МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Брянский государственный технический университет

УТВЕРЖДАЮ
Ректор университета
___________ О.Н. Федонин
«___» ___________ 2014 г.

ПРОЕКТ

ИНФОРМАТИКА (ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ)

НАХОЖДЕНИЕ ЭКСТРЕМУМА ФУНКЦИИ ДВУХ ПЕРЕМЕННЫХ СРЕДСТВАМИ ПРОГРАММ EXCEL И MATHCAD

Методические указания
к выполнению лабораторной работы № ???
для студентов очной формы обучения
всех технических специальностей
(квалификация «бакалавр»)

Брянск 2014

УДК 004.2

Информатика (информационные технологии). Нахождение экстремума функции двух переменных средствами программ EXCEL и MATHCAD: [Текст]+[Электронный ресурс]: метод. указания к выполнению лабораторной работы № ?? для студентов очной формы всех технических специальностей (квалификация «бакалавр»). – Брянск: БГТУ, 2014. –42 с.

Разработал: Зернин М.В., канд. техн. наук, доц.

Рекомендовано кафедрой «Информатика и программное обеспечение»
(протокол № ?? от ??.??.??)

1. ЦЕЛЬ И ЗАДАЧИ РАБОТЫ

При решении некоторых технических задач требуется определение оптимального решения, сводящихся к математической формулировке отыскания экстремума функции нескольких переменных (нескольких аргументов). В настоящей лабораторной работе рассмотрено решение этой задачи с применением численных методов, реализованных в программах EXCEL и MATHCAD. Решение таких задач проще всего продемонстрировать для функции двух переменных, т.к. можно показать понятные графические схемы алгоритмов различных численных методов. Отыскание экстремумов функции более чем двух переменных принципиально не отличается от разбираемых примеров. Но усложняется выполнение первого этапа задачи – этапа локализации решений из-за невозможности построения многомерных диаграмм.

Цель лабораторной работы заключается в повторении и закреплении знаний, полученных в дисциплине «Высшая математика» и в приобретении практических навыков в применения возможностей программ EXCEL и MATHCAD для нахождения экстремумов функции двух переменных.

Задачи лабораторной работы:

1. Изучить постановку задачи нахождения экстремумов функций многих переменных и основы численных методов ее решения.

2. Освоить процедуру «**Поиск решения**» программы EXCEL и применение ее для отыскания экстремумов функций двух переменных.

3. Освоить процедуры «**minimize**» и «**maximize**» отыскания минимумов и максимумов функции двух переменных в программе MATHCAD.

Продолжительность лабораторной работы – 8 часов, из них:

а) изучение методических указаний (МУ)– 4 часа (самостоятельная работа студентов);

б) выполнение лабораторной работы – 4 часа:

работа в EXCEL – 2 часа,

работа в MATHCAD – 2 часа.

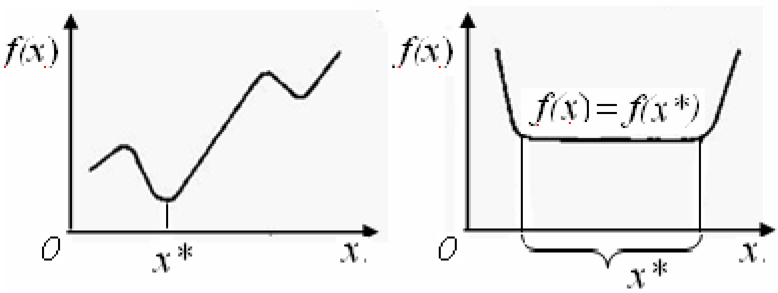
2. ТЕОРЕТИЧЕСКИЕ СВЕДЕНИЯ

Теория по указанной теме изложена во многих учебниках по курсам «Вычислительная математика» и «Численные методы». Данный раздел составлен на основе электронного ресурса [3] с дополнениями из других источников и изменениями, внесенными автором данных методических указаний.

2.1. Постановка задачи отыскания экстремумов функции одной и многих переменной

Нелинейная функция одной переменной $f(x)$ может иметь точки экстремумов. Решением задачи отыскания точки экстремума называют точку x^*, в которой целевая функция $f(x)$ достигает своего минимального или максимального значения.

Точка x^* является точкой **глобального минимума функции** одной переменной $f(x)$, заданной на числовой прямой R, если $x^*\in$ R и $f(x^*)<f(x)$ для всех $x^*\in R$ (рис. 1а). То есть среди всех минимумов функции $f(x)$ в этой точке достигается самое малое значение функции. Точка x^* называется точкой **строгого глобального минимума**, если это неравенство выполняется как строгое. Если же в выражении $f(x^*)\le f(x)$ равенство возможно при $x\neq x^*$, то имеем нестрогий минимум, а под решением в этом случае понимают множество $x^*=[x^*\in x{:}f(x)=f(x^*)]$ (рис. 1б).



а) б)

Рис. 1. Глобальный минимум:
а – строгий, б – нестрогий

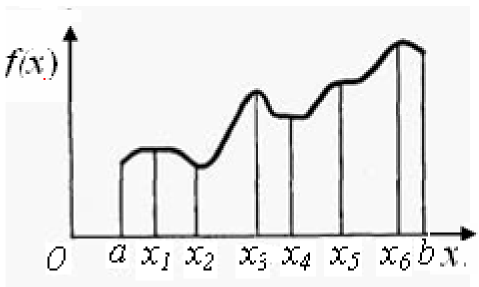


Рис. 2. Экстремумы функции

Точка $x^*\in$ M является точкой **локального минимума функции** $f(x)$ на множестве M, если при некотором достаточно малом $\delta>0$ для всех x, не равных x^*, $x\in$ M, удовлетворяющих условию $|x-x^*|\le \delta$, выполняется неравенство $f(x^*)<f(x)$. Если неравенство строгое, то x^*

является точкой **строгого локального минимума**. Все определения для максимума функции получаются заменой знаков предыдущих неравенств на обратные. На рис. 2 показаны экстремумы функции одной переменной *f(x)* на отрезке [*a, b*]. Здесь x_1, x_3, x_5 - точки локального максимума, а x_4 - локального минимума. В точке x_6 имеем глобальный максимум, а в точке x_2 – глобальный минимум.

Некоторые практические задачи сводятся к математической формулировке, в которой присутствует нелинейная **функция нескольких переменных** (нескольких аргументов)

$$z = f(x_1, x_2, ..., x_n). \tag{1}$$

Здесь **переменные** обозначены как x_1, x_2, x_i, x_n. Обычно требуется отыскать оптимальное решение, т.е. такое сочетание значений переменных $x_1^*, x_2^*, ... x_i^*, ... x_n^*$, которое соответствует минимуму или максимуму функции. Т.е. должно быть решена задача ***поиска оптимального решения*** (***задача оптимизации***). В редких случаях имеется возможность получить точное решение этой задачи аналитическими методами. Значительно чаще такие задачи решают численными методами и находят **приближенное решение**, соответствующее **допустимой погрешности.** Т.е. результатом решения этих задач является не точное, а приближенное (с заданной погрешностью) значение переменных $x_1^*, x_2^*, ... x_i^*, ... x_n^*$. Эти задачи решают в два этапа: **этап локализации точек экстремума** (определение интервала значений переменных, в пределах которого имеется лишь одна точка экстремума) и **этап уточнения** их положения (с допустимой погрешностью).

