образом, чтобы $F_n > \frac{b_0 - a_0}{\Delta}$, где $\Delta > 0$ – конечная длина отрезка.

Затем рассчитываются координаты двух точек:

$$x_1^{(0)} = a_0 + \frac{F_{n-2}}{F_n} \cdot (b_0 - a_0),$$

$$x_2^{(0)} = a_0 + \frac{F_{n-1}}{F_n} \cdot (b_0 - a_0)$$

и значение функции в этих точках $y_1=f(x_1^{(0)})$ и $y_2=f(x_2^{(0)})$.

В случае $y_1 < y_2$

$a_1=a_0$ и $b_1=x_2^{(0)}$, $\qquad x_1^{(1)} = a_1 + \frac{F_{n-3}}{F_{n-1}} \cdot (b_1 - a_1)$, $\qquad x_2^{(1)}= x_1^{(0)}$.

В случае $y_1 > y_2$

$a_1= x_1^{(0)}$ и $b_1= b_0$, $\qquad x_1^{(1)}= x_2^{(0)}$, $\qquad x_2^{(1)} = a_1 + \frac{F_{n-2}}{F_{n-1}} \cdot (b_1 - a_1)$.

Таким образом данная процедура повторяется $(n - 2)$ раза.

При $k = (n - 2)$ точки x_k и y_k совпадают и соответствуют середине отрезка, поэтому чтобы обеспечить дальнейшее сокращение отрезка, точка последнего вычисления функции перемещается вправо на величину константы различимости δ >0, которая выбирается заранее существенно меньше заданной точности.

ЗАДАНИЕ НА КОНТРОЛЬНУЮ РАБОТУ 1

1 Для заданной целевой функции найти аналитическое решение задачи одномерной минимизации $f(x) \to \min, \quad x \in X \ (X \subset R)$ и найти промежуток $(X \subset R)$, на котором функция унимодальна.

2 Произвести графический анализ функции с отображением первой и второй ее производных.

3 Найти минимум функции методом обратного переменного шага для заданной точности ε и начальной точки x_0 .

ПРИМЕР ВЫПОЛНЕНИЯ КОНТРОЛЬНОЙ РАБОТЫ 1

Найти минимум функции $f(x)= x^2+2x$ с точностью $\varepsilon= 0,8$ для начальной точки $x_0 = 10$. Константу различимости примем равной $\delta= 0,2$.

А.1 Аналитический анализ функции

Представленная функция и ее производные непрерывны, поэтому определяем первую производную и приравниваем ее к 0, т.е

$$f'(x) = 2x + 2 = 0;$$

откуда следует, что функция имеет один экстремум в точке $x^* = -1$.

Далее находим значение второй производной в точке x^*: $f''(x^*) = 2 > 0$, т.е. в указанной точке имеем глобальный минимум функции, который составляет: $f(x^*)= (-1)^2+ 2\cdot(-1)= -1$

Поскольку вторая производная всегда положительна, функция унимодальна на интервале $(-\infty, \infty)$.

А.2 Графический анализ функции

Построим график функции, ее первой и второй производных в окрестности точки x^*, как показано на рисунке А.1.

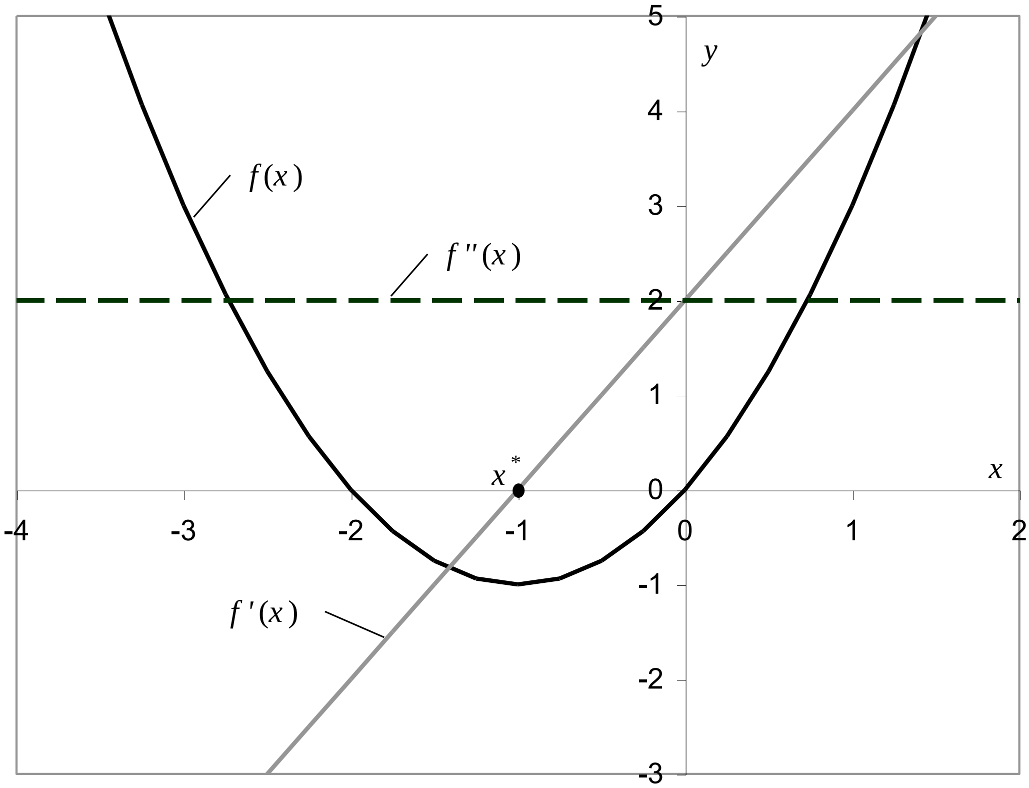


Рисунок А.1 – Графический анализ функции

Полученные кривые подтверждают выводы, сделанные в разделе А.1.

А.3 Поиск минимума методом обратного переменного шага

Принимаем для начальной точки x_0 = 10 величину начального шага Δ_0 = 5 , а коэффициент сжатия β = 0,3.

А.3.1 Для определения знака Δ в начальной точке x_0 = 10 сравним значения

$f(x_0) = f(10) = 120$,

$f(x_0 + h_0) = f(15) = 255$ и

$f(x_0 - h_0) = f(5) = 35$.

Поскольку $f(x_0 + \Delta_0) > f(x_0) > f(x_0 - \Delta_0)$, то величина шага должна быть отрицательной, т.е. $\Delta_0 = -5$.

Таким образом имеем $x_1 = x_0 + \Delta_0 = 10 – 5 = 5$, $x_0^{(1)} = x_1$.

А.3.2 Координаты следующей точки $x_1^{(1)} = x_0^{(1)} - \Delta = 0$, в которой $y_1^{(1)} = f(0) = 0$. Т.к. $y_1^{(1)} = f(x_1^{(1)}) = 0 < y_1^{(1)} = f(x_0^{(1)}) = 35$ принимаем $x_0^{(2)} = x_1^{(1)}$ и продолжаем движение.

А.3.3 $x_1^{(2)} = x_0^{(2)} - \Delta_0 = -5$; $y_1^{(2)} = f(x_1^{(2)}) = 15$. Значение функции в новой точке увеличилось, т.е. $y_1^{(2)} = f(x_1^{(2)}) = 15 > y_0^{(2)} = f(x_0^{(2)}) = 0$, и т.к. $\Delta_0 > \varepsilon = 0,8$ уменьшаем шаг $\Delta_1 = - 0,3\ h_0 = 1,5$ и продолжаем движение.

А.3.4 Дальнейшую последовательность вычислений представим в виде пар координат $\begin{bmatrix} x_i^{(k)} \\ y_i^{(k)} \end{bmatrix}$, где i= 0; 1 , а верхний индекс k является номером очередного шага сканирования.

$x_0^{(3)} = x_1^{(2)} = -5$; $x_1^{(3)} = x_0^{(3)} + \Delta_1 = -3{,}5$; $y_1^{(3)} = 5{,}25 < y_0^{(3)} = 15$.

А.3.5 $x_0^{(4)} = x_1^{(3)}$; $x_1^{(4)} = x_0^{(4)} + \Delta_1 = -2$; $y_1^{(4)} = 0 < y_0^{(4)} = 5{,}25$.

А.3.6 $x_0^{(5)} = x_1^{(4)}$; $x_1^{(5)} = x_0^{(5)} + \Delta_1 = -0{,}5$; $y_1^{(5)} = -0{,}75 < y_0^{(5)} = 0$.

А.3.7 $x_0^{(6)} = x_1^{(5)}$; $x_1^{(6)} = x_0^{(6)} + \Delta_1 = 1$; $y_1^{(6)} = 3 > y_0^{(6)} = -0{,}75$; $\Delta_1 > \varepsilon$, поэтому уменьшаем шаг $\Delta_2 = -0{,}3\,\Delta_1 = -0{,}45$.

А.3.8 $x_0^{(7)} = x_1^{(6)}$; $x_1^{(7)} = x_0^{(7)} + \Delta_2 = 1 - 0{,}45 = 0{,}55$; $y_1^{(7)} = 1{,}4 < y_0^{(7)} = 3$.

А.3.9 $x_0^{(8)} = x_1^{(7)}$; $x_1^{(8)} = x_0^{(8)} + \Delta_2 = 0{,}1$; $y_1^{(8)} = 0{,}21 < y_0^{(8)} = 1{,}4$.

А.3.10 $x_0^{(9)} = x_1^{(8)}$; $x_1^{(9)} = x_0^{(9)} + \Delta_2 = -0{,}35$; $y_1^{(9)} = -0{,}58 < y_0^{(9)} = 0{,}21$.

А.3.11 $x_0^{(10)} = x_1^{(9)}$; $x_1^{(10)} = x_0^{(10)} + \Delta_2 = -0{,}8$; $y_1^{(10)} = -0{,}96 < y_0^{(10)} = -0{,}58$.

А.3.12 $x_0^{(11)} = x_1^{(10)}$; $x_1^{(11)} = x_0^{(11)} + \Delta_2 = -1{,}25$; $y_1^{(11)} = -0{,}94 > y_0^{(11)} = -0{,}96$.

Значение функции в последней точке возросло, а поскольку $\Delta_2 = 0{,}45 < \varepsilon$**,** поиск можно прекратить.

Вывод: Минимум функции достигается в точке $x^* \approx x_0^{(11)} = -0{,}8$ с погрешностью $\varepsilon = 0{,}8$. Количество итераций равно 11 при 14 вычислениях функции. Значение функции в этой точке равно -0,96.

ЗАДАНИЕ НА КОНТРОЛЬНУЮ РАБОТУ 2

1 Найти минимум функции методом Пауэлла для заданной точности ε и начальной точки x_0.

2 Определить начальный промежуток унимодальности $[a_0, b_0]$, взяв за основу первые несколько шагов метода обратного переменного шага.

3 Найти минимум функции методом локализации оптимума для заданной точности ε .

ПРИМЕР ВЫПОЛНЕНИЯ КОНТРОЛЬНОЙ РАБОТЫ 2

А.4 Поиск минимума методом Пауэлла

Принимаем величину начального шага $\Delta x = 5$ для начальной точки $x_0 = 10$.
А.4.1 Вычисляем $x_1 = x_0 + \Delta x = 10+5 = 15$; $y_0 = f(10) = 120$ и $y_1 = f(x_1) = 255$.

А.4.2 Поскольку $y_0 \le y_1$, $x_2 = x_0 - \Delta x = 10 – 5 = 5$; $y_2 = f(x_2) = f(5) = 35$.

А.4.3 Используя значения x_0 , x_1 , x_2 и y_0 , y_1 , y_2 вычисляем x^* с помощью квадратичной аппроксимации

$$a_0 = y_0 = 120 \quad ;$$

А.4.4 $$a_1 = \frac{y_1 - y_0}{x_1 - x_0} = \frac{255 - 120}{15 - 10} = 27 \quad ;$$

$$a_2 = \frac{1}{x_2 - x_1}\left(\frac{y_2 - y_0}{x_2 - x_0} - \frac{y_1 - y_0}{x_1 - x_0}\right) = \frac{1}{5 - 15}\left(\frac{35 - 120}{5 - 10} - \frac{255 - 120}{15 - 10}\right) = 1 \quad .$$

$$x^* \approx \frac{x_1 + x_0}{2} - \frac{a_1}{2a_2} = \frac{15 + 10}{2} - \frac{27}{2 \cdot 1} = 12{,}5 - 13{,}5 = -1$$

А.4.5 Проверяем условие окончания поиска $| x_{\min} - x^*| = | x_2 - x^*| = | 5 + 1| > \varepsilon = 0{,}8$ и продолжаем поиск.

А.4.6 Отбрасываем точку с наибольшим значением целевой функции, т.е. $y_1 = y_{\max} = 255$ при $x_1 = 15$.

А.4.7 Для оставшихся точек $x_0 = 10$, $x_1 = -1$, $x_2 = 5$ рассчитываем новые коэффициенты полинома $a_0 = 120$, $a_1 = 11$, $a_2 = 1$ и оптимальное значение $x^* = -1$.

А.4.8 Условие окончания поиска выполняется $|x_{\min} - x^*| = |x_1 - x^*| = |-1 + 1| < \varepsilon = 0{,}8$, поэтому поиск можно прекратить.

Вывод: Минимум функции достигается в точке $x^* \approx = -1$. Количество итераций равно 2 при 4 вычислениях функции. Значение функции в этой точке равно - 1. Следует отметить, что заданная функция является квадратичной, поэтому при квадратичной аппроксимации на первом же шаге определяется оптимальное значение.

А.5 Определение начального отрезка унимодальности

Для определения отрезка унимодальности используем начальные вычисления метода обратного переменного шага, описанные в пункте А.3 для заданной начальной точки $x^{(0)} = 10$ и начальном шаге $\Delta = -5$.

$f(x_0) = f(10) = 120$;

$f(x_0 + \Delta_0) = f(10 - 5) = 35$; $f(x_0 + \Delta_0) < f(x_0)$;

$f(x_0 + 2\Delta_0) = f(10 - 10) = 0$; $f(0) < f(5)$;

$f(x_0 + 3\Delta_0) = f(10 - 15) = 15$; $f(-5) > f(0)$ и т.о. точка минимума $5 > x^* > -5$.

Вывод: минимум функции находится на отрезке [-5, 5] , который можно принять в качестве начального отрезка унимодальности.

А.6 Расчет минимума функции методом локализации оптимума

А.6.1 Делим исходный отрезок на четыре равные части и рассчитываем значения функции в полученных точках

$y_1 = f(x_1) = f(-2{,}5) = 1{,}25$;

$y_2 = f(x_2) = f(0) = 0$;

$y_3 = f(x_3) = f(2{,}5) = 11{,}25$.

Учитывая, что значения функции на концах исходного отрезка известны и минимум достигается в точке $x_2 = 0$, получаем новый отрезок унимодальности [-2,5 ; 2,5]. Проверяем условие окончания поиска: $\Delta_1 = b_1 - a_1 = 2{,}5 + 2{,}5 > \varepsilon = 0{,}8$ и продолжаем решение.

А.6.2 Делим новый отрезок на четыре равные части и рассчитываем значения функции в полученных точках. При этом в середине отрезка значение функции уже известно, т.е. $y_2 = f(x_2) = f(0) = 0$

$y_1 = f(x_1) = f(-1{,}25) = -0{,}94$;

$y_3 = f(x_3) = f(1{,}25) = 4{,}06$.

Поскольку минимум достигается в точке $x_1 = -1{,}25$, получаем новый отрезок унимодальности [-2,5 ; 0]. Проверяем условие окончания поиска: $\Delta_2 = b_2 - a_2 = 0 + 2{,}5 > \varepsilon = 0{,}8$ и продолжаем решение.

А.6.3 Вновь делим новый отрезок на четыре равные части и рассчитываем значения функции в полученных точках

$y_1 = f(x_1) = f(-1{,}88) = -0{,}23$; $y_3 = f(x_3) = f(-0{,}625) = -0{,}86$.

Поскольку минимум достигается в точке x_2 =-1,25, получаем новый отрезок унимодальности [-1,88 ; -0,625]. Проверяем условие окончания поиска: $\Delta_3 = b_3 - a_3 = -0{,}625+1{,}88 = 1{,}25 > \varepsilon = 0{,}8$ и продолжаем решение.

А.6.4 Вновь делим новый отрезок на четыре равные части и рассчитываем значения функции в полученных точках $y_1 = f(x_1) = f(-1{,}56) = -0{,}23$; $y_3 = f(x_3) = f(-0{,}94) = -0{,}86$.
Поскольку минимум достигается в точке x_2 =-1,25, получаем новый отрезок унимодальности [-1,25 ; -0,625]. Длина нового отрезка составляет $\Delta_4 = b_4 - a_4 = -0{,}625+1{,}25 = 0{,}625 < \varepsilon = 0{,}8$ и условие окончания поиска выполняется.

Вывод: Конечный отрезок унимодальности [-1,25;-0,625]. Принимаем за оптимальную точку его середину $x^* = -0{,}9375$. Значение функции в этой точке равно - 0,9961, количество итераций – 4 при 9 вычислениях функции.

ЗАДАНИЕ НА КОНТРОЛЬНУЮ РАБОТУ 3

1 Определить минимум функции методом половинного деления для заданной точности ε и константы различимости δ.

2 Найти минимум функции методом золотого сечения с заданной точностью.

3 Произвести поиск минимума функции методом Фибоначчи для заданной точности ε и константы различимости δ.

ПРИМЕР ВЫПОЛНЕНИЯ КОНТРОЛЬНОЙ РАБОТЫ 3

А.7 Расчет минимума функции методом половинного деления

Выбираем значение окрестности равной константе различимости $\delta = 0{,}2 < \varepsilon$.

А.7.1 Выбираем две точки, симметрично расположенные относительно середины отрезка $[a_0, b_0]$:

$x_1^{(0)} = (-5+5)/2 - 0{,}2 = -0{,}2$,

$x_2^{(0)} = (-5+5)/2 + 0{,}2 = 0{,}2$

и рассчитываем значение функции в этих точках:

$y_1^{(0)} = f(x_1^{(0)}) = (-0{,}2)^2 + 2\cdot(-0{,}2) = -0{,}36$,

$y_2^{(0)} = f(x_2^{(0)}) = 0.2^2 + 2\cdot 0.2 = 0.44$.

Так как $y_1^{(0)} < y_2^{(0)}$, определяем границы нового отрезка:

$a_1 = a_0 = -5$,

$b_1 = x_2^{(0)} = 0{,}2$.

А.7.2 Вновь выбираем две точки на отрезке $[a_1, b_1]$:

$x_1^{(1)} = (-5+0{,}2)/2 - 0{,}2 = -2{,}6$,

$x_2^{(1)} = (-5+0{,}2)/2 + 0{,}2 = -2{,}2$

и рассчитываем значение функции в этих точках:

$y_1^{(1)} = (-2{,}6)^2 + 2\cdot(-2{,}6) = 1{,}56$,

$y_2^{(1)} = (-2{,}2)^2 + 2\cdot(-2{,}2) = 0{,}44$.

Поскольку $y_1^{(1)} > y_2^{(1)}$, границы нового отрезка-

$a_2 = x_1^{(1)} = -2{,}6$,

$b_2 = b_1 = 0{,}2$.

Проверяем условие окончания оптимизации и так как

$\Delta_2 = b_2 - a_2 = 0{,}2+2{,}6 = 2{,}8 > 0{,}8$,

продолжаем поиск.

А.7.3 Отрезок $[a_2 , b_2] = [x_1^{(1)} , b_1] = [-2{,}6; 0{,}2]$.

$x_1^{(2)} = -1{,}4$, $x_2^{(2)} = -1{,}0$, $y_1^{(2)} = -0{,}84 > y_2^{(2)} = -1{,}00$,

$a_3 = x_1^{(2)} = -1{,}4$, $b_3 = b_2 = 0{,}2$

и проверяем условие окончания оптимизации:

$\Delta_3 = b_3 - a_3 = 0{,}2+1{,}4 > 0{,}8$.

А.7.4 Отрезок $[a_3, b_3] = [x_1^{(2)}, b_2] = [-1{,}4; 0{,}2]$.
$x_1^{(3)} = -0{,}8$, $x_2^{(3)} = -0{,}4$, $y_1^{(3)} = -0{,}96 < y_2^{(3)} = -0{,}64$.

А.7.5 Отрезок $[a_4, b_4] = [a_3, x_2^{(3)}] = [-1{,}4; -0{,}4]$.
$x_1^{(4)} = -1{,}1$, $x_2^{(4)} = -0{,}7$, $y_1^{(4)} = -0{,}99 < y_2^{(4)} = -0{,}91$.

Новый отрезок имеет границы:

$[a_5, b_5] = [a_4, x_2^{(4)}] = [-1{,}4; -0{,}7]$,
его длина составляет $\Delta_5 = b_5 - a_5 = -0{,}7+1{,}4 = 0{,}7 < \varepsilon = 0{,}8$
и таким образом условие окончания поиска выполняется.

За точку локального минимума, найденную с заданной точностью принимаем середину отрезка $[a_5, b_5]$: $x^* \approx (a_5+b_5)/2 = (-1{,}4-0{,}7)/2 = -1{,}05$.
Значение функции в этой точке

$f(x^*) = (-1{,}05)^2 + 2\cdot(-1{,}05) = -0{,}9975$.

Вывод: Конечный отрезок унимодальности [-1,4 , -0,7]. Принимаем за оптимальную точку середину этого отрезка $x^* = -1{,}05$. Значение функции в этой точке равно - 0,9975, количество итераций равно 5 при 10 вычислений функции.

А.8 Поиск минимума функции методом золотого сечения

А.8.1 Для начального отрезка $[a_0, b_0] = [-5, 5]$ рассчитываем две точки:

$x_1^{(0)} = -5 + (1-0{,}618)\cdot(5+5) = -1{,}18$,
$x_2^{(0)} = -5 + 0{,}618\cdot(5+5) = 1{,}18$
и значение функции в этих точках
$y_1^{(0)} = f(x_1^{(0)}) = (-1{,}18)^2 + 2\cdot(\cdot-1{,}18) = -0{,}968$,
$y_2^{(0)} = f(x_2^{(0)}) = (1{,}18)^2 + 2\cdot(1{,}18) = 3{,}7524$.
$y_1^{(0)} < y_2^{(0)}$, следовательно новый отрезок имеет границы
$[a_1, b_1] = [a_0, x_2^{(0)}] = [-5; 1{,}18]$.

Проверяем условие окончания поиска: $\Delta_1 = b_1 - a_1 = 1{,}18+5 > \varepsilon = 0{,}8$ и продолжаем решение.

А.8.2 Для отрезка $[a_1, b_1] = [-5; 1{,}18]$ рассчитываем новые точки
$x_1^{(1)} = -5 + (1-0{,}618)\cdot(1{,}18+5) = -2{,}639$,
$x_2^{(1)} = x_1^{(0)} = -1{,}18$
и значение функции в точке $x_1^{(1)}$ -
$y_1^{(1)} = f(x_1^{(1)}) = (-2{,}639)^2 - 2\cdot 2{,}639 = 1{,}686$.
Значение функции в точке $x_2^{(1)}$ совпадает со значением в $x_1^{(0)}$.
$y_2^{(1)} = y_1^{(0)} = -0{,}968$,
$y_1^{(1)} > y_2^{(1)}$, следовательно новый отрезок составит:
$[a_2, b_2] = [x_1^{(1)}, b_1] = [-2{,}639; 1{,}18]$.

Проверяем условие окончания поиска: $\Delta_2 = b_2 - a_2 = 1{,}18+2{,}639 > 0{,}8$ и продолжаем решение.

A.8.3 Отрезок $[a_2, b_2] = [x_1^{(1)}, b_1] = [-2{,}639; 1{,}18]$.

$x_1^{(2)}= x_2^{(1)}= x_1^{(0)}=-1{,}18$, $x_2^{(2)}= -2{,}639+ 0{,}618\cdot(1{,}18+2{,}639)= -0{,}279$,

$y_1^{(2)} = y_2^{(1)} = -0{,}968 <$ $y_2^{(2)} ==(-0{,}279)^2+ 2\cdot(-0{,}279)= -0{,}480$.

A.8.4 Отрезок $[a_3, b_3] = [a_2 , x_2^{(2)}] = [-2{,}639; -0{,}279]$.

$x_1^{(3)}=-2{,}639+0{,}382\cdot(-0{,}279+2{,}639) =-1{,}737$, $x_2^{(3)}=x_1^{(2)}=x_2^{(1)}=x_1^{(0)}=-1{,}18$,

$y_1^{(3)} =(-1{,}737)^2+ 2\cdot(-1{,}737) = -0{,}457 > y_2^{(3)} = y_1^{(2)} = - 0{,}968$.

A.8.5 Отрезок $[a_4, b_4] = [x_1^{(3)}, b_3] = [-1{,}737; -0{,}279]$.

$x_1^{(4)}= x_2^{(3)}=-1{,}18$, $x_2^{(4)}= -1{,}737+ 0{,}618\cdot(-0{,}279+1{,}737)= -0{,}836$,

$y_1^{(4)} = y_2^{(3)} = -0{,}968 >$ $y_2^{(4)} ==(-0{,}836)^2+ 2\cdot(-0{,}836)= -0{,}973$.

A.8.6 Отрезок $[a_5, b_5] = [x_1^{(4)}, b_4] = [-1{,}18 ; -0{,}279]$.

$x_1^{(5)} = x_2^{(4)} = - 0{,}836$, $x_2^{(5)}= - 1{,}18+ 0{,}618\cdot(-0{,}279+1{,}18)= -0{,}623$,

$y_1^{(5)} = y_2^{(4)} = -0{,}973 <$ $y_2^{(5)} ==(-0{,}623)^2+ 2\cdot(-0{,}623)= -0{,}858$.

Поскольку длина нового отрезка

$[a_6, b_6] = [a_5 , x_2^{(5)}] = [-1{,}18; - 0{,}623]$ составляет

$\Delta_6 = b_6 - a_6 = -0{,}623+1{,}18= 0{,}557 < = 0{,}8$,

условие окончания поиска выполняется.

Вывод: Минимум функции находится на отрезке [-1,18 , -0,623]. Количество итераций равно 6 при 7 вычислениях функции. В качестве искомой точки выбираем середину отрезка $[a_6, b_6]$, $x^* = -0{,}9015$. Значение функции в этой точке равно -0,9903.

A.9 Расчет минимума функции методом Фибоначчи

Задаемся константой различимости $\delta = 0{,}2 < \varepsilon$ и выбираем количество расчетов n из условия $F_n > \dfrac{b_0 - a_0}{\varepsilon}$ = (5+5)/0,8= 12,5:

Ряд Фибоначчи:

i	0	1	2	3	4	5	6	7	8	9	10	…
F_i	1	1	2	3	5	8	13	21	34	55	89	…

$F_6 = 13 > 12{,}5$ следовательно количество расчетов $n = 6$.

A.9.1 Рассчитываем точки на отрезке $[a_0, b_0]$; k=0 :

$x_1^{(0)}= -5+ (5+5) \cdot F_4/F_6= -5+10\cdot5/13=- 1{,}154$

$x_2^{(0)}= -5+ (5+5) \cdot F_5/F_6= -5+10\cdot8/13= 1{,}154$

и значение функции в этих точках-

$y_1^{(0)} = f(x_1^{(0)}) =(-1{,}154)^2 +2\cdot(-1{,}154)=-0{,}976$

$y_2^{(0)} = f(x_2^{(0)}) = (1{,}154)^2 +2\cdot (1{,}154) = 3{,}64$.