Охарактеризуем первый этап – локализацию положения точек экстремума, т.е. определение их начальных координат, которые на втором этапе должны быть уточнены. Уравнение (1) может иметь множество точек экстремумов и на первом этапе необходимо определить, какую из этих точек мы будем искать, т.е. необходимо выполнить **отделения искомого экстремума** от остальных и кроме того требуется **выявление факта, является ли он максимумом или минимумом функции**. Этот этап заключается в том, что указываются приближенные значения координат $x_1^*, x_2^*, ... x_i^*, ... x_n^*$ или для каждой из *i*-ой координаты выбирается отрезок $[a_i; b_i]$, в пределах которого содержится одна точка экстремума и не содержится других подобных точек. Для отделения искомого экстремума используют различные способы.

Для функции двух переменных можно построить диаграмму – поверхность и по ней определить вид экстремума (т. е выявить, это минимум или максимум) и приближенное положение точки экстремума или границы области, в пределах которой он локализован. Можно получить достаточно подробную таблицу значений функции (например, с помощью табличного процессора) и по ней определить то же самое. Современные программные средства позволяют качественно выполнить этап локализации за счет табуляции функции с мелким шагом изменения аргументов и построения ее диаграммы. Для функции, зависящей от более двух аргументов, этот этап выполняется сложнее, т.к. графические представления таких функций получить нельзя.

На этапе локализации точки экстремума выбраны некоторые начальные значения вектора переменных (**начальное приближение**). Далее строят схему перехода к следующему (более точному) приближению, посредством которой выполняется уточнение координат экстремума. Разработано достаточно много алгоритмов уточнения положения экстремальной точки. Но в основе всех методов решения задачи уточнения лежит **принцип последовательных приближений.** Поясним смысл этого принципа для задачи отыскания экстремума функции одной переменной.

Итак, необходимо найти приближенное значение координаты экстремума x_a^*, абсолютная погрешность которого не превышает заданного положительного числа ε. Приближенное значение x_a^* не единственно, так как заданной точности удовлетворяют все точки отрезка $\left[x^* - \varepsilon\ ,\ x^* + \varepsilon\right]$ и все они годятся в качестве искомого приближенного решения задачи. Мы будем искать любое из этих приближенных решений.

Сформулируем **принцип последовательных приближений**. Пусть известна числовая последовательность $\{x_n\}$, сходящаяся к x^* ($\lim\limits_{n\to\infty} x_n = x^*$). Согласно определению предела, для заданного $\varepsilon > 0$ найдется номер *N* такой, что для всех номеров *n*, больших или равных *N*, будет выполнено неравенство $\left|x^* - x_n\right| \le \varepsilon$. Иными словами, $x_n \in \left[x^* - \varepsilon\ ,\ x^* + \varepsilon\right]$ для всех номеров *n*, больших или равных *N*. Таким образом, все x_n при $n \ge N$ являются приближенными решениями задачи с погрешностью, не превышающей ε. Поэтому члены последовательности x_n называются **приближениями** (или **итерациями**). Все **методы**, в которых используется принцип последовательных приближений, называются **итерационными**.

Таким образом, решение задачи уточнения положения точки экстремума сводится к построению последовательности приближений и отысканию номера *N*. Аналогично и для функции многих переменных разные итерационные методы уточнения положения точки экстремума отличаются друг от друга только способом построения последовательности приближений и выбора значения *N*. Для вычисления приближенного решения, удовлетворяющего допустимой погрешности, необходимо построить такую последовательность приближений (которая в данном случае будет представлять собой последовательность векторов $\mathbf{x}^{(k)} = (x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)})$) сходящуюся к искомому точному решению $\mathbf{x}^* = (x_1^*, x_2^*, \dots, x_n^*)$. Здесь k – номер члена последовательности (**номер итерации**).

Как и большинство нелинейных задач, эти задачи могут иметь несколько решений. Поэтому прежде чем использовать процедуру уточнения решения в соответствии с каким-либо методом, необходимо корректно назначить начальное приближение. Обычно результат получается такой. Если начальное приближение ближе к какому-то из возможных решений, то будет найдено именно это решение. И только вводя дополнительные ограничения, можно регулировать процесс сходимости к тому или другому искомому решению.

2.2. Простейший алгоритм уточнения положения экстремума функции одной переменной

Методов и соответствующих алгоритмов уточнения значений характерных точек разработано много. Опишем простейший из методов – **метод деления отрезка пополам** (имеющий также другие названия: метод половинного деления, метод бисекции, метод дихотомии). Цель всех методов уточнения – п**остроение последовательности приближений, сходящейся к искомой точке.** Подробно метод деления отрезка пополам для отыскания корня и точки экстремума уравнения описан в методических указаниях к ранее выполняемой лабораторной работе № ???. Здесь кратко опишем этот алгоритм для уточнения положения точки экстремума.

Будем считать, что функция $f(x)$ непрерывна на отрезке $[a;b]$ и имеет на нем единственную точку экстремума (рис. 3). После вычисления положения середины участка $x_n = (a_n + b_n)/2$ необходимо определить, на какой из двух половин имеется экстремум, на рис. 3 – это максимум. Для этого необходимо вычислить два значения функции, на небольшом удалении от точки, разделяющей отрезок пополам, например, в точках $x_n – ε$ и $x_n + ε$. В последующей итерации оставляют тот участок, в котором вычисленное значений функции большее из двух. На рис. 3 это левый участок.

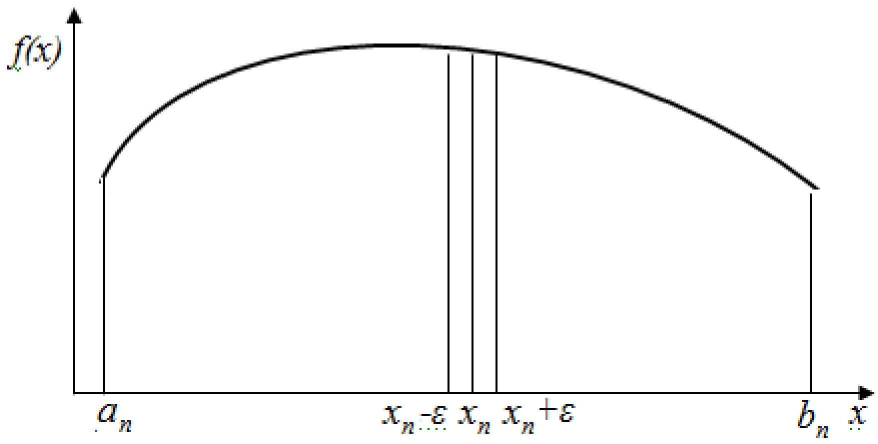


Рис. 3. Схема метода половинного деления при отыскании точки экстремума

Условие, позволяющее получить член последовательности приближений, погрешность которого не превышает заданного положительного числа ε (**условие окончания итераций**) состоит в следующем. Если $(b_n - a_n)/2 \le \varepsilon$, то в качестве приближенного значения корня с погрешностью, не превышающей ε, можно выбрать $x_a^* = (a_n + b_n)/2$.

Ранее применялись многие методы уточнения значений характерных точек, т.к. разные методы сходятся с разной скоростью. На скорость сходимости влияло как количество вычислений на одной итерации, так и требуемое количество итераций для решения задачи. При современной производительности ЭВМ отыскание характерных точек нелинейной функции одной переменной выполняется столь быстро, что необходимость применения более сложных, но быстрее сходящихся методов, становится сомнительной. Конечно, существуют некоторые вычислительные задачи, при решении которых находить экстремумы необходимо многократно. Для таких задач имеет смысл использовать более сложные итерационные методы, обеспечивающие быструю сходимость. Но эти методы требуют более тщательного обоснования возможности их применения и выбора начального приближения в том или ином случае. В большинстве же практических случаев, когда экстремумы нужно найти несколько раз, рационально применять метод деления отрезка пополам. Хотя он не обладает самой быстрой сходимостью, но условие его применимости и условие окончания итерационного процесса самые простые.