$y_1^{(0)} < y_2^{(0)}$, рассчитываем границы нового отрезка:

$a_1=a_0= -5$

$b_1 = x_2^{(0)} = 1,154$.

А.9.2 Для отрезка $[a_1 , b_1] = [-5 ; 1,154]$ рассчитываем новые точки; $k = 1$:

$x_1^{(1)} = -5+ (1,154+5) \cdot 3/8= - 2,692 \qquad x_2^{(1)} = x_1^{(0)} = -1,154$

$y_1^{(1)} =(-2,692)^2+ 2\cdot(-2,692) =1,863 \;\; > \;\; y_2^{(1)} = y_1^{(0)} =-0,976$.

А.9.3 Отрезок $[a_2 , b_2] = [x_1^{(1)} , b_1] = [-2,692 ; 1,154]$; $\quad k = 2$:

$x_1^{(2)} = x_2^{(1)} = -1,154 \qquad x_2^{(2)} =-2,692+ (1,154+2,692) \cdot 3/5 = -0,384$

$y_1^{(2)} = y_2^{(1)} = -0,976 \;\; < \;\; y_2^{(2)} =(-0,384)^2+ 2\cdot(-0,384) =-0,621.$

А.9.4 Отрезок $[a_3 , b_3] = [a_2 , x_2^{(2)}] = [-2,692 ; 0,384]$; $\quad k = 3$:

$x_1^{(3)} = -2,692 +(-0,384+2,692)\, 1/3= -1,923 \qquad x_2^{(3)} = x_1^{(2)} = -1,154$

$y_1^{(3)} =(-1,923)^2+ 2\cdot(-1,923) = -0,148 \;\; >y_2^{(3)} = y_1^{(2)} =-0,976$.

А.9.5 Отрезок $[a_4 , b_4] = [x_1^{(3)} , b_3] = [-1,923 ; -0,384]$.

Поскольку на данном шаге $k=n$-2= 4 точки $x_1^{(4)}$ и $x_2^{(4)}$ совпадают и соответствуют середине отрезка

$x_1^{(4)} = x_2^{(3)} = x_2^{(4)} = -1,154$,

для обеспечения заданной точности перемещаем точку последнего вычисления функции вправо на величину константы различимости δ:

$x_2^{(4)} = -1,154+0.2= -0.954$

$y_2^{(4)} =(-0,954)^2+2\cdot(-0,954)=-0,998$.

Так как $y_1^{(4)}= y_2^{(3)}= -0.976 > y_2^{(4)}$, получаем конечный отрезок $[a_5 , b_5] = [-1,154; -0,384]$. За точку минимума принимаем середину отрезка

$x^* = (a_5+ b_5)/2= (-1,154-0,384)/2= -0,769.$

Вывод: Минимум функции соответствует точке -0,769, значение функции в этой точке -0,947. Количество итераций равно 5 при 6 расчетах целевой функции.

ВАРИАНТЫ ИНДИВИДУАЛЬНЫХ ЗАДАНИЙ ДЛЯ КОНТРОЛЬНЫХ РАБОТ 1-3

№ варианта	Вид функции $f(x)$	Начальная точка $x^{(0)}$	Точность, ε	Константа различимости, δ
1	x^2-2x+1	-10	0,6	0,05
2	x^2+15x+5	5	0,5	0,04
3	x^2+30x-7	0	0,5	0,02
4	x^2+5x	-15	0,5	0,01
5	5x^2+35x-1	18	0,6	0,03
6	2x^2+25x	12	0,5	0,01
7	3x^2-15x	-8	0,7	0,02
8	5x^2-2x	10	0,6	0,02
9	3x^2+20x-1	7	0,5	0,04
10	2x^2-45x+5	-2	0,6	0,02
11	6x^2+88x	5	0,7	0,02
12	27/x +3x	20	0,45	0,03
13	x^2+12x+10	5	0,45	0,03
14	x^2-16x+13	-5	0,7	0,04
15	x^2+25x+8	3	0,5	0,02
16	x^2+9x	14	0,5	0,03
17	x^2-6x	-12	0,6	0,04
18	16/x +4x	15	0,6	0,03
19	9x^2+5x-3	-15	0,5	0,02
20	3x^2+38x+2	8	0,5	0,01
21	3x^2+18x-6	13	0,6	0,03
22	x^2+15x	10	0,45	0,02
23	x^2+30x+3	0	0,6	0,02
24	x^2-28x-2	-2	0,5	0,01
25	8/x +2x	-10	0,55	0,04

Раздел Методы безусловной многомерной оптимизации

1. ПОСТАНОВКА ЗАДАЧИ

Пусть задана функция n действительных переменных

$f(x_1, x_2, x_3, ..., x_n) = f(\mathbf{x})$, определенная на множестве $\mathbf{X} \in \mathbf{R}^n$,

где $\mathbf{x}$ - вектор- столбец, обозначающий точку в n-мерном евклидовом пространстве с координатами $x_1, x_2, x_3, ..., x_n$.

Функция $f(\mathbf{x})$ имеет локальный минимум в точке $\mathbf{x}^* \in \mathbf{X}$, если существует окрестность точки $\mathbf{x}^*$ такая, что $f(\mathbf{x}^*) \le f(\mathbf{x})$ во всех точках этой окрестности. В случае глобального минимума в точке $\mathbf{x}^*$ для всех $\mathbf{x} \in \mathbf{R}^n$ справедливо неравенство $f(\mathbf{x}^*) \le f(\mathbf{x})$.

Далее будем рассматривать задачу отыскания точек минимума функции $f(\mathbf{x})$, т.е. $f(\mathbf{x}) \to \min, \mathbf{x} \in \mathbf{R}^n$. Для приведения же задачи максимизации к задаче минимизации достаточно изменить знак целевой функции.

2. АНАЛИТИЧЕСКИЙ АНАЛИЗ ЭКСТРЕМУМА ФУНКЦИИ

Необходимым условием существования экстремума функции нескольких переменных в точке $\mathbf{x}^*$ является равенство нулю всех частных производных в этой точке:

$$\frac{\partial f}{\partial x_i}(x^*) = 0 \qquad , \quad i=1, 2, 3, ..., n$$

т.е. градиент функции равен нулевому вектору.

Данная система может иметь как одно, так и несколько решений. Точки $\mathbf{x}^*$ называются стационарными точками. Для проверки полученных точек на экстремум необходимо провести исследование вторых частных производных. При этом, рассчитывается матрица Гессе $H(\mathbf{x}^*)$, представляющая квадратную матрицу вторых частных производных $f(\mathbf{x})$, взятых в точке $\mathbf{x}^*$. Достаточным условием минимума является положительно определенная матрица H, а максимума - отрицательно определенная.

Для функции двух переменных введем следующие обозначения

$$\frac{\partial^2 f}{\partial x_1^2}(\mathbf{x}^*) = A \qquad \frac{\partial^2 f}{\partial x_2^2}(\mathbf{x}^*) = B \qquad \frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{x}^*) = C$$

Возможны два случая: $AB - C^2 < 0$ и $AB - C^2 > 0$. В первом случае вывода о наличии экстремума функции сделать нельзя. Во втором случае при $A > 0$ найденная точка является минимумом функции, при $A < 0$ - максимумом функции.

3. ГРАФИЧЕСКИЙ АНАЛИЗ ФУНКЦИИ. ПОСТРОЕНИЕ ЛИНИЙ УРОВНЯ

Область функции, в которой находится оптимальное решение, представляет собой некоторую поверхность в многомерном пространстве. Эта поверхность называется поверхностью отклика. Данную поверхность даже для случая n=2 трудно изобразить графически, поэтому на плоскости ее обычно отображают с помощью линий уровня, которые представляют собой множество точек с одинаковым значением целевой функции.

Для построения линий уровня необходимо выразить одну переменную через другую переменную и целевую функцию $x_1=F(f(x_1, x_2), x_2)$. Затем необходимо, задаваясь значениями функции, провести сканирование по второй переменной, рассчитывая при этом первую. По полученным точкам можно построить линию уровня. Затем необходимо изменить значение функции и вновь повторить процедуру. Операция повторяется столько раз, сколько необходимо провести линий уровня.

В случае неявно заданного уравнения линии уровня необходимо использовать более сложные методы для графического отображения функции.

4. ПОИСКОВЫЕ МНОГОМЕРНЫЕ МЕТОДЫ

Все методы, которые изложены далее носят шаговый характер. Одна итерация метода может включать в себя либо один шаг, либо множество шагов. Шаг считается «удачным», если значение целевой функции в новой точке не больше, чем в старой, т.е. если $f(\mathbf{x}^{(k)}) \le f(\mathbf{x}^{(k-1)})$; в противном случае шаг считается «неудачным».

4.1 Методы на основе пошаговой одномерной оптимизации

4.1.1 Метод Гаусса- Зейделя (покоординатного спуска)

В основу метода Гаусса- Зейделя положены принципы более раннего метода поочередного изменения переменных. Идея последнего заключается в следующем: из начальной точки делается шаг по первой переменной, если он «удачный», то переходят к следующей переменной. Если шаг оказался «неудачным», то делается шаг в противоположном направлении. Эта процедура повторяется до тех пор, пока во всех направлениях не будут получаться одни «неудачные шаги». В этом случае величина шага уменьшается. Поиск продолжается до тех пор, пока абсолютное значение величины шага оказывается меньше заданной точности.

В методе Гаусса- Зейделя при выполнении шага по каждой переменной ищут минимум целевой функции в ее направлении, при этом значения остальных переменных остаются постоянными. Этот поиск по направлению можно производить любым известным методом одномерной оптимизации (например, методом обратного переменного шага, методом «золотого сечения»

и т.п.). Таким образом, в методе Гаусса-Зейделя задача многомерной оптимизации сводится к многократному использованию метода одномерной оптимизации. Очередность варьирования переменных при этом устанавливается произвольно и обычно не меняется в процессе оптимизации.
Таким образом, алгоритм метода заключается в следующем.

1. Для некоторого начального значения $\mathbf{x}^{(0)}$ фиксируют все координаты вектора $\mathbf{x}$, кроме одной (для определенности x_1) и проводят операцию одномерного поиска минимума функции $F(x_1)=f(x_1, x_2^{(0)}, x_3^{(0)}, \dots x_n^{(0)})$, в результате чего получают точку $\mathbf{x}^{(1)}=(x_1^*, x_2^{(0)}, x_3^{(0)}, \dots x_n^{(0)})$, где $x_1^* = \arg\min_{x_1} F(x_1)$.
2. Фиксируя в точке $\mathbf{x}^{(1)}$ все координаты кроме второй, повторяют п. 1 по x_2. И так до последней составляющей x_n. Цикл алгоритма завершается после n – кратной операции одномерной оптимизации вдоль каждой из координат, после чего этот цикл повторяют, получая точки $\mathbf{x}^{(1)}$, $\mathbf{x}^{(2)}$, …, в каждой из которых значение целевой функции не больше, чем в предыдущей.

Условием прекращения вычислительной процедуры при достижении заданной точности ε может служить неравенство $\left\| \mathbf{x}^{(k)} - \mathbf{x}^{(k-1)} \right\| < \varepsilon$. При пошаговом движении, например, в алгоритме поочередного изменения переменных поиск прекращается в точке, для которой $\mathbf{x}^{(k)}$ совпало с $\mathbf{x}^{(k-1)}$, т.е. цикл оказался нерезультативным.

Недостатком метода Гаусса-Зейделя является жесткое направление изменения каждой из составляющих решения, не зависящее от характера функции, что может привести к неоправданной остановке алгоритма в случае “овражных” функций.

4.1.2 Метод Хука и Дживса (метод конфигураций)

Этот метод можно рассматривать как модификацию метода Гаусса-Зейделя. Идея метода заключается в следующем: из начальной (базовой) точки выполняется одна итерация метода Гаусса-Зейделя; там, где получено уточненное значение функции, помещается временная базовая точка. После этого дальнейший поиск проводят вдоль прямой, соединяющий две базовые точки. Этот поиск проводится любым методом одномерного поиска. Найдя точку с минимальным значением целевой функции, из нее снова выполняют одну итерацию метода Гаусса-Зейделя, и дальнейший поиск снова проводят вдоль прямой, соединяющий две последние базовые точки т.д.

Рассмотрим простейшую модификацию метода Хука и Дживса. Процедура включает в себя два циклически повторяющихся этапа: исследующий поиск вокруг базисной точки и поиск по образцу.

1. Исследующий поиск. Задается начальная базисная точка $\mathbf{x}^{(0)}$ и приращения по каждой координате Δx_i . Рассчитывается значение целевой функции в базисной точке. Затем в циклическом порядке изменяется каждая координата:

$x_i^{(1)} = x_i^{(0)} + \Delta x_i$

Если приращение улучшает целевую функцию, то шаг считается "удачным" и дается приращение по другой координате. В противном случае- "неудачным" и делается шаг в противоположном направлении:

$x_i^{(1)} = x_i^{(0)} - \Delta x_i$

Если он также оказывается "неудачным", то значение $x_i^{(0)}$ оставляют неизменным, и дается приращение по следующей координате и т.д., пока не будут изменены все координаты. На этом заканчивается исследующий поиск, найдена точка $\mathbf{x}^{(1)}$.

Если "неудачными" оказались шаги по всем направлениям производится уменьшение приращений Δx_i (обычно в два раза) и исследующий поиск повторяется. Процедура заканчивается, когда величина приращений не станет меньше заданной точности.

2. Поиск по образцу осуществляется вдоль направления, соединяющего точки $\mathbf{x}^{(0)}$ и $\mathbf{x}^{(1)}$. Совершается один или несколько шагов до тех пор, пока шаги будут "удачными", т.е. приводят к улучшению целевой функции. Величина шагов обычно равна расстоянию между $\mathbf{x}^{(0)}$ и $\mathbf{x}^{(1)}$, т.е.

$$\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + (\mathbf{x}^{(1)}-\mathbf{x}^{(0)})= 2\cdot\mathbf{x}^{(1)}-\mathbf{x}^{(0)}.$$

Если после последнего "удачного" шага условие окончания поиска не выполнено, т.е. $\Delta x_i > \varepsilon$, то эту точку принимают в качестве новой базисной точки, и всю процедуру повторяют.

Недостатком метода Хука – Дживса, как и метода Гаусса- Зейделя, является его плохая сходимость при оптимизации “овражных” функций, что на стадии исследующего поиска может привести к неоправданной остановке алгоритма.

4.2 Симплексные алгоритмы

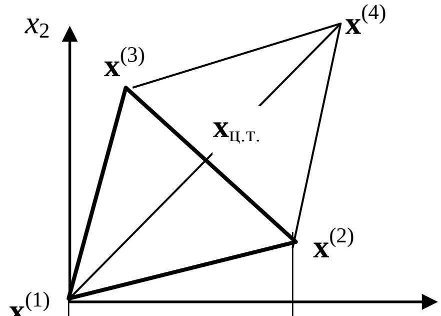
4.2.1 Обычный симплекс- метод

Симплексом в пространстве n переменных называют выпуклый многогранник, имеющий n+1 вершину. В пространстве двух переменных это треугольник, в пространстве трех переменных - тетраэдр. В обычном симплекс-методе используется правильный симплекс (все ребра которого равны).

Идея симплекс-метода, которая далее рассматривается на примере двумерного случая, заключается в следующем. Выбирается начальный симплекс с вершинами $\mathbf{x}^{(1)}$- $\mathbf{x}^{(2)}$- $\mathbf{x}^{(3)}$. Размещение правильного симплекса в пространстве может быть осуществлено двумя путями (рис.1).

1. Одна вершина симплекса помещается в начало координат, а остальные вершины располагаются так, чтобы ребра, выходящие из первой вершины, образовывали одинаковые углы с соответствующими координатными осями. Тогда для двумерного случая координаты вершин будут равны:

№ вершины	x_1	x_2
$\mathbf{x}^{(1)}$	0	0



$\mathbf{x}^{(2)}$	$\frac{\sqrt{3}+1}{2\sqrt{2}}$	$\frac{\sqrt{3}-1}{2\sqrt{2}}$
$\mathbf{x}^{(3)}$	$\frac{\sqrt{3}-1}{2\sqrt{2}}$	$\frac{\sqrt{3}+1}{2\sqrt{2}}$

Рис. 1

В общем случае координаты вершин симплекса определяются матрицей:

№ вершины	x_1	x_2	x_3	…	x_n
1	0	0	0	…	0
2	P	Q	Q	…	Q
3	Q	P	Q	…	Q
…	…	…	…	…	…
n+1	Q		Q	…	P

$$P = \frac{1}{n\sqrt{2}}(\sqrt{n+1} + n - 1);$$

$$Q = \frac{1}{n\sqrt{2}}\left(\sqrt{n+1} - 1\right).$$

2. Центр симплекса помещается в начало координат, а (n+1)-я вершина на ось x_n. Остальные вершины располагаются симметрично относительно координатных осей. В двумерном случае координаты вершин будут равны:

№ вершины	x_1	x_2
$\mathbf{x}^{(1)}$	$-\frac{1}{2}$	$-\frac{1}{2\sqrt{3}}$
$\mathbf{x}^{(2)}$	$\frac{1}{2}$	$-\frac{1}{2\sqrt{3}}$
$\mathbf{x}^{(3)}$	0	$\frac{1}{\sqrt{3}}$

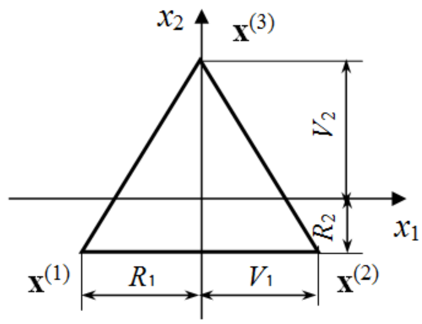


Рис. 2

В общем случае координаты вершин симплекса определяются матрицей:

№ вершины	x_1	x_2	x_3	…	x_n
1	$-R_1$	$-R_2$	$-R_3$	…	$-R_n$
2	V_1	$-R_2$	$-R_3$	…	$-R_n$
3	0	V_2	R_3	…	$-R_n$
4	0	0	V_3		$-R_n$
…	…	…	…	…	…
n+1	0	0	0	…	V_n

$$R_i = \frac{1}{\sqrt{2 \cdot i(i+1)}};$$

$$V_i = \sqrt{\frac{i}{2(i+1)}}.$$

В первом и во втором случаях формулы получены для симплекса, длина ребра которого равна единице. Для произвольной длины каждую формулу нужно умножить на длину ребра. Если поиск осуществляется не из начала координат, а из начальной точки $\mathbf{x}^{(0)}$, то к координатам вершин симплекса необходимо добавить координаты начальной точки- $x_1^{(0)}$ и $x_2^{(0)}$.

В вершинах исходного симплекса рассчитывается значение целевой функции $f(\mathbf{x}^{(1)})$, $f(\mathbf{x}^{(2)})$, $f(\mathbf{x}^{(3)})$. Из этих трех значений выбирается "наихудшая" точка (при поиске минимума это та точка, в которой функция принимает максимальное значение). Допустим, что это точка $\mathbf{x}^{(1)}$. Через центр тяжести противолежащей грани $\mathbf{x}_{ц.т.}=(\mathbf{x}^{(2)} + \mathbf{x}^{(3)})/2$ строится новая вершина симплекса $\mathbf{x}^{(4)}$, симметричная "наихудшей" вершине $\mathbf{x}^{(1)}$ (рис.1). Координаты новой вершины $\mathbf{x}^{(4)}$ рассчитываются по формуле:

$$\mathbf{x}^{(4)}= \mathbf{x}_{ц.т.}+ (\mathbf{x}_{ц.т.}- \mathbf{x}^{(1)})= \mathbf{x}^{(2)} + \mathbf{x}^{(3)} - \mathbf{x}^{(1)}$$

В результате получается новый симплекс $\mathbf{x}^{(2)}$ - $\mathbf{x}^{(3)}$ - $\mathbf{x}^{(4)}$, причем значение целевой функции в двух точках $\mathbf{x}^{(2)}$ и $\mathbf{x}^{(3)}$ уже известно. Поэтому вычисляется значение функции в точке $\mathbf{x}^{(4)}$ и среди всех вершин ищется вершина с "наихудшим" значением. Эта вершина вновь отображается через середину противолежащей грани и вся процедура повторяется. Признаком окончания поиска является так называемая процедура зацикливания, когда вновь отображенная вершина оказывается "наихудшей". В этом случае, если заданная точность не достигается, (точность определяется длиной ребра симплекса) необходимо уменьшить размеры симплекса. Процедура повторяется до тех пор, пока длина ребра симплекса не станет меньше заданной точности.

В общем n – мерном случае, если обозначить отображаемую вершину симплекса за $\mathbf{x}^{(1)}$, остальные вершины - $\mathbf{x}^{(2)}$, $\mathbf{x}^{(3)}$, $\mathbf{x}^{(n+1)}$, отображенную вершину за $\mathbf{x}^{(n+2)}$, координаты центра тяжести грани, относительно которой производится отображение, определяются по формуле:

$$\mathbf{x}_{ц.т.} = \frac{1}{n}\sum_{i=2}^{n+1}\mathbf{x}^{(i)} \quad , \quad \text{а отображаемой вершины}$$

$$\mathbf{x}^{(n+2)} = \mathbf{x}_{ц.т.} + \alpha\left(\mathbf{x}_{ц.т.} - \mathbf{x}^{(1)}\right). \qquad (4.1)$$

Здесь множитель $\alpha > 0$. При $\alpha =1$ получаем зеркальное отображение $\mathbf{x}^{(1)}$ и если исходный симплекс был правильным, то при зеркальном отображении и новый симплекс окажется правильным.

4.2.2 Метод Нельдера – Мида (деформируемых многогранников)

Данный метод является существенно более эффективным, чем простейший алгоритм регулярных симплексов, за счет того, что симплекс меняет свою форму от цикла к циклу. Рабочий цикл алгоритма состоит из следующих операций.

1. Выбирают начальный (обычно регулярный) симплекс $\mathbf{x}^{(1)}$ -$\mathbf{x}^{(2)}$-…- $\mathbf{x}^{(n+1)}$ и, как и в предыдущем алгоритме, рассчитывают значения целевой функции в его вершинах.
2. Из найденных значений целевой функции ищут максимальное $\mathbf{y}_{max}=f(\mathbf{x}_{max})$ и минимальное значение $\mathbf{y}_{min}=f(\mathbf{x}_{min})$.
3. Отображают “наихудшую” вершину относительно центра тяжести $\mathbf{x}_{ц.т}$ противоположной грани с коэффициентом $\alpha =1$ в формулах (4.1).
4. Анализируют результат отображения, сравнивая значение функции в отображенной вершине с ее значениями в вершинах предыдущего симплекса.

a) Если при этом значение функции в новой вершине оказалось меньше, чем наилучшее значение предыдущего симплекса, т.е. $\mathbf{y}^{(n+2)} < \mathbf{y}_{\min}$, то проводят ***растяжение*** симплекса, увеличивая в β>1 (обычно β =2) раз расстояние от отраженной вершины до центра тяжести $\mathbf{x}_{ц.т}$ - $$\mathbf{x}^{(n+3)} = \beta\mathbf{x}^{(n+2)} + (1-\beta)\mathbf{x}_{ц.т.}. \quad (4.2)$$

   Если после операции растяжения значение функции в новой точке уменьшается, т.е. $\mathbf{y}^{(n+3)} < \mathbf{y}^{(n+2)}$, то эта вершина принимается за вершину нового симплекса, если же увеличивается, то в качестве новой вершины берется точка, полученная после отображения $\mathbf{y}^{(n+2)}$.

b) Когда значение функции в отображенной вершине меньше, чем в наихудшей вершине предыдущего симплекса, но больше, чем во всех остальных, то производят операцию ***сжатия*** с коэффициентом β < 1 (обычно β = 0,5) в формуле (4.2).

c) Если значение функции в отображенной точке больше, чем в наихудшей вершине предыдущего симплекса, т.е. $\mathbf{y}^{(n+2)} > \mathbf{y}_{\max}$, то производят ***редукцию*** (уменьшение размеров симплекса обычно в два раза), т.е. координаты всех вершин симплекса сдвигаются на половину расстояния до наилучшей точки $\mathbf{x}^{(k)} = \mathbf{x}_{\min} + 0{,}5(\mathbf{x}^{(k)} - \mathbf{x}_{\min})$.

В качестве критерия остановки авторы алгоритма рекомендуют среднеквадратичную величину разности значений функции в вершинах симплекса и среднего ее значения, т.е. $\sqrt{\frac{1}{n+1}\sum_{i=1}^{n+1}\left[f\left(\mathbf{x}^{(k)}\right) - f\left(\mathbf{x}_{ц.т.}\right)\right]^2} \le \varepsilon$,

где ε - заданная точность.

4.3 Градиентные методы

Суть всех градиентных методов заключается в использовании вектора градиента для определения направления движения к оптимуму. Вектор градиента $\nabla f(\mathbf{x})$ = $\left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right)$ обладает несколькими свойствами, которые и обуславливают его эффективное применение при поиске экстремальных значений функции многих переменных. Выделим некоторые из них.

- Вектор градиента всегда направлен в сторону наиболее быстрого возрастания функции в данной точке. Поэтому очевидно, что при поиске минимальных значений функции необходимо двигаться в противоположную сторону. Такое направление движения называют антиградиентом
  ($-\nabla f(\mathbf{x})$) или отрицательным градиентом и оно характеризует направление наиболее быстрого убывания функции.

- Градиент всегда ортогонален линии равного уровня, проходящей через данную точку $\mathbf{x}^{(k)}$.
- Согласно необходимому условию существования экстремума функции многих переменных в точке экстремума градиент функции обращается в ноль. Это свойство часто используется для проверки условия окончания поиска в градиентных методах, т.е. $\left\| \nabla f(\mathbf{x}^{(k)}) \right\| \le \varepsilon$.

Общий алгоритм всех градиентных методов заключается в построении из некоторой начальной точки $\mathbf{x}^{(0)}$ последовательности приближений (рассматриваем задачу минимизации):

$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \lambda^{(k)} S^{(k)}$ $\qquad (k=0, 1, \ldots)$,

где $S^{(k)}$- единичный вектор в направлении градиента целевой функции $f(\mathbf{x})$ в точке $\mathbf{x}^{(k)}$:

$$S^{(k)} = \frac{\nabla f(\mathbf{x}^{(k)})}{\left\| \nabla f(\mathbf{x}^{(k)}) \right\|} = \frac{\nabla f(\mathbf{x}^{(k)})}{\sqrt{\left(\frac{\partial f}{\partial x_1}\right)^2 + \left(\frac{\partial f}{\partial x_2}\right)^2 + \cdots + \left(\frac{\partial f}{\partial x_n}\right)^2}};$$

$\lambda^{(k)}$ – величина шага в направлении градиента.

4.3.1 Метод крутого восхождения Бокса – Уилсона

Метод крутого восхождения Бокса – Уилсона представляет собой пошаговую процедуру движения по поверхности отклика, в которой для оценки составляющих градиента $\nabla f(\mathbf{x}) = [b_0^{(k)}, b_1^{(k)}, b_2^{(k)}, \ldots, b_n^{(k)}]$ используется линейное уравнение регрессии $f(\mathbf{x}) = b_0^{(k)} + b_1^{(k)} x_1 + b_2^{(k)} x_2 + \ldots + b_n^{(k)} x_n$, полученное в результате планирования эксперимента в окрестности точки $\mathbf{x}^{(k)}$.