2.3. Основные определения и классификация методов безусловной оптимизации функции многих переменных

Методы **безусловной оптимизации** изложены в большом количестве учебной и научной литературы, ссылки на которую здесь не приводится. Создано и используется десятки методов и сотни вариантов алгоритмов их реализующих. Нами за основу выбрана классификация этих методов, приведенная в электронном учебнике А.Г.Трифонова [1]. Кроме того, добавлена дополнительная информация из других источников. Причем для лучшего понимания непростого материала добавлены имеющиеся в других источниках и нарисованные нами графические схемы сходимости процессов отыскания экстремума функции двух переменных. Алгоритмы кратко описаны на основе этих схем.

Основные определения. Решение многих теоретических и практических задач сводится к отысканию экстремума (наибольшего или наименьшего значения) скалярной функции $f(x)$ n-мерного векторного аргументах. В дальнейшем под x будем понимать вектор-столбец (точку в n-мерном пространстве).

Оптимизируемую функцию $f(x)$ называют целевой **функцией, функцией цели или критерием оптимальности.** В дальнейшем без ограничения общности будем говорить о поиске минимального значения функции $f(x)$ и записывать эту задачу следующим образом

$f(x) \to min.$

Вектор x^*, дающий минимум целевой функции, называют **оптимальным**.

Отметим, что задачу максимизации $f(x)$ можно заменить эквивалентной ей задачей минимизации или наоборот. Рассмотрим это на примере функции одной переменной (рис. 4). Если x^* - точка минимума функции $y=f(x)$, то для функции $y=-f(x)$ она является точкой максимума, так как графики функций $f(x)$ и $-f(x)$ симметричны относительно оси абсцисс. Итак, минимум функции $f(x)$ и максимум функции $-f(x)$ достигаются при одном и том же значении переменной x^*. Минимальное же значение функции $f(x)$ равно максимальному значению функции $-f(x)$, взятому с противоположным знаком, т.е. $\min f(x)=-max(f(x))$.

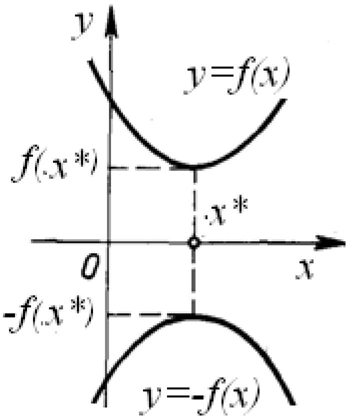


Рис. 4. Симметрия экстремумов функции одной переменной

Рассуждая аналогично, этот вывод нетрудно распространить на случай функции многих переменных. Если требуется заменить задачу минимизации функции $f(x_1, \ldots, x_n)$ задачей максимизации, то достаточно вместо отыскания минимума этой функции найти максимум функции - $f(x_1, \ldots, x_n)$. Экстремальные значения этих функций достигаются при одних и тех же значениях переменных. Минимальное значение функции $f(x_1, \ldots, x_n)$ равно максимальному значению функции - $f(x_1, \ldots, x_n)$, взятому с обратным знаком, т.е. $min\ f(x_1, \ldots, x_n) =max(-f(x_1, \ldots, x_n))$. Поэтому ниже рассматриваем только задачу минимизации.

В реальных условиях на сами переменные x_j, $j=1,2, \ldots. n$, и некоторые функции от этих переменных $g_i(x)$, $h_i(x)$, характеризующие качественные свойства объекта, системы, процесса, могут быть наложены **ограничения** (условия) вида $a \leq x \leq b$, где вектор x ограничивается соответствующими векторами констант, а кроме того возможны ограничения в виде равенств или неравенств

$g_i(x)=0, i=1, \ldots. n,$

$h_i(x) \leq 0, i=1, \ldots. n,$

Такую задачу называют задачей **условной оптимизации**. При отсутствии ограничений имеет место задача **безусловной оптимизации**.

Каждая точка x в n-мерном пространстве переменных $x_1, \ldots, x_n$, в которой выполняются ограничения, называется **допустимой точкой** задачи. Множество всех допустимых точек называют **допустимой областью** G. Решением задачи (оптимальной точкой) называют допустимую точку x^*, в которой целевая функция $f(x)$ достигает своего минимального значения. Понятия **локального и глобального мини-**

мума (максимума), строгого и нестрого минимума (максимума) приведены выше для нелинейной функции одной переменой. Существуют аналогичные понятия для функции нескольких переменных.

Мы в учебных целях рассматриваем задачу безусловной оптимизации функций двух переменных, как более легко реализуемую и достаточную для понимания алгоритмов решения таких задач в многомерных случаях. При этом предполагается, что предварительно выполнена задача локализации решения. То есть предварительно выделена область допустимых значений двух параметров, в которой имеется единственный экстремум. Для данной локальной области этот экстремум является глобальным.

Кроме того, задача оптимизации функции двух переменных позволяет наглядно представлять схемы сходимости на плоском рисунке с **поверхностями равного уровня**. Поверхностью уровня (на плоскости - линией уровня) является поверхность, получаемая приравниванием выражения функции $f(x)$ некоторой постоянной величине C, т. е. $f(x)=C$. Во всех точках этой поверхности функция имеет одно и то же значение C. Давая величине C различ-ные значения C_1, ..., C_n, получают ряд поверхностей, геометрически иллюстрирующих характер функции. В качестве примера на рис. 5 приведена поверхность функции двух переменных $f(x_1,x_2)$, имеющая минимум. Эта поверхность рассечена серией плоскостей, соответствующих различным значениям функции $f(x_1,x_2)$. Если спроецировать линии пересечения таких плоскостей на плоскость координат x_1x_2 , то получим серию линий равного уровня.

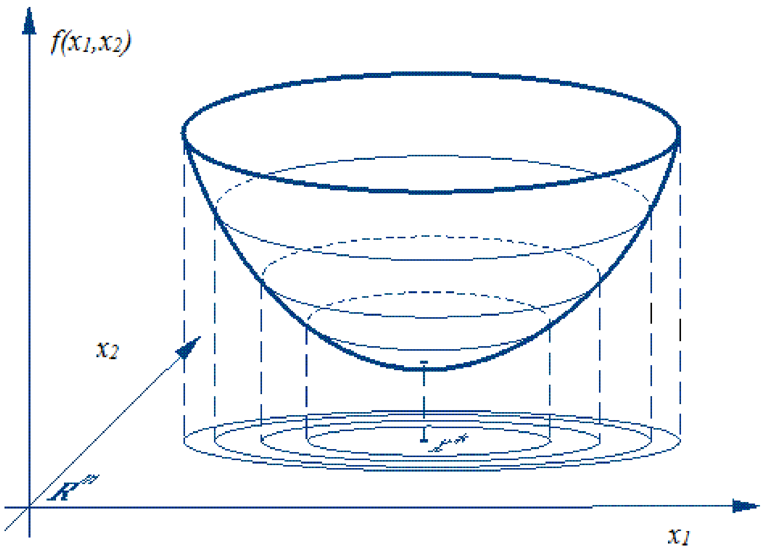


Рис. 5. Схема получения линий равного уровня в плоскости двух координат

Классификация методов безусловной оптимизации. Возможны два подхода к решению задачи отыскания минимума функции многих переменных $f(x)= f(x_1, ..., x_n)$ при отсутствии ограничений на диапазон изменения неизвестных. Первый подход лежит в основе