Затем совершается движение по поверхности отклика в направлении градиента с величиной шага, пропорциональной произведению коэффициента $b_j^{(k)}$ на интервал варьирования Δx_j. Движение по поверхности осуществляется до тех пор, пока параметр оптимизации не начнет увеличиваться (в случае поиска минимума). В полученной точке вновь производится планирование эксперимента и оценка нового направления движения. Процедура поиска продолжается до тех пор, пока величина вектора градиента не станет меньше заданной точности.

Пусть мы имеем функцию $f(x_1, x_2)$. Необходимо найти минимум функции методом крутого восхождения.

Алгоритм поиска следующий.

Выбираем начальную точку поиска $\mathbf{x}^{(0)}$ $(x_1^{(0)}, x_2^{(0)})$. В окрестности точки проводим полный факторный эксперимент ПФЭ 2^2. Матрица планирования в кодированных переменных для ПФЭ 2^2 имеет вид:

№	X_1	X_2	y
1	-1	-1	y_1
2	+1	-1	y_2
3	-1	+1	y_3
4	+1	+1	y_4

Переход от кодированных пере-менных к натуральным осуществляется по формуле: $\quad x_j = x_j^{(0)} + X_j \cdot \Delta x_j, \qquad j = 1, 2,$

где $\quad x_j^{(0)}$- опорный уровень;

X_j - кодированное значение переменной;

Δx_j - интервал варьирования.

Опорному уровню соответствуют координаты начальной точки поиска, кодированные значения переменных приведены в матрице ПФЭ 2^2. Интервал варьирования Δx_j для расчетов можно принять равным 1. С учетом этого матрица ПФЭ 2^2 в натуральных переменных будет иметь вид:

№	x_1	x_2	y
1	$x_1^{(0)}-1$	$x_1^{(0)}-1$	y_1
2	$x_1^{(0)}+1$	$x_1^{(0)}-1$	y_2
3	$x_1^{(0)}-1$	$x_1^{(0)}+1$	y_3
4	$x_1^{(0)}+1$	$x_1^{(0)}+1$	y_4

Значения y_1, y_2, y_3, y_4 определяем, рассчитывая значение функции $y=f(x_1, x_2)$ для полученных четырех точек. Метод ПФЭ позволяет получить линейную зависимость y от x_1 и x_2:

$$y= b_0+ b_1x_1+ b_2x_2 \; .$$

Необходимо оценить значения коэффициентов уравнения b_1, b_2

$$b_1= (-y_1+y_2-y_3+y_4)/4 \; ; \qquad b_2= (-y_1-y_2+y_3+y_4)/4 \; .$$

Определив значение коэффициентов регрессии, начинаем движение по поверхности отклика в направлении вектора градиента $[-b_1,-b_2]$. Знак (-) указывает на поиск минимума. Расчет ведем по формуле:

$$x_j^{(k)}= x_j^{(0)} - k\cdot a\cdot b_j\cdot\Delta x_j \, ,$$

где: $x_j^{(k)}$ - значение j-го фактора на k-том шаге движения в направлении градиента;

a - коэффициент пропорциональности (выбираем с учетом предполагаемого вида поверхности отклика).

Движение по поверхности отклика осуществляется до тех пор, пока значение функции не начнет увеличиваться. В этом случае за исходную точку берется точка, в которой значение функции будет минимальным. В этой точке вновь реализуется ПФЭ 2^2, рассчитываются коэффициенты регрессии и вновь осуществляется движение по поверхности отклика. Процедура повторяется до тех пор, пока длина вектора градиента $\|\nabla f(\mathbf{x})\| = \sqrt{b_1^2 + b_2^2}$ не станет меньше заданного ε.

5. ПРИМЕНЕНИЕ МЕТОДОВ ОПТИМИЗАЦИИ К РЕШЕНИЮ НЕЛИНЕЙНЫХ УРАВНЕНИЙ И СИСТЕМ

Пусть требуется найти решение системы

$$\begin{cases} f_1(x_1, x_2, \; \ldots \; , x_n) = 0 \\ f_2(x_1, x_2, \; \ldots \; , x_n) = 0 \\ \quad \ldots\ldots\ldots\ldots\ldots \\ f_n(x_1, x_2, \; \ldots \; , x_n) = 0 \end{cases}$$

которая при $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix}$, $\mathbf{f}(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \cdots \\ f_n(\mathbf{x}) \end{pmatrix}$

записывается в векторной форме как $\mathbf{f}(\mathbf{x})=0$.

Введем неотрицательную функцию

$$\Phi(\mathbf{x}) = \sum_{i=1}^{n} \left(f_i(x_1, x_1, \ldots, x_n) \right)^2 .$$

Тогда точка минимума $\mathbf{x}^*$ функции $\Phi(\mathbf{x})$ является решением данной системы уравнений, и наоборот, решение системы реализует минимум функции $\Phi(\mathbf{x})$. Действительно, пусть $\mathbf{x}^*$- решение системы уравнений. Тогда $\Phi(\mathbf{x}^*)=0$, т.е. $\mathbf{x}^*$ является точкой локального минимума функции $\Phi(\mathbf{x})$ $(\Phi(\mathbf{x}) \geq 0 \ \ \forall \mathbf{x} \in \mathbf{R}^n)$. Таким образом, можно получить решение системы нелинейных уравнений, решая задачу об определении локального минимума функции $\Phi(\mathbf{x})$, т.е. задачу $\Phi(\mathbf{x}) \to \min$, $\mathbf{x} \in \mathbf{R}^n$.

Пусть дана система уравнений

$$\begin{cases} x_1^2 + x_2^2 - 1 = 0 \\ x_2 - x_1 \sin x_1 = 0 \end{cases} .$$

Решение этой системы можно найти, отыскивая точку локального минимума целевой функции:

$\Phi(\mathbf{x}) = f_1^2(x_1, x_2) + f_2^2(x_1, x_2) = (x_1^2 + x_2^2 - 1)^2 + (x_2 - x_1 \sin x_1)^2$.

ЗАДАНИЕ НА КОНТРОЛЬНУЮ РАБОТУ 4

1. Найти решение задачи безусловной оптимизации для заданной целевой функции (прил. 2, 3), используя теоремы о необходимых и достаточных условиях экстремума. Провести анализ найденного решения и установить, на каком множестве **D** оно является глобальным.
2. Провести графический анализ функции, отобразив ее в виде совокупности линий уровня.

ПРИМЕР ВЫПОЛНЕНИЯ КОНТРОЛЬНОЙ РАБОТЫ 4

Найти минимум функции $f(x_1,x_2) = (x_1 + x_2)^2 + (x_2 - 1)^2$
с точностью $\varepsilon = 0{,}2$ для начальной точки с координатами $\mathbf{x}^{(0)}= (5, 6)$.

Аналитический анализ функции

Находим частные производные первого порядка и приравниваем их нулю.

$$\frac{\partial f}{\partial x_1} = 2(x_1 + x_2) = 0 \quad ;$$

$$\frac{\partial f}{\partial x_1} = 2(x_1 + x_2) + 2(x_2 - 1) = 0 \ .$$

Решаем систему уравнений:

$$\begin{cases} x_1 + x_2 = 0 \\ 2x_1 + 4x_2 - 2 = 0 \end{cases} ; \quad \begin{cases} x_1 = -x_2 \\ -2x_2 + 4x_2 - 2 = 0 \end{cases}$$

и получаем $x_1^* = -1$; $x_2^* = 1$.

Находим частные производные второго порядка:

$$\frac{\partial^2 f}{\partial x_1^2} = 2 = A$$

$$\frac{\partial^2 f}{\partial x_2^2} = 4 = B \qquad \text{и матрицу Гессе } \mathrm{H} = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} A & C \\ C & B \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 4 \end{bmatrix}.$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = 2 = C$$

Поскольку матрица Гессе ($A \cdot B - C^2 = 8 - 4 = 4 > 0$ и $A = 2 > 0$), точка $\mathbf{x}^* = (-1, 1)$ является точкой минимума. Значение функции в этой точке $f(-1; 1) = (-1+1)^2 + (1-1)^2 = 0$.

Матрица Гессе положительно определена независимо от координат точки **x** и, следовательно, рассматриваемая функция является выпуклой на множестве $\mathbf{R}^2$, а единственная стационарная точка $\mathbf{x}^*$ - глобальным минимумом $f(x_1,x_2)$.

Графический анализ функции

Для построения линий уровня функции $F = f(x_1, x_2) = (x_1+x_2)^2+(x_2-1)^2$ выбираем следующие значения функции: 25, 50, 100, 200.

Выражаем переменную x_1 через x_2 и функцию F

$$x_1 = \pm\sqrt{F-(x_2-1)^2} - x_2 \ .$$

Результаты расчетов сводим в таблицу

$f(x_1, x_2)$	x_2	$+x_1$	$- x_1$
25	1	4	-6
	-1	5,58	-3,58
	3	1,58	-7,58
	-3	6,00	0,00
	5	-2,00	-8,00
50	1	6,07	-8,07
	-1	7,78	-5,78
	3	3,78	-9,78
	-3	8,83	-2,88
	5	0,88	-10,8
	-5	8,74	2,26
	6	3,26	9,74
100	0	9,95	-9,95
	2	7,95	-11,95
	-2	11,54	-7,54
	4	5,54	-13,54
	-4	12,66	-4,66
	6	2,66	-14,66
	-6	13,14	-1,14
	8	-1,86	-15,14
	-8	12,36	4,64
	10	-6,64	-14,36

$f(x_1, x_2)$	x_2	$+x_1$	$- x_1$
200	0	14,11	-14,11
	2	12,11	16,11
	-2	15,82	-11,82
	4	9,82	-17,82
	-4	17,23	-9,23
	6	7,23	-19,23
	-6	18,29	-6,29
	8	4,29	-20,89
	-8	18,91	-2,91
	10	0,91	-20,91
	-10	18,89	2,11
	12	-4,11	-20,89
	-12	17,56	7,44
	14	7,00	-11,56

Полученные линии уровня приведены на рис.3

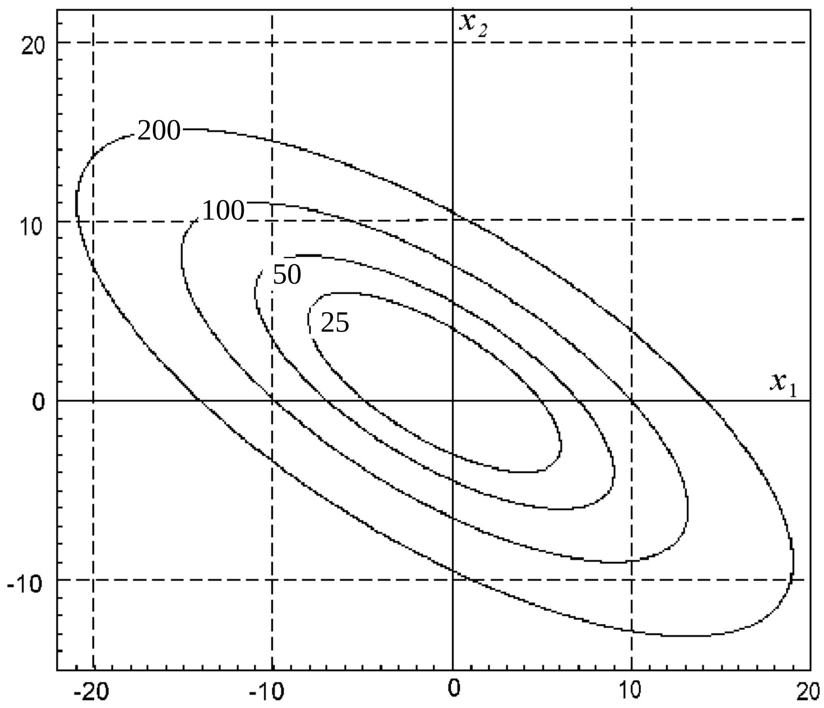


Рис.3

ЗАДАНИЕ НА КОНТРОЛЬНУЮ РАБОТУ 5

1. Найти приближенное решение задачи безусловной минимизации $f(\mathbf{x}) \to \min$, $\mathbf{x} \subset \mathbf{D}$ для заданной начальной точки $\mathbf{x}^{(0)}$ с заданной точностью ε:
   - методом Гаусса-Зейделя;
   - симплекс-методом;

ПРИМЕР ВЫПОЛНЕНИЯ КОНТРОЛЬНОЙ РАБОТЫ 5

Поиск экстремума методом Гаусса - Зейделя

Определяем значение функции в начальной точке $\mathbf{x}^{(0)}= (5, 6)$

$$f(\mathbf{x}^{(0)})= (5+6)^2+(6-1)^2= 146.$$

Выбираем шаг по каждой координате: $\Delta x_1=2$; $\Delta x_2=2$.

1.1. Осуществляем одномерный поиск по координате x_1 при $x_2=6$.

$x_1^{(0)1} = 5+2 = 7$;

$f(\mathbf{x}^{(0)1}) = (7+6)^2+(6-1)^2= 194$ - шаг неудачный и движемся в обратном направлении;

$x_1^{(0)1} = 5-2 = 3$;

$f(\mathbf{x}^{(0)1}) = (3+6)^2+(6-1)^2= 106$ - шаг удачный и продолжаем движение в том же направлении ;

$\mathbf{x}^{(0)2} = (1; 6)$; $f(\mathbf{x}^{(0)2}) = (1+6)^2+(6-1)^2= 106$;

$\mathbf{x}^{(0)3} = (-1; 6)$; $f(\mathbf{x}^{(0)3}) = (-1+6)^2+(6-1)^2= 50$;

$\mathbf{x}^{(0)4} = (-3; 6)$; $f(\mathbf{x}^{(0)4}) = (-3+6)^2+(6-1)^2= 34$;

$\mathbf{x}^{(0)5} = (-5; 6)$; $f(\mathbf{x}^{(0)5}) = (-5+6)^2+(6-1)^2= 26$;

$\mathbf{x}^{(0)6} = (-7; 6)$; $f(\mathbf{x}^{(0)6}) = (-7+6)^2+(6-1)^2= 26$;

$\mathbf{x}^{(0)7} = (-9; 6)$; $f(\mathbf{x}^{(0)7}) = (-9+6)^2+(6-1)^2= 34$ – шаг неудачный, возвращаемся в точку $\mathbf{x}^{(0)6}$ и осуществляем одномерный поиск по координате x_2.