косвенных методов оптимизации и сводит решение задачи оптимизации к решению системы нелинейных уравнений, являющихся следствием условий экстремума функции многих переменных. Как известно, эти условия определяют, что в точке экстремума x^* все первые частные производные функции по независимым переменным равны нулю:

$$\left.\frac{\partial f}{\partial x_i}\right/_{x = x^*} = 0, \quad i=1,2, \ldots, n.$$

Эти условия образуют систему *n* нелинейных уравнений, после решений которой находятся точки минимума. Вектор $f'(x)$, составленный из первых частных производных функции по каждой переменной, т.е. вектор функций

$$f'(x) = (\partial f(x) / \partial x_1, \ldots, \partial f(x) / \partial x_n)^T,$$

называют **градиентом скалярной функции** $f(x)$. В точке минимума градиент равен нулю. Не всегда задача отыскания такой точки решается точно. Если используется метод последовательных приближений (итерационный метод), то процесс отыскания решения завершается при достижении заданной погрешности значений производных.

Решение систем нелинейных уравнений - задача весьма сложная и трудоемкая. Вследствие этого на практике чаще используют второй подход к минимизации функций, составляющий основу **прямых методов.** Суть их состоит в построении последовательности векторов $x[0], x[1], \ldots, x[n]$, таких, что $f(x[0]) > f(x[1]) > f(x[n]) > \ldots$ В качестве начальной точки $x[0]$ может быть выбрана произвольная точка, однако стремятся использовать всю имеющуюся информацию о функции $f(x)$, чтобы точка $x[0]$ располагалась как можно ближе к точке минимума. Эта задача решается на этапе локализации искомых точек экстремумов. Переход (итерация) от точки $x[k]$ к точке $x[k+1]$, $k = 0, 1, 2, \ldots$, состоит из **двух этапов**:

- **выбор направления движения** из точки $x[k]$;
- **определение шага вдоль этого направления**.

Методы построения таких последовательностей называют **методами спуска**, так как осуществляется переход от больших значений функций к меньшим.

Математически методы спуска описываются соотношением

$x[k+1] = x[k] + a_k\, p[k],\ k = 0, 1, 2, \ldots,$

где $p[k]$ - вектор, определяющий направление спуска; a_k - длина шага.

Различные методы спуска отличаются друг от друга способами выбора двух параметров - направления спуска и длины шага вдоль этого направления. На практике применяются только методы, обладающие **сходимостью**. Они позволяют за конечное число шагов получить точку минимума или подойти к точке, достаточно близкой к точке минимума. Качество сходящихся итерационных методов оценивают по **скорости сходимости**.

В методах спуска решение задачи теоретически получается за бесконечное число итераций. На практике вычисления прекращаются при выполнении некоторых **критериев (условий) останова итерационного процесса.** Например, это может быть **условие малости приращения аргумента**

$$|x[k] - x[k-1]| < \varepsilon$$

или **функции**

$$|f(x[x]) - f(x[k-1])| < \gamma.$$

Здесь k - номер итерации; справа заданные допустимые погрешности решения задачи. Также иногда используют **условия малости первых частных производных функции**.

Методы поиска точки минимума называются **детерминированными**, если оба элемента перехода от $x[k]$ к $x[k+l]$ (направление движения и величина шага) выбираются однозначно по доступной в точке $x[k]$ информации. Если же при переходе используется какой-либо случайный механизм, то алгоритм поиска называется **случайным поиском минимума**.

Детерминированные алгоритмы безусловной минимизации делят на **классы** в зависимости от вида используемой информации. Если на каждой итерации используются лишь значения минимизируемых функций, то метод называется **методом нулевого порядка**. Если, кроме того, требуется вычисление первых производных минимизируемой функции, то это **методы первого порядка**, при необходимости дополнительного вычисления вторых производных - **методы второго порядка**.

В настоящее время разработано множество численных методов решения задач как безусловной, так и условной оптимизации. Естественным является стремление выбрать для решения конкретной задачи наилучший метод, позволяющий за наименьшее время использования ЭВМ получить решение с заданной точностью.

Качество численного метода характеризуется многими факторами: необходимостью и сложностью предварительно выполняемых аналитически операций, скоростью сходимости, временем выполнения одной итерации, объемом памяти ЭВМ, необходимым для реализации метода, классом решаемых задач и т. д. Решаемые задачи также весьма разнообразны: они могут иметь высокую и малую размерность, быть унимодальными (обладающими одним экстремумом) и многоэкстремальными и т. д. Один и тот же метод, эффективный для решения задач одного типа, может оказаться совершенно неприемлемым для задач другого типа. Очевидно, что разумное сочетание разнообразных методов, учет их свойств позволят с наибольшей эффективностью решать поставленные задачи. Применяются также многометодные алгоритмы решения, в которых по ходу решения возможен переход с одного метода на другой, более эффективный для данного этапа решения.

Ниже из большого количества методов оптимизации кратко описаны те из них, которые реализованы в программных пакетах EXCEL и MATHCAD.

2.4. Некоторые численные методы безусловной оптимизации нулевого порядка

Общая характеристика методов нулевого порядка. В этих методах для определения направления спуска не требуется вычислять производные целевой функции. Направление минимизации в данном случае полностью определяется последовательными вычислениями значений функции. Следует отметить, что при решении задач безусловной минимизации методы первого и второго порядков обладают, как правило, более высокой скоростью сходимости, чем методы нулевого порядка. Однако на практике вычисление первых и вторых производных функции большого количества переменных весьма трудоемко. В ряде случаев они не могут быть получены в виде аналитических функций. Определение производных с помощью различных численных методов осуществляется с ошибками, которые могут ограничить применение таких методов. Кроме того, на практике встречаются задачи, решение которых возможно лишь с помощью методов нулевого порядка, например задачи минимизации функций с разрывными первыми производными. Критерий оптимальности может быть задан не в явном виде, а системой уравнений. В этом случае аналитическое или численное определение производных становится очень сложным, а иногда невозможным. Для решения таких практических задач оптимизации могут быть успешно применены методы нулевого порядка. Кратко охарактеризуем некоторые из них.

Метод прямого поиска (метод Хука-Дживса). Суть этого метода состоит в следующем. Задаются некоторой начальной точкой *x[0]*. Изменяя компоненты вектора *x[0]*, обследуют окрестность данной точки, в результате чего находят направление, в котором происходит уменьшение минимизируемой функции *f(x)*. В выбранном направлении осуществляют спуск до тех пор, пока значение функции уменьшается. После того как в данном направлении не удается найти точку с меньшим значением функции, уменьшают величину шага спуска. Если последовательные дробления шага не приводят к уменьшению функции, от выбранного направления спуска отказываются и осуществляют новое обследование окрестности и т. д. Достоинством метода прямого поиска является простота его программирования на компьютере. Он не требует знания целевой функции в явном виде, а также легко учитывает ограничения на отдельные переменные, а также сложные ограничения на область поиска. Недостаток метода прямого поиска состоит в том, что в случае сильно вытянутых, изогнутых или обладающих острыми углами линий уровня целевой функции он может оказаться неспособным обеспечить продвижение к точке минимума.

Метод деформируемого многогранника (метод Нелдера—Мида). Данный метод состоит в том, что для минимизации функции *n* переменных *f(x)* в *n*-мерном пространстве строится многогранник, содержащий *(n+1)* вершину. Очевидно, что каждая вершина соответствует некоторому вектору *x*. Вычисляются значения целевой функции *f(x)* в каждой из вершин многогранника, определяются максимальное из этих значений и соответствующая ему вершина *x[h]*. Через эту вершину и центр тяжести остальных вершин проводится проецирующая прямая, на которой находится точка *x[q]* с меньшим значением целевой функции, чем в вершине *x[h]* (рис. 6). Затем исключается вершина *x[h]*. Из оставшихся вершин и точки *x[q]* строится новый многогранник, с которым повторяется описанная процедура.

В процессе выполнения данных операций многогранник изменяет свои размеры, что и обусловило название метода.

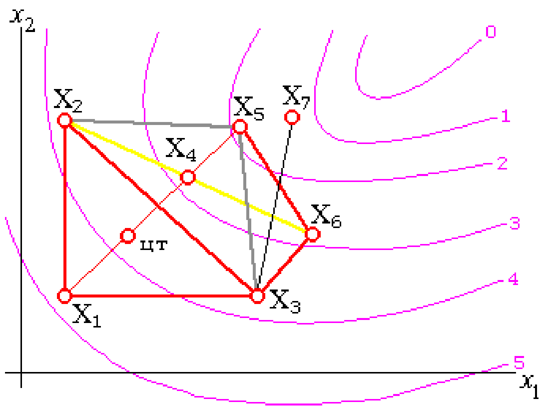


Рис. 6. Геометрическая интерпретация алгоритма метода деформируемого многогранника

С помощью операции растяжения и сжатия размеры и форма деформируемого многогранника адаптируются к топографии целевой функции. В результате многогранник вытягивается вдоль длинных наклонных поверхностей, изменяет направление в изогнутых впадинах, сжимается в окрестности минимума, что определяет эффективность рассмотренного метода.

Методы циклического покоординатного спуска. Согласно этом методам направления спуска выбирается параллельно координатным осям. Т.е. сначала спуск осуществляется вдоль первой оси OX_1, затем вдоль оси OX_2 и т.д. до последней оси OX_n. Потом эти действия повторяются в цикле вплоть до достижения оптимума. Варианты метода отличаются выбором шага *a*. На рис. 7. приведена схема метода циклического покоординатного спуска **с постоянным шагом** для функции двух переменных.

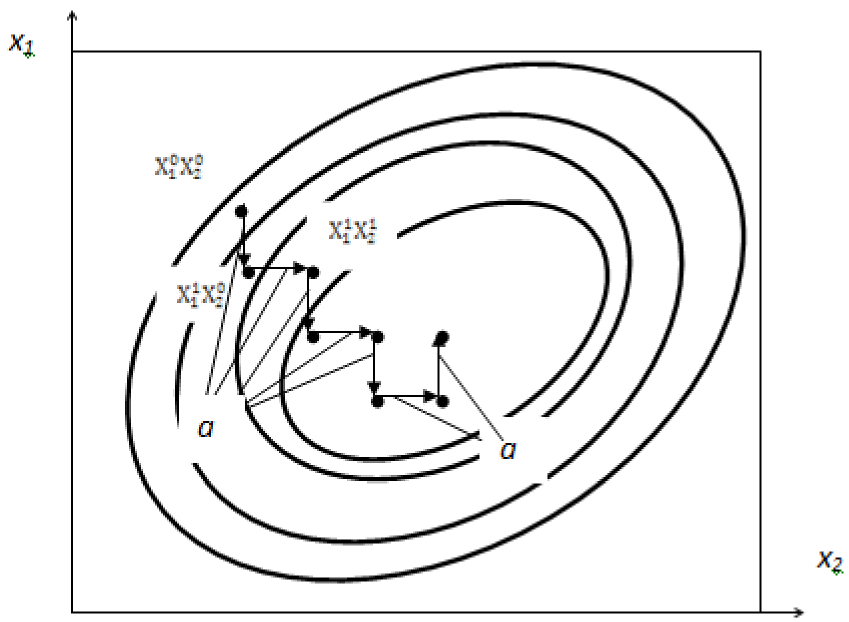


Рис. 7. Геометрическая интерпретация метода циклического покоординатного спуска с постоянным (малым) шагом

Пусть $x(0)$ – начальная точка, a – некоторое положительное число, выбранное как начальное значение шага. Вычисляют значение функции в этой точке $f(x(0))$. Далее вычисляют значение функции при измененном значении аргумента $x_1=x_1+a$. Если значение функции уменьшилось, то полагают, что точка первого приближения $x(1)$ сдвинута от начальной точки в направлении оси OX_1 на величину a.

Иначе вычисляют значение функции при $x_1=x_1-a$. Если значение функции уменьшилось, то полагают, что точка первого приближения $x(1)$ сдвинута от начальной точки против направления оси OX_1 на величину a. Если и в этом направлении значение функции не уменьшилось то точку не сдвигают из начального направления вдоль оси OX_1. Аналогичным образом осуществляют передвижение точки вдоль других осей. В этом состоит первая итерация.

Сходимость этого метода обеспечивается при малой величине шага, например, при шаге равном допустимой погрешности для значений координат $a\leq \varepsilon$. Однако на первых итерациях шаги имеет смысл назначать большими, что быстрее приведет к отысканию точки экстремума. В модификациях метода **с дробным шагом** изначально выбирают величину шага достаточно большую. Но, если в результате очередной итерации точка не сдвинулась из того положения, где она находилась до начала цикла по координатным осям, следует уменьшить величину шага, например разделив его на 2.

Другой вариант метода характеризуется таким же циклическим перебором направлений поиска поочередно вдоль всех n координатных осей, но шаг рассчитывается **на основе одномерной оптимизации.** На рис. 8 показан один шаг итерации, состоящий из двух этапов. Сначала фиксируется значение координаты x_2 и решается задача одномерной оптимизации по координате x_1. Т. е. определяется минимум одномерной функции, изображенной на выносном графике слева. Далее фиксируется значений координаты x_1 сообщающее минимум этой функции. И отыскивается минимум однопараметрической функции, график которой показан на выносном элементе внизу.