$x_2^{(0)8} = 6+2 = 8$;

$f(\mathbf{x}^{(0)8}) = (-7+8)^2+ (8-1)^2= 50$ - шаг неудачный, меняем шаг на обратный;

$x_2^{(0)9} = 6-2 = 4$; $f(\mathbf{x}^{(0)9}) = (-7+4)^2+(4-1)^2=18$ - шаг удачный.

Получили точку $\mathbf{x}^{(1)}= \mathbf{x}^{(0)9}$ с координатами (-7; 4) и переходим к следующей итерации.

1.2. Осуществляем одномерный поиск по координате x_1 .

$\mathbf{x}^{(1)1} = (-5; 4)$; $f(\mathbf{x}^{(1)1}) = (-5+4)^2+(4-1)^2= 10$;

$\mathbf{x}^{(1)2} = (-3; 4)$; $f(\mathbf{x}^{(1)2}) = (-3+4)^2+(4-1)^2= 10$;

$\mathbf{x}^{(1)3} = (-1; 4)$; $f(\mathbf{x}^{(1)3}) = (-1+4)^2+(4-1)^2= 18$ - шаг неудачный, возвращаемся в точку $\mathbf{x}^{(1)2}$ и осуществляем одномерный поиск по координате x_2.

$\mathbf{x}^{(1)4} = (-3; 6)$; $f(\mathbf{x}^{(1)4}) = (-3+6)^2+(6-1)^2= 34$ –неудача;

$\mathbf{x}^{(1)5} = (-3; 2)$; $f(\mathbf{x}^{(1)5}) = (-3+2)^2+(2-1)^2= 2$ –удача;

$\mathbf{x}^{(1)6} = (-3; 0)$; $f(\mathbf{x}^{(1)6}) = (-3+0)^2+(0-1)^2= 10$ –неудача и переходим к следующей итерации.

1.3. Базовая точка $\mathbf{x}^{(2)}$ = (-3, 2). Осуществляем одномерный поиск по координате x_1 .
$\mathbf{x}^{(2)1}$ = (-1; 2); $f(\mathbf{x}^{(2)1})$ = 2 – удача; продолжаем по x_2;
$\mathbf{x}^{(2)2}$ = (-1; 4) ; $f(\mathbf{x}^{(2)2})$ = 18 – неудача;
$\mathbf{x}^{(2)3}$ = (-1; 0) ; $f(\mathbf{x}^{(2)3})$ = 2 – удача, переходим к следующей итерации.

1.4. Базовая точка $\mathbf{x}^{(3)}$ = (-1, 0).
$\mathbf{x}^{(3)1}$ = (-3; 0) ; $f(\mathbf{x}^{(3)1})$ = 10 – неудача;
$\mathbf{x}^{(3)2}$ = (1; 0); $f(\mathbf{x}^{(3)2})$ = 2 – удача, продолжаем по x_2;
$\mathbf{x}^{(3)3}$ = (1; 2); $f(\mathbf{x}^{(3)3})$ = 10 – неудача;
$\mathbf{x}^{(3)4}$ = (1;-2); $f(\mathbf{x}^{(3)4})$ = 10 – неудача, переходим к следующей итерации.

1.5. Базовая точка $\mathbf{x}^{(4)}$ = (1, 0).
$\mathbf{x}^{(4)1}$ = (-3; 0); $f(\mathbf{x}^{(4)1})$ = 10 – неудача;
$\mathbf{x}^{(4)2}$ = (3; 0); $f(\mathbf{x}^{(4)2})$ = 10 – неудача, продолжаем по x_2;
$\mathbf{x}^{(4)3}$ = (1; 2); $f(\mathbf{x}^{(4)3})$ = 10 – неудача;
$\mathbf{x}^{(4)4}$ = (1;-2); $f(\mathbf{x}^{(4)4})$ = 10 – неудача.

Поскольку в данной точке одномерный поиск не приводит к успеху ни по одной координате проверяем условие остановки алгоритма Δx_i=2> ε=0,2, уменьшаем шаг по каждой координате в два раза:

Δx_1=1 ; Δx_2=1 и переходим к следующей итерации.

1.6. Базовая точка $\mathbf{x}^{(5)}$ = (1, 0).
$\mathbf{x}^{(5)1}$ = (2; 0); $f(\mathbf{x}^{(5)1})$ = 5 – неудача;
$\mathbf{x}^{(5)2}$ = (0; 0); $f(\mathbf{x}^{(5)2})$ = 1 – удача, продолжаем по x_2;
$\mathbf{x}^{(5)3}$ = (0; 1); $f(\mathbf{x}^{(5)3})$ = 1 – удача; переходим к следующей итерации.

1.7. Базовая точка $\mathbf{x}^{(6)}$ = (0, 1).
$\mathbf{x}^{(6)1}$ = (1; 1); $f(\mathbf{x}^{(6)1})$ = 4 – неудача;
$\mathbf{x}^{(6)2}$ = (-1; 1); $f(\mathbf{x}^{(6)2})$ = 0 – удача, продолжаем по x_2;
$\mathbf{x}^{(6)3}$ = (-1; 2) ; $f(\mathbf{x}^{(6)3})$ = 2 – неудача;
$\mathbf{x}^{(6)4}$ = (-1; 0) ; $f(\mathbf{x}^{(6)4})$ = 2– неудача; переходим к следующей итерации.

1.8. Базовая точка $\mathbf{x}^{(7)}$ = (-1, 1).

Все последующие шаги из данной точки неудачны, поэтому сокращаем шаг в два раза до 0,5 и поскольку дальнейшие шаги также неудачны (произошло случайное попадание в точку экстремума) сокращаем шаг еще в два раза до 0,125. Последующие шаги также не улучшают целевую функцию и, поскольку условие остановки алгоритма Δx_i=0,125 < ε=0,2 выполняется, прекращаем вычисления.

Таким образом, за точку минимума принимаем значение $\mathbf{x}^* \approx \mathbf{x}^{(7)} = (-1; 1)$.

Траектория поиска показана на рис. 4.

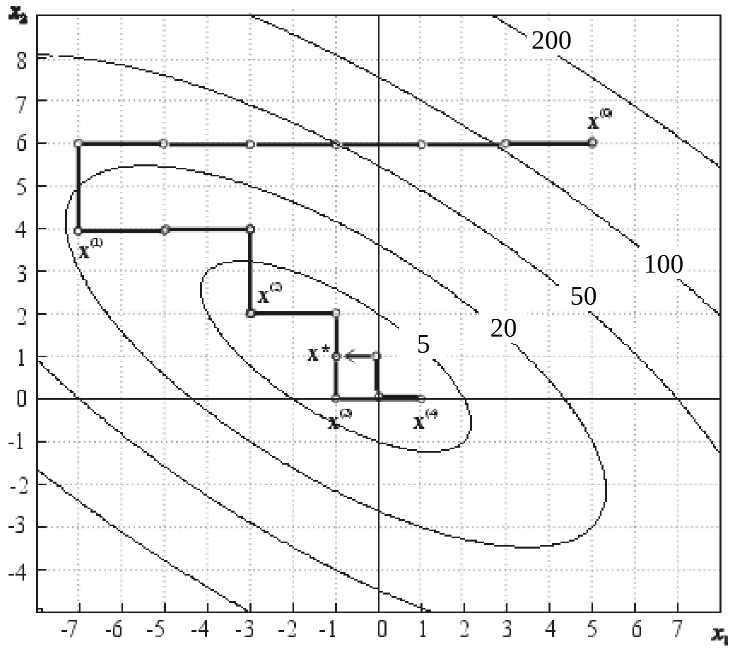


Рис.4

Поиск экстремума методом Гаусса - Зейделя

Определяем значение функции в начальной точке $\mathbf{x}^{(0)} = (5, 6)$

$$f(\mathbf{x}^{(0)}) = (5+6)^2+(6-1)^2 = 146.$$

Выбираем шаг по каждой координате: $\Delta x_1 = 2$; $\Delta x_2 = 2$.

1.9. Осуществляем одномерный поиск по координате x_1 при $x_2 = 6$.

$x_1^{(0)1} = 5+2 = 7$;

$f(\mathbf{x}^{(0)1}) = (7+6)^2+(6-1)^2 = 194$ - шаг неудачный и движемся в обратном направлении;

$x_1^{(0)1} = 5-2 = 3$;

$f(\mathbf{x}^{(0)1}) = (3+6)^2+(6-1)^2 = 106$ - шаг удачный и продолжаем движение в том же направлении ;

$\mathbf{x}^{(0)2} = (1; 6)$; $f(\mathbf{x}^{(0)2}) = (1+6)^2+(6-1)^2 = 106$;

$\mathbf{x}^{(0)3} = (-1; 6)$; $f(\mathbf{x}^{(0)3}) = (-1+6)^2+(6-1)^2 = 50$;

$\mathbf{x}^{(0)4} = (-3; 6)$; $f(\mathbf{x}^{(0)4}) = (-3+6)^2+(6-1)^2 = 34$;

$\mathbf{x}^{(0)5} = (-5; 6)$; $f(\mathbf{x}^{(0)5}) = (-5+6)^2+(6-1)^2 = 26$;

$\mathbf{x}^{(0)6} = (-7; 6)$; $f(\mathbf{x}^{(0)6}) = (-7+6)^2+(6-1)^2 = 26$;

$\mathbf{x}^{(0)7} = (-9; 6)$; $f(\mathbf{x}^{(0)7}) = (-9+6)^2+(6-1)^2 = 34$ – шаг неудачный, возвращаемся в точку $\mathbf{x}^{(0)6}$ и осуществляем одномерный поиск по координате x_2.

$x_2^{(0)8} = 6+2 = 8$;

$f(\mathbf{x}^{(0)8}) = (-7+8)^2+ (8-1)^2 = 50$ - шаг неудачный, меняем шаг на обратный;

$x_2^{(0)9} = 6-2 = 4$; $f(\mathbf{x}^{(0)9}) = (-7+4)^2+(4-1)^2 = 18$ - шаг удачный.

Получили точку $\mathbf{x}^{(1)} = \mathbf{x}^{(0)9}$ с координатами (-7; 4) и переходим к следующей итерации.

1.10. Осуществляем одномерный поиск по координате x_1 .

$\mathbf{x}^{(1)1} = (-5; 4)$; $f(\mathbf{x}^{(1)1}) = (-5+4)^2+(4-1)^2 = 10$;

$\mathbf{x}^{(1)2} = (-3; 4)$; $f(\mathbf{x}^{(1)2}) = (-3+4)^2+(4-1)^2 = 10$;

$\mathbf{x}^{(1)3}$ = (-1; 4) ; $f(\mathbf{x}^{(1)3})$ = $(-1+4)^2+(4-1)^2$= 18 - шаг неудачный, возвращаемся в точку $\mathbf{x}^{(1)2}$ и осуществляем одномерный поиск по координате x_2.

$\mathbf{x}^{(1)4}$ = (-3; 6); $f(\mathbf{x}^{(1)4})$ = $(-3+6)^2+(6-1)^2$= 34 –неудача;
$\mathbf{x}^{(1)5}$ = (-3; 2); $f(\mathbf{x}^{(1)5})$ = $(-3+2)^2+(2-1)^2$= 2 –удача;
$\mathbf{x}^{(1)6}$ = (-3; 0); $f(\mathbf{x}^{(1)6})$ = $(-3+0)^2+(0-1)^2$= 10 –неудача и переходим к следующей итерации.

1.11. Базовая точка $\mathbf{x}^{(2)}$ = (-3, 2). Осуществляем одномерный поиск по координате x_1 .
$\mathbf{x}^{(2)1}$ = (-1; 2); $f(\mathbf{x}^{(2)1})$ = 2 – удача; продолжаем по x_2;
$\mathbf{x}^{(2)2}$ = (-1; 4) ; $f(\mathbf{x}^{(2)2})$ = 18 – неудача;
$\mathbf{x}^{(2)3}$ = (-1; 0) ; $f(\mathbf{x}^{(2)3})$ = 2 – удача, переходим к следующей итерации.

1.12. Базовая точка $\mathbf{x}^{(3)}$ = (-1, 0).
$\mathbf{x}^{(3)1}$ = (-3; 0) ; $f(\mathbf{x}^{(3)1})$ = 10 – неудача;
$\mathbf{x}^{(3)2}$ = (1; 0); $f(\mathbf{x}^{(3)2})$ = 2 – удача, продолжаем по x_2;
$\mathbf{x}^{(3)3}$ = (1; 2); $f(\mathbf{x}^{(3)3})$ = 10 – неудача;
$\mathbf{x}^{(3)4}$ = (1;-2); $f(\mathbf{x}^{(3)4})$ = 10 – неудача, переходим к следующей итерации.

1.13. Базовая точка $\mathbf{x}^{(4)}$ = (1, 0).
$\mathbf{x}^{(4)1}$ = (-3; 0); $f(\mathbf{x}^{(4)1})$ = 10 – неудача;
$\mathbf{x}^{(4)2}$ = (3; 0); $f(\mathbf{x}^{(4)2})$ = 10 – неудача, продолжаем по x_2;
$\mathbf{x}^{(4)3}$ = (1; 2); $f(\mathbf{x}^{(4)3})$ = 10 – неудача;
$\mathbf{x}^{(4)4}$ = (1;-2); $f(\mathbf{x}^{(4)4})$ = 10 – неудача.

Поскольку в данной точке одномерный поиск не приводит к успеху ни по одной координате проверяем условие остановки алгоритма Δx_i=2> ε=0,2, уменьшаем шаг по каждой координате в два раза:

Δx_1=1 ; Δx_2=1 и переходим к следующей итерации.