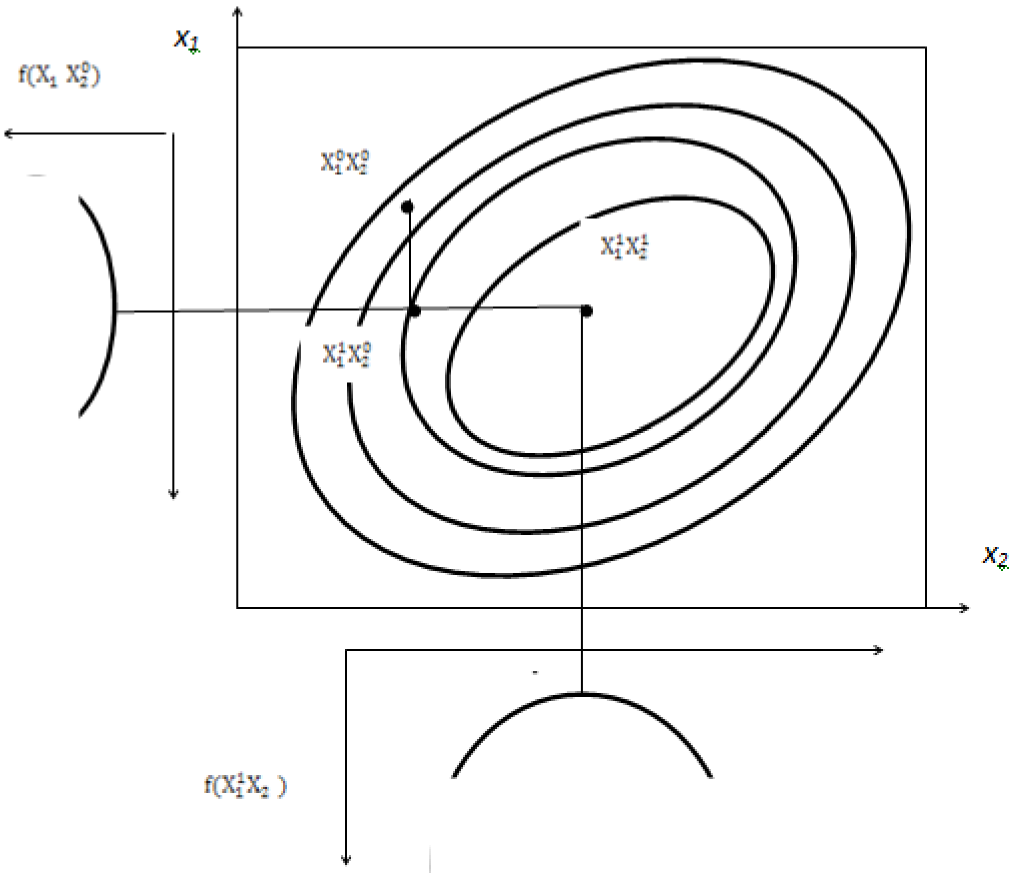


Рис. 8.Схема сходимости метода циклического покоординатного спуска с решением задачи одномерной оптимизации на каждом шаге

Но не для всех гладких функций применение этого метода гарантирует сходимость. Некоторый класс функций отображается поверхностями равного уровня, сильно вытянутыми в одном из направлений (как природные овраги). **Оврагом** называют часть пространства управляемых параметров, в которой наблюдаются слабые изменения производных целевой функции по одним направлениям и значительные изменения с переменой знака – по некоторым другим направлениям. Такие функции называют **овражными.** При использовании метода покоординатного спуска велика вероятность "застревания" поиска на дне оврага вдали от точки экстремума.. Для таких «овражных функций» разработаны специальные методы, отличающиеся изменением направления координат. А именно после «спуска на дно оврага», ось координат ориентируется вдоль оврага. Подробнее эти методы не рассматриваем.

2.5. Некоторые численные методы безусловной оптимизации первого порядка

Градиентом дифференцируемой функции *f(x)* в точке *x[0]* называется *n*-мерный вектор *f(x[0])*, компоненты которого являются частными производными функции *f(x)*, вычисленными в точке *x[0]*, т. е.

$$f'(x[0]) = (\partial f(x[0])/\partial x_1,\ \dots,\ \partial f(x[0])/\partial x_n)^T.$$

Этот вектор перпендикулярен к плоскости, проведенной через точку *x[0]*, и касательной к поверхности уровня функции *f(x)*, проходящей через точку *x[0]*. В каждой точке такой поверхности функция *f(x)* принимает одинаковое значение. Приравнивая функцию различным постоянным величинам $C_0 > C_1 > C_2, > \dots > C_n$, получим серию поверхностей равного уровня, характеризующих ее топологию. Если переменных две, то получаем серию линий равного уровня в плоскости координат (рис. 5). Вектор-градиент направлен в сторону наискорейшего возрастания функции в данной точке (рис. 9). Вектор, противоположный градиенту *(-f'(x[0]))*, называется **антиградиентом** и направлен в сторону наискорейшего убывания функции. В точке минимума градиент функции равен нулю. На свойствах градиента основаны методы первого порядка, называемые также градиентными методами минимизации. Использование этих методов в общем случае позволяет определить точку локального минимума функции.

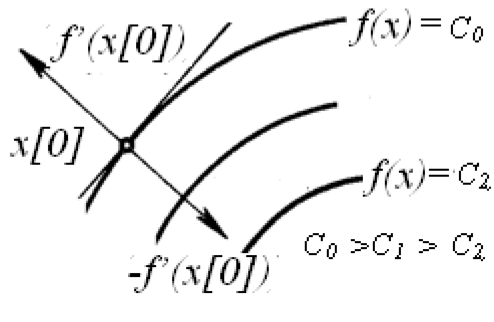


Рис. 9. Вектор градиента и вектор антиградиента функции цели

Очевидно, что если нет дополнительной информации, то из начальной точки *x[0]* разумно перейти в точку *x[1]*, лежащую в направлении антиградиента - наискорейшего убывания функции. Выбирая в качестве направления спуска *p[k]* антиградиент *-f'(x[k])* в точке *x[k]*, получаем итерационный процесс вида

$$x[k+1] = x[k] - a_k f'(x[k]),\ a_k > 0;\ k=0, 1, 2, \dots$$

В качестве критерия останова итерационного процесса используют либо выполнение условия малости приращения аргумента

$|x[k+l]-x[k]|<=\varepsilon$, либо выполнение условия малости градиента $|f'(x[k+l])|<=\xi$., где ε и ξ - заданные малые величины.

Возможен и комбинированный критерий, состоящий в одновременном выполнении двух указанных условий. Градиентные методы отличаются друг от друга способами выбора величины шага a_k. При методе с постоянным шагом для всех итераций выбирается некоторая постоянная величина шага a_k. Достаточно малый шаг a_k обеспечит убывание функции. Однако это может привести к необходимости проводить неприемлемо большое количество итераций для достижения точки минимума. С другой стороны, слишком большой шаг может вызвать неожиданный рост функции либо привести к колебаниям около точки минимума (зацикливанию). Из-за сложности получения необходимой информации для выбора величины шага методы с постоянным шагом применяются на практике редко. Более экономичны в смысле количества итераций и надежности градиентные методы с переменным шагом, когда в зависимости от результатов вычислений величина шага некоторым образом меняется.

В третьем варианте **метода наискорейшего спуска** на каждой итерации величина шага a_k выбирается из условия минимума функции $f(x)$ в направлении спуска, т. е.

$$f(x[k]-a_k f'(x[k])) = \min_{a>0} f(x[k]-af'(x[k])).$$

Это условие означает, что движение вдоль антиградиента происходит до тех пор, пока значение функции $f(x)$ убывает. С математической точки зрения на каждой итерации необходимо решать задачу одномерной минимизации по a функции $f(a) = f(x[k] - af'(x[k]))$. Геометрическая интерпретация такого метода приведена на рис 10. Причем на выносных элементах рисунка приведены графики одномерных функций, подлежащих оптимизации с целью определения шага.