1.14. Базовая точка $\mathbf{x}^{(5)}$ = (1, 0).
$\mathbf{x}^{(5)1}$ = (2; 0); $f(\mathbf{x}^{(5)1})$ = 5 – неудача;
$\mathbf{x}^{(5)2}$ = (0; 0); $f(\mathbf{x}^{(5)2})$ = 1 – удача, продолжаем по x_2;
$\mathbf{x}^{(5)3}$ = (0; 1); $f(\mathbf{x}^{(5)3})$ = 1 – удача; переходим к следующей итерации.

1.15. Базовая точка $\mathbf{x}^{(6)}$ = (0, 1).
$\mathbf{x}^{(6)1}$ = (1; 1); $f(\mathbf{x}^{(6)1})$ = 4 – неудача;
$\mathbf{x}^{(6)2}$ = (-1; 1); $f(\mathbf{x}^{(6)2})$ = 0 – удача, продолжаем по x_2;
$\mathbf{x}^{(6)3}$ = (-1; 2) ; $f(\mathbf{x}^{(6)3})$ = 2 – неудача;
$\mathbf{x}^{(6)4}$ = (-1; 0) ; $f(\mathbf{x}^{(6)4})$ = 2– неудача; переходим к следующей итерации.

1.16. Базовая точка $\mathbf{x}^{(7)}$ = (-1, 1).

Все последующие шаги из данной точки неудачны, поэтому сокращаем шаг в два раза до 0,5 и поскольку дальнейшие шаги также неудачны (произошло случайное попадание в точку экстремума) сокращаем шаг еще в два раза до 0,125. Последующие шаги также не улучшают целевую функцию и, поскольку условие остановки алгоритма Δx_i=0,125 < ε=0,2 выполняется, прекращаем вычисления.

Таким образом, за точку минимума принимаем значение $\mathbf{x}^* \approx \mathbf{x}^{(7)}$ =(-1; 1).

Траектория поиска показана на рис. 4.

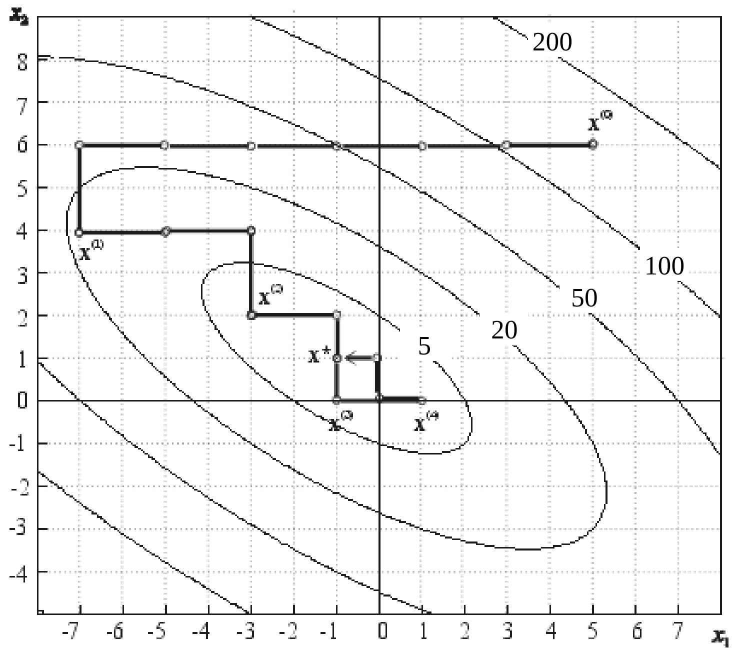


Рис.4

ЗАДАНИЕ НА КОНТРОЛЬНУЮ РАБОТУ 6

1. Найти приближенное решение задачи безусловной минимизации $f(\mathbf{x}) \to \min$, $\mathbf{x} \subset \mathbf{D}$ для заданной начальной точки $\mathbf{x}^{(0)}$ с заданной точностью ε:
   - методом Хука и Дживса;
   - методом Нельдера-Мида.

   Построить траектории поиска, совместив их в одних осях координат с линиями уровня.
2. Сделать выводы об эффективности методов, сравнивая количество расчетов функции для достижения заданной точности.

ПРИМЕР ВЫПОЛНЕНИЯ КОНТРОЛЬНОЙ РАБОТЫ 6

Поиск экстремума методом Хука и Дживса

Определяем значение функции в начальной точке

$$f(\mathbf{x}^{(0)})= (5+6)^2+(6-1)^2= 146.$$

Выбираем приращения:

$$\Delta x_1=2 ; \qquad \Delta x_2=2$$

1.1. Осуществляем исследующий поиск по координате x_1

$x_1^{(0)1} = 5+2 = 7$;

$f(\mathbf{x}^{(0)1}) = (7+6)^2+(6-1)^2= 194$ - шаг неудачный.

$x_1^{(0)2} = 5-2 = 3$;

$f(\mathbf{x}^{(0)2}) = (3+6)^2+(6-1)^2= 106$ - шаг удачный.

Осуществляем исследующий поиск по координате x_2.

$x_2^{(0)3} = 6+2 = 8$;

$f(\mathbf{x}^{(0)3}) = (3+8)^2+ (8-1)^2= 170$ - шаг неудачный.

$x_2^{(0)4} = 6-2 = 4$;

$f(\mathbf{x}^{(0)4}) = (3+4)^2+(4-1)^2=58$ - шаг удачный.

Получили точку $\mathbf{x}^{(1)}$ с координатами (3; 4).

Осуществляем пошаговый поиск по образцу:

Точка $\mathbf{x}^{(2)}$ имеет координаты (1; 2);$f(\mathbf{x}^{(2)})= (1+2)^2+(2-1)^2=10$.

Точка $\mathbf{x}^{(3)}$ имеет координаты (-1; 0); $f(\mathbf{x}^{(3)}) =(-1+0)^2+(0-1)^2=2$.

Точка $\mathbf{x}^{(4)}$ имеет координаты (-3,-2);

$f(\mathbf{x}^{(4)}) =(-3-2)^2+(-2-1)^2=34$ - шаг неудачный.

Переходим ко второй итерации.

1.2. Координаты базисной точки $\mathbf{x}^{(3)} = (-1; 0)$.

Осуществляем исследующий поиск по координате x_1:

$\mathbf{x}^{(3)1} = (1; 0)$; $f(\mathbf{x}^{(3)1}) = (1+0)^2+(0-1)^2=2$ - шаг удачный.

Осуществляем исследующий поиск по координате x_2.

$\mathbf{x}^{(3)2} = (1; 2)$; $f(\mathbf{x}^{(3)2}) = (1+2)^2+(2-1)^2=10$ - шаг неудачный.

$\mathbf{x}^{(3)3} = (1; -2)$; $f(\mathbf{x}^{(3)3}) = (1-(-2))^2+(-2-1)^2=10$ - шаг неудачный.

Получили точку $\mathbf{x}^{(4)}$ с координатами (1; 0).

Осуществляем пошаговый поиск по образцу:

Точка $\mathbf{x}^{(5)}$ имеет координаты (3; 0);

$f(\mathbf{x}^{(5)})= (3+0)^2+(0-1)^2=10$ - шаг неудачный.

Поскольку неудачными оказались шаги по всем направлениям, проверяем условие окончания алгоритма $\Delta x_i=2> \varepsilon=0,2$, уменьшаем приращения в два раза: $\Delta x_1=1$; $\Delta x_2=1$

и переходим ко второй итерации.

1.3. Координаты базисной точки $\mathbf{x}^{(4)} = (1; 0)$.

Осуществляем исследующий поиск по координате x_1:

$\mathbf{x}^{(4)1} = (2; 0)$; $f(\mathbf{x}^{(4)1}) = (2+0)^2+(0-1)^2=5$ - шаг неудачный.

$\mathbf{x}^{(4)2} = (0; 0)$; $f(\mathbf{x}^{(4)2}) = (0+0)^2+(0-1)^2=1$ - шаг удачный.

Осуществляем исследующий поиск по координате x_2.

$\mathbf{x}^{(4)3} = (0; 1)$; $f(\mathbf{x}^{(4)3}) = (0+1)^2+(1-1)^2=1$ - шаг удачный.

Получили точку $\mathbf{x}^{(5)}$ с координатами (0; 1); $f(\mathbf{x}^{(5)}) =1$.

Осуществляем пошаговый поиск по образцу.

Точка $\mathbf{x}^{(6)}$ имеет координаты (-1; 2); $f(\mathbf{x}^{(3)})= (-1+2)^2+(2-1)^2=2$.

Шаг неудачный, поэтому возвращаемся к точке $\mathbf{x}^{(5)}$ и повторяем исследующий поиск.

1.4. Координаты базисной точки $\mathbf{x}^{(5)} = (0; 1)$; $f(\mathbf{x}^{(5)}) =1$.

Осуществляем исследующий поиск по координате x_1:

$\mathbf{x}^{(5)1} = (1; 1)$; $f(\mathbf{x}^{(5)1}) = (1+1)^2+(1-1)^2=4$ - шаг неудачный.

$\mathbf{x}^{(5)2} = (-1; 1)$; $f(\mathbf{x}^{(5)2}) = (-1+1)^2+(1-1)^2=0$ - шаг удачный.

Осуществляем исследующий поиск по координате x_2.

$\mathbf{x}^{(5)3} = (-1; 2)$; $f(\mathbf{x}^{(5)3}) = (-1+2)^2+(2-1)^2=2$ - шаг неудачный.

$\mathbf{x}^{(5)4} = (-1; 0)$; $f(\mathbf{x}^{(5)4}) = (-1+0)^2+(0-1)^2=2$ - шаг неудачный.

Получили точку $\mathbf{x}^{(6)}$ с координатами (-1; 1); $f(\mathbf{x}^{(6)}) =0$.

Последующие неудачные шаги объясняются случайным попаданием в точку минимума.

Поскольку пошаговый поиск по образцу не приносит положительных результатов, а также неудачными оказываются шаги по всем направлениям, проверяем условие окончания алгоритма $\Delta x_i=1> \varepsilon=0,2$, уменьшаем приращения в два раза:

$\Delta x_1=0,5$; $\Delta x_2=0,5$

и переходим к следующей итерации итерации.

1.5. Координаты базисной точки $\mathbf{x}^{(6)} = (-1; 1)$; $f(\mathbf{x}^{(6)}) =0$.

Исследующий поиск в окрестности базисной точки не приносит улучшения функции, поэтому уменьшаем шаг до 0,25 и далее до 0,125, что также не приводит к положительным результатам. Проверяем условие окончания поиска Δx_i=0,125 < ε=0,2 и заканчиваем вычисления.

Таким образом, за точку минимума принимаем значение
x*≈ $\mathbf{x}^{(6)}$ = (-1; 1).

Траектория поиска приведена на рис.5.

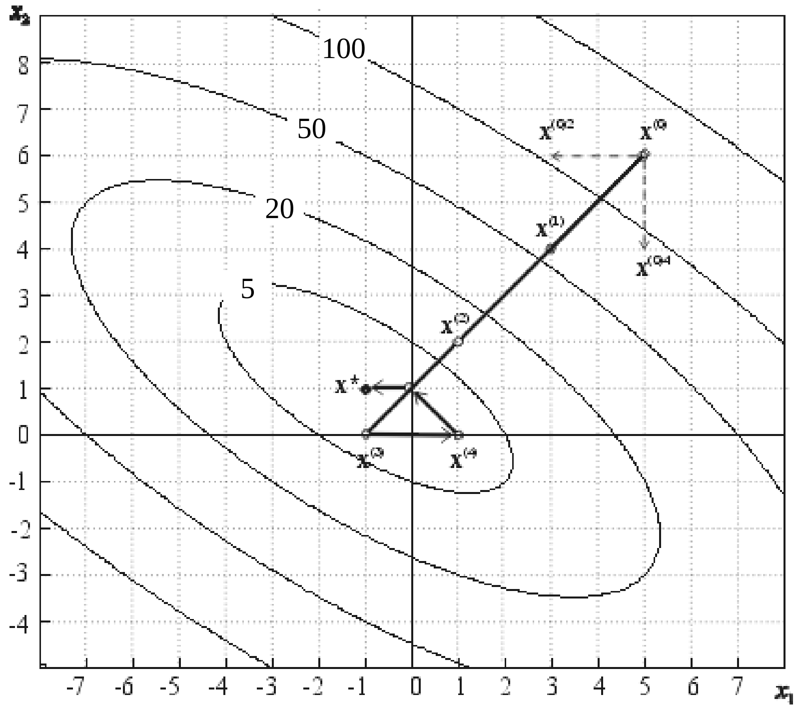


Рис. 5

Поиск экстремума методом Нельдера - Мида

В качестве начального возьмем такой же симплекс, что и в предыдущем примере.

№ точки	x_1	x_2	y
$\mathbf{x}^{(1)}$	4	5,42	108,27
$\mathbf{x}^{(2)}$	6	5,42	149,95
$\mathbf{x}^{(3)}$	5	7,16	185,81

"Наихудшей" вершиной является $\mathbf{x}^{(3)}$ = (5; 7,16), а "наилучшей" - $\mathbf{x}^{(1)}$ = (4; 5,42). После отражения имеем новый симплекс с вершинами:

№ точки	x_1	x_2	y
$\mathbf{x}^{(1)}$	4	5,42	108,27
$\mathbf{x}^{(2)}$	6	5,42	149,95
$\mathbf{x}^{(4)}$	5	3,68	82,52

Поскольку отраженная точка $\mathbf{x}^{(4)}$ является "наилучшей", производим растяжение симплекса. Координаты новой вершины рассчитываем по формулам:

$$x_1^{(5)}=2\,x_1^{(4)} - (x_1^{(1)} + x_1^{(2)})/2= 5\ ;$$
$$x_2^{(5)}= 2\,x_2^{(4)} - (x_2^{(1)}+ x_2^{(2)})/2=1{,}94\ .$$

Теперь " наихудшей" точкой будет вершина симплекса с координатами $\mathbf{x}^{(2)}$ = (6; 5,42). Дальнейшие расчеты приведены в таблице.