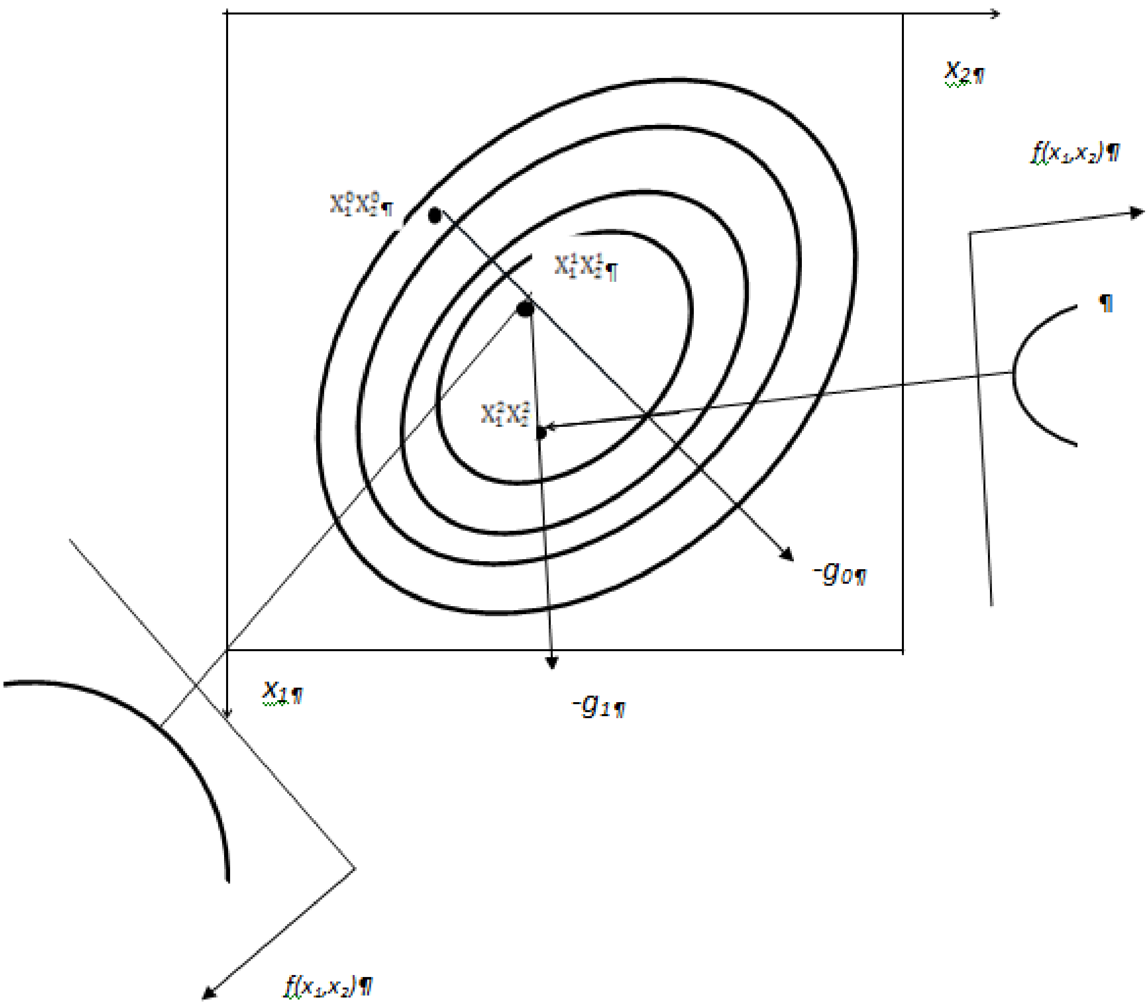


Рис. 10. Геометрическая интерпретация метода наискорейшего спуска (с решением задачи одномерной оптимизации на каждой итерации)

Градиентные методы сходятся к минимуму с высокой скоростью (со скоростью геометрической прогрессии) для гладких выпуклых функций. Однако на практике минимизируемые функции часто имеют плохо обусловленные матрицы вторых производных *(матрицы Гессе)*. Значения таких функций вдоль некоторых направлений изменяются гораздо быстрее (иногда на несколько порядков), чем в других направлениях. Их поверхности уровня в простейшем случае сильно вытягиваются, а сами такие функции называют **овражными.** Направление антиградиента этих функций может существенно отклоняется от направления в точку минимума, что приводит к замедлению скорости сходимости. Скорость сходимости градиентных методов существенно зависит также от точности вычислений градиента. Потеря точности, а это обычно происходит в окрестности точек минимума или в овражной ситуации, может вообще нарушить сходимость процесса градиентного спуска. Вследствие перечисленных причин градиентные методы зачастую используются в комбинации с другими, более эффективными методами на начальной стадии решения задачи. В этом случае точка $x[0]$ находится далеко от точки минимума, и шаги в направлении антиградиента позволяют достичь существенного убывания функции.

Другие варианты методов первого порядка здесь не рассматриваем.

2.6. Некоторые численные методы безусловной оптимизации второго порядка.

Особенности методов второго порядка. Методы безусловной оптимизации второго порядка используют вторые частные производные минимизируемой функции $f(x)$. Суть этих методов состоит в следующем. Необходимым условием экстремума функции многих переменных $f(x)$ в точке x^* является равенство нулю ее градиента в этой точке:

$$f'(x^*) = 0.$$

Разложение $f'(x)$ в окрестности точки $x[k]$ **в ряд Тейлора** с точностью до членов первого порядка позволяет переписать предыдущее уравнение в виде

$$f'(x) = f'(x[k]) + f''(x[k])\,(x - x[k]) = 0.$$

Здесь $f''(x[k]) = H(x[k])$ - матрица вторых производных (**матрица Гессе**) – матрица вторых частных производных минимизируемой функции по координатам

$$H(x) = \begin{vmatrix} \dfrac{\partial^2 f(x)}{\partial x_1 \partial x_1} & \dfrac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \dfrac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f(x)}{\partial x_2 \partial x_2} & \cdots & \dfrac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \dfrac{\partial^2 f(x)}{\partial x_n \partial x_1} & \dfrac{\partial^2 f(x)}{\partial x_n \partial x_2} & \cdots & \dfrac{\partial^2 f(x)}{\partial x_n \partial x_n} \end{vmatrix}$$

Следовательно, итерационный процесс для построения последовательных приближений к решению задачи минимизации функции $f(x)$ описывается выражением

$$x[k+l] = x[k] - H^{-1}(x[k])\, f'(x[k]),$$

где $H^{-1}(x[k])$ - обратная матрица для матрицы Гессе, а $-H^{-1}(x[k])\, f'(x[k]) = p[k]$ - направление спуска.

Полученный метод минимизации называют методом Ньютона. Очевидно, что в данном методе величина шага вдоль направления *p[k]* полагается равной единице. Последовательность точек *{x[k]}*, получаемая в результате применения итерационного процесса, при определенных предположениях сходится к некоторой стационарной точке *x** функции *f(x)*. Если матрица Гессе *H(x*)* положительно определена, точка *x** будет точкой строгого локального минимума функции *f(x)*. Последовательность *x[k]* сходится к точке *x** только в том случае, когда матрица Гессе целевой функции положительно определена на каждой итерации.

Если функция *f(x)* является квадратичной, то, независимо от начального приближения *x[0]* и степени овражности, с помощью метода Ньютона ее минимум находится за один шаг. Это объясняется тем, что направление спуска *p[k] = H-1(x[k])f'(x[k])* в любых точках *x[0]* всегда совпадает с направлением в точку минимума *x**. Если же функция *f(x)* не квадратичная, но выпуклая, метод Ньютона гарантирует ее монотонное убывание от итерации к итерации. При минимизации овражных функций скорость сходимости метода Ньютона более высока по сравнению с градиентными методами. В таком случае вектор *p[k]* не указывает направление в точку минимума функции *f(x)*, однако имеет большую составляющую вдоль оси оврага и значительно ближе к направлению на минимум, чем антиградиент.