№ точки	x1	x2	y
$\mathbf{x}^{(1)}$	4	5,42	108,27
$\mathbf{x}^{(2)}$	6	5,42	149,95
$\mathbf{x}^{(5)}$	5	1,94	49,05
Производим отражение вершины $\mathbf{x}^{(2)}$			
$\mathbf{x}^{(6)}$	3	1,94	25,28
Производим растяжение от вершины $\mathbf{x}^{(6)}$			
$\mathbf{x}^{(7)}$	1,5	0,2	3,53
Производим отражение вершины $\mathbf{x}^{(1)}$			
$\mathbf{x}^{(8)}$	2,5	-3,28	18,92
Производим отражение вершины $\mathbf{x}^{(5)}$			
$\mathbf{x}^{(9)}$	-1	-5,02	72,48

Координаты последнего симплекса приведены в таблице:

№ точки	x_1	x_2	y
$\mathbf{x}^{(7)}$	1,5	0,2	3,53
$\mathbf{x}^{(8)}$	2,5	-3,28	18,92
$\mathbf{x}^{(9)}$	-1	-5,02	72,48

Таким образом, вершина $\mathbf{x}^{(9)}$ оказалась "наихудшей". Для продолжения поиска необходимо произвести редукцию последнего симплекса.

Выбирается вершина, в которой функция принимает минимальное значение $\mathbf{x}^{(7)}$ = (1,5; 0,2). Две другие вершины будут расположены на серединах прилежащих к ней граней.

№ точки	x_1	x_2	y
$\mathbf{x}^{(7)}$	1,5	0,2	3,53
$\mathbf{x}^{(10)}$	2	-1,54	6,66
$\mathbf{x}^{(11)}$	0,25	-2,41	16,29

Дальнейшие расчеты приведены в таблице:

Последняя точка является “лучшей”, по сравнению с отраженной, но тем не менее, является “худшей” среди оставшихся, поэтому производим операцию сжатия. Координаты новой вершины рассчитываем по формулам:

№ точки	x_1	x_2	y
Производим отражение вершины $\mathbf{x}^{(11)}$			
$\mathbf{x}^{(12)}$	3,25	1,07	18,66
Производим редукцию относительно $\mathbf{x}^{(7)}$			
$\mathbf{x}^{(7)}$	1,5	0,2	3,53
$\mathbf{x}^{(13)}$	1,75	-0,67	3,96
$\mathbf{x}^{(14)}$	2,375	0,635	9,19
Производим отражение вершины $\mathbf{x}^{(14)}$			
$\mathbf{x}^{(15)}$	0,875	-1,105	4,48

$x_1^{(16)} = 0{,}5x_1^{(15)} + (x_1^{(7)} + x_1^{(13)})/4 = 1{,}25$; $x_2^{(16)} = 0{,}5x_2^{(15)} + (x_2^{(7)} + x_2^{(13)})/4 = -0{,}67$.

Дальнейшие расчеты сведем в таблицу:

№ точки	x_1	x_2	y
$\mathbf{x}^{(7)}$	1,5	0,2	3,53
$\mathbf{x}^{(13)}$	1,75	-0,67	3,96
$\mathbf{x}^{(16)}$	1,25	-0,67	3,13
Производим отражение вершины $\mathbf{x}^{(13)}$			
$\mathbf{x}^{(17)}$	1	0,2	2,08
Растяжение от вершины $\mathbf{x}^{(17)}$			
$\mathbf{x}^{(18)}$	0,625	0,635	1,72
Отражение вершины $\mathbf{x}^{(7)}$			
$\mathbf{x}^{(19)}$	0,375	-0,235	1,54
Растяжение ничего не дает, поэтому производим отражение вершины $\mathbf{x}^{(16)}$			
$\mathbf{x}^{(20)}$	-0,25	1,07	0,68
№ точки	x_1	x_2	y
Отражение вершины $\mathbf{x}^{(19)}$			
$\mathbf{x}^{(21)}$	-0,50	0,2	0,73
Отражение вершины $\mathbf{x}^{(18)}$			
$\mathbf{x}^{(22)}$	-1,125	1,505	0,40
Отражение вершины $\mathbf{x}^{(21)}$			
$\mathbf{x}^{(23)}$	-0,88	2,375	4,14
Редукция относительно $\mathbf{x}^{(22)}$			
$\mathbf{x}^{(22)}$	-1,125	1,505	0,4
$\mathbf{x}^{(24)}$	-1	1,94	1,767
$\mathbf{x}^{(25)}$	-0,688	1,29	0,443
Отражение вершины $\mathbf{x}^{(24)}$			
$\mathbf{x}^{(26)}$	-0,81	0,8525	0,023
Отражение вершины $\mathbf{x}^{(25)}$			
$\mathbf{x}^{(27)}$	-1,25	1,07	0,037
Отражение вершины $\mathbf{x}^{(22)}$			
$\mathbf{x}^{(28)}$	-0,938	0,418	0,61
Редукция относительно $\mathbf{x}^{(26)}$			
$\mathbf{x}^{(26)}$	-0,81	0,8525	0,023
$\mathbf{x}^{(29)}$	-0,875	0,635	0,19
$\mathbf{x}^{(30)}$	-1,03	0,96	0,0064
Отражение вершины $\mathbf{x}^{(29)}$			
$\mathbf{x}^{(31)}$	-0,969	1,18	0,076

В последней точке происходит выполнение условия окончания алгоритма, т.е. среднеквадратичная величина разности значений функции в вершинах симплекса и среднего ее значения составляет $0{,}015 < \varepsilon = 0{,}2$.

Таким образом, за точку минимума принимаем вершину симплекса, соответствующую минимуму целевой функции, т.е. $\mathbf{x}^* \approx \mathbf{x}^{(30)} = (-1{,}03;\ 0{,}96)$.

Траектория поиска показана на рис.7.

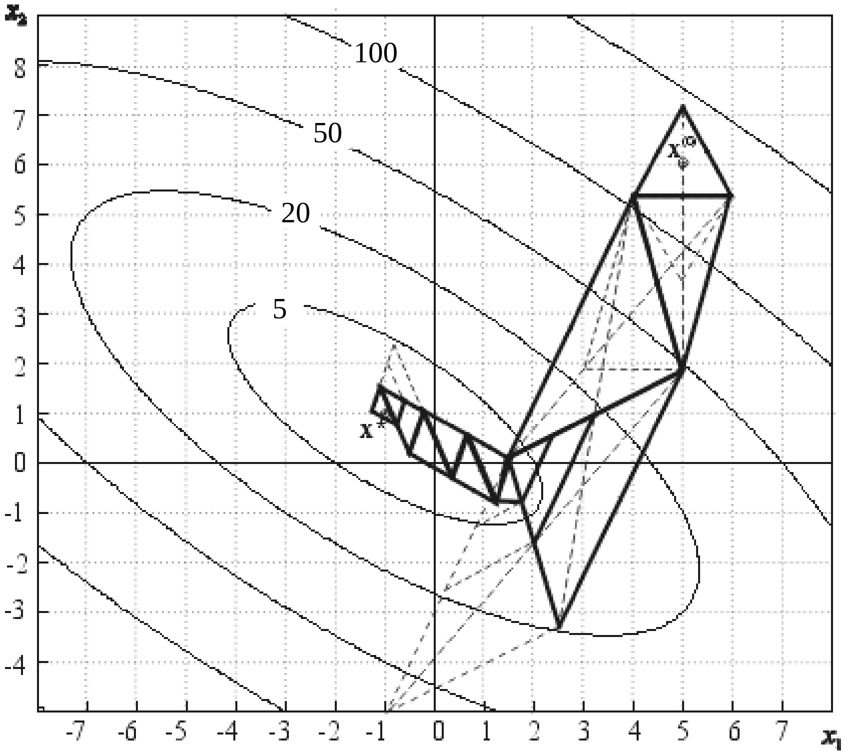


Рис.7

ВАРИАНТЫ ИНДИВИДУАЛЬНЫХ ЗАДАНИЙ ДЛЯ КОНТРОЛЬНЫХ РАБОТ 4-6

№ варианта	Вид целевой функции $f(x_1,x_2)$	Начальная точка		Точность ε
		$x_1^{(0)}$	$x_2^{(0)}$	
1	$(x_1 - 4x_2)^2 + (x_2 + 5)^2$	10	-5	0,15
2	$(x_1 + x_2)^2 + (x_2 + 4)^2$	9	5	0,3
3	$(x_1 - 3x_2)^2 + (x_2 - 2)^2$	4	10	0,25
4	$(x_1 - 3x_2)^2 + (x_2 + 1)^2$	0	8	0,15
5	$(x_1 + 5x_2)^2 + (x_2 - 1)^2$	8	10	0,35
6	$(x_1 + 2x_2)^2 + (x_2 - 3)^2$	0	10	0,18
7	$(x_1 - 2x_2)^2 + (x_2 - 3)^2$	-7	-7	0,2
8	$(x_1 + x_2)^2 + (x_2 + 2)^2$	6	-1	0,18
9	$(x_1 + 3x_2)^2 + (x_2 + 5)^2$	10	10	0,35
10	$(x_1 + 9x_2)^2 + (x_2 - 1)^2$	-6	5	0,25
11	$(x_1 - 2x_2)^2 + (x_2 - 9)^2$	15	12	0,15
12	$(x_1 + x_2)^2 + (x_2 + 6)^2$	10	8	0,18
13	$(x_1 + x_2)^2 + (x_2 - 1)^2$	5	6	0,15
14	$(x_1 - 2x_2)^2 + (x_2 - 3)^2$	7	6	0,25
15	$(x_1 + 2x_2)^2 + (x_2 - 4)^2$	-4	7	0,3
16	$(x_1 - 2x_2)^2 + (x_2 + 5)^2$	-15	5	0,2
17	$(x_1 - 6x_2)^2 + (x_2 + 1)^2$	-5	-3	0,2
18	$(x_1 - 5x_2)^2 + (x_2 + 6)^2$	-10	-5	0,18
19	$(x_1 + 4x_2)^2 + (x_2 - 3)^2$	-5	6	0,22
20	$(x_1 + 6x_2)^2 + (x_2 + 2)^2$	-10	7	0,22
21	$(x_1 - 7x_2)^2 + (x_2 - 2)^2$	8	6	0,3
22	$(x_1 - 8x_2)^2 + (x_2 + 1)^2$	-5	-5	0,2
23	$(x_1 - x_2)^2 + (x_2 - 7)^2$	10	2	0,25
24	$(x_1 + 8x_2)^2 + (x_2 - 2)^2$	-10	5	0,35
25	$(x_1 - 5x_2)^2 + (x_2 + 3)^2$	-10	-5	0,21

ВОПРОСЫ К ЭКЗАМЕНУ

1. Общая постановка задачи оптимизации. Классификация задач оптимизации.
2. Общая классификация методов скалярной оптимизации.
3. Основные этапы решения задач оптимизации.
4. Необходимые и достаточные условия экстремума функции одной переменной.
5. Классификация численных методов одномерной оптимизации. Методы сканирования и локализации оптимума.
6. Общая схема сужения промежутка унимодальности для одномерной функции. Методы половинного деления, золотого сечения и Фибоначчи.
7. Методы точечного оценивания экстремума одномерной функции. Метод обратного переменного шага, квадратичной аппроксимации, Пауэлла.
8. Необходимые и достаточные условия экстремума функции нескольких переменных.
9. Классификация численных методов многомерной оптимизации. Методы сканирования и локализации оптимума.
10. Методы покоординатного поиска экстремума функции нескольких переменных.
11. Метод Хука и Дживса.
12. Метод деформируемых многогранников Нельдера- Мида.
13. Обычные градиентные методы.
14. Методы наискорейшего спуска (крутого восхождения).
15. Методы случайного поиска экстремума.
16. Сравнительный анализ численных методов многомерной оптимизации.
17. Классификация задач и объектов. Задачи оптимизации при идентификации объектов и планировании эксперимента.
18. Технико-экономические задачи. Оптимальное проектирование, планирование и анализ функционирования объекта.
19. Задачи о рациональной загрузке оборудования, раскрое материалов, составлении расписаний.

РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

Основная учебная литература

1 Аттетков, А.В. Методы оптимизации [Электронный ресурс]: учебн. пособие / А.В. Аттетков, В.С. Зарубин, А.Н. Канатников - Электрон. текстовые дан.– Москва: РИОР: ИНФРА-М, 2019. – 270 с. Режим доступа: https://znanium.com/catalog/product/1002733

Дополнительная учебная литература

2 Сухарев, А.Г. Курс методов оптимизации [Электронный ресурс]: учебн. пособие / А. Г. Сухарев, А.В. Тимохов, В.В. Федоров - 2-е изд. – Электрон. текстовые дан.– Москва: ФИЗМАТЛИТ, 2011. – 384 с. Режим доступа: http://e.lanbook.com/view/book/2330/

3 Сдвижков, О.А. Практикум по методам оптимизации [Электронный ресурс]: учебн. пособие / О.А. Сдвижков - Электрон. текстовые дан.– Москва: Вузовский учебник, 2015. – 200 с. Режим доступа: http://znanium.com/bookread2.php?book=459517

4 Васильев, Ф. П. Методы оптимизации. Кн.2: Учебное пособие / Васильев Ф.П. - Москва :МЦНМО, 2011. - 433 с.: ISBN 978-5-94057-708-9. - Текст : электронный. - URL: https://znanium.com/catalog/product/958697

5 Бабенышев, С. В. Методы оптимизации : учебное пособие / С. В. Бабенышев, Е. Н. Матеров. - Железногорск : ФГБОУ ВО Сибирская пожарно-спасательная академия ГПС МЧС России, 2019. - 134 с. - Текст : электронный. - URL: https://znanium.com/catalog/product/1082159