Существенным недостатком метода Ньютона является зависимость сходимости для невыпуклых функций от начального приближения *x[0]*. Если *x[0]* находится достаточно далеко от точки минимума, то метод может расходиться, т. е. при проведении итерации каждая следующая точка будет более удаленной от точки минимума, чем предыдущая. Сходимость метода, независимо от начального приближения, обеспечивается выбором не только направления спуска $p[k] = H^{-1}(x[k])f'(x[k])$, но и величины шага *a* вдоль этого направления. Соответствующий алгоритм называют методом Ньютона с регулировкой шага. Итерационный процесс в таком случае определяется выражением

$$x[k+l] = x[k] - a_k H^{-1}(x[k])f'(x[k]).$$

Величина шага a_k выбирается из условия минимума функции *f(x)* по а в направлении движения, т. е. в результате решения задачи одномерной минимизации:

$$f(x[k] - a_k H^{-1}(x[k])f'(x[k]) = \min_{a\geq 0} (f(x[k] - aH^{-1}(x[k])f'(x[k])).$$

Такой вариант алгоритма называют также методом Ньютона-Рафсона. Графическая интерпретация этого варианта метода Ньютона представлена на рис. 11. На выносных элементах рисунка приведены графики одномерных функций, подлежащих оптимизации с целью определения шага.

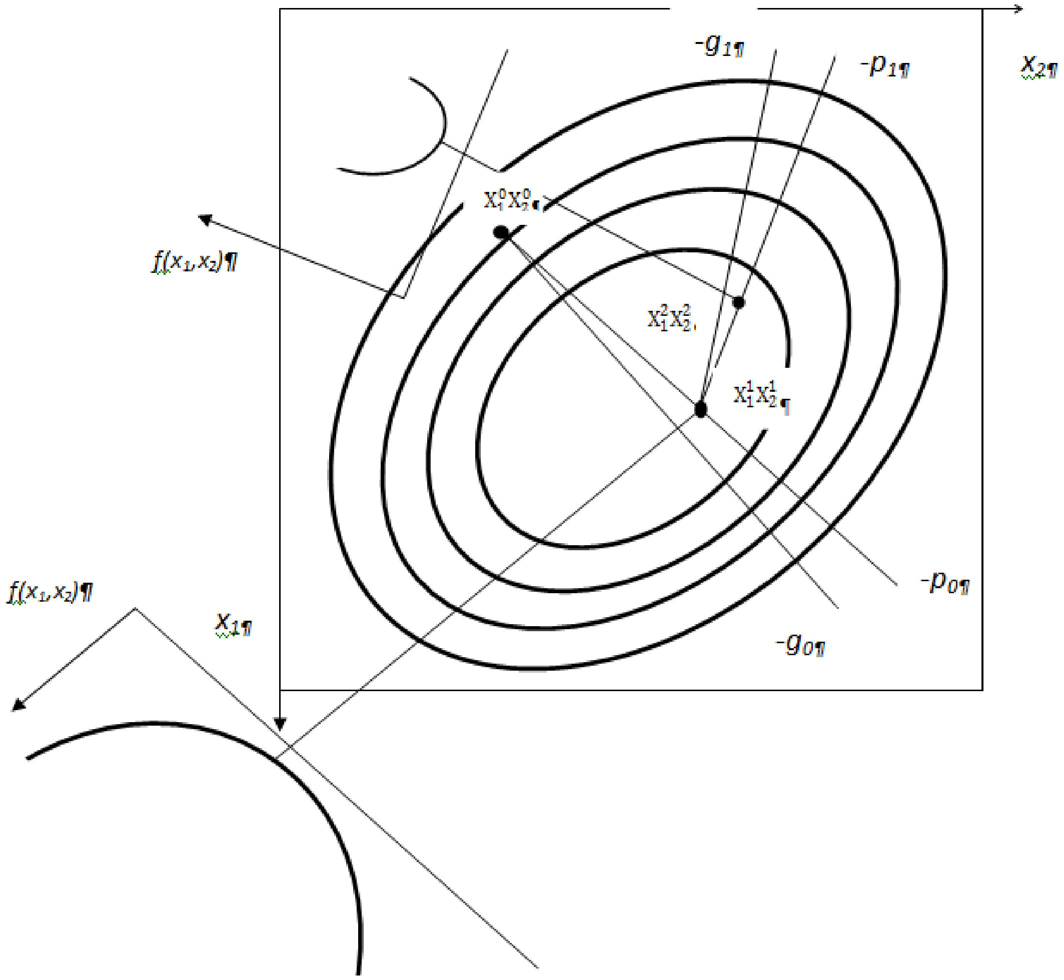


Рис. 11. Геометрическая интерпретация метода Ньютона-Рафсона

Алгоритм метода Ньютона-Рафсона состоит в следующем. Часть действий выполняется до начала итерационного процесса. А именно необходимо получить вектор формул, составляющих градиент $f'([x])$ (т.е. вектор первых частных производных) и матрицу формул, составляющих матрицу Гессе $H(x)$ (т.е. матрицу вторых частных производных). Далее в итерационном цикле в эти формулы подставляются значения компонент вектора x и эти массивы становятся массивами чисел.

1. В начальной точке $x[0]$ вычисляется вектор, определяющий направление спуска $p[0] = -H^{-1}(x[0])f'([0])$. Тем самым задача многомерная сводится к задаче одномерной оптимизации.

2. На k-й итерации определяется шаг a_k по схеме, изображенной на рис. 11, для этого решается задача одномерной оптимизации) и осуществляется переход в точку $x[k+1]$.

3. Вычисляется величина $f(x[k+1])$.

4. Проверяются условия завершения процесса поиска минимума функции. Эти условия аналогичны условиям, описанным в алгоритме метода наискорейшего спуска. Если эти условия выполняются, вычисления прекращаются. В противном случае вычисляется новое направление

$p[k+1] = -H^{-1}(x[k])f'([k])$

и осуществляется переход к следующей итерации, т. е. на шаг 2.

Количество вычислений на одной итерации методом Ньютона, как правило, значительно больше, чем в градиентных методах. Это объясняется необходимостью вычисления и обращения матрицы вторых производных целевой функции (или решения системы уравнений, что требует меньше трудозатрат). Однако на получение решения с достаточно высокой степенью точности с помощью метода Ньютона обычно требуется намного меньше итераций, чем при использовании градиентных методов. В силу этого метод Ньютона существенно более эффективен. Тем не менее в некоторых задачах трудоемкость итерации методом Ньютона может оказаться очень большой и применение его становится неэффективным.

3. ПРИМЕР НАХОЖДЕНИЯ ЭКСТРЕМУМА ФУНКЦИИ ДВУХ ПЕРЕМЕННЫХ В ПРОГРАММЕ EXCEL

3.1. Процедура «Поиск решения», реализующая итерационные методы уточнения решения нелинейных задач

В программе EXCEL среди надстроек имеется процедура «Поиск решения», реализующая итерационные методы уточнения решения нелинейных задач. При стандартной установке EXCEL эта процедура обычно недоступна, ее необходимо дополнительно один раз подключить. Для этого необходимо открыть закладку «**Файл**» и нажав на кнопку «**Параметры**», выбрать категории «**Надстройки**» (рис. 12). Убедившись, что в нижней части панели в поле «**Управле-**

ние» выбран элемент «**Надстройки EXCEL**», нажать кнопку «**Перейти**» (на рис. 12 выделено красным).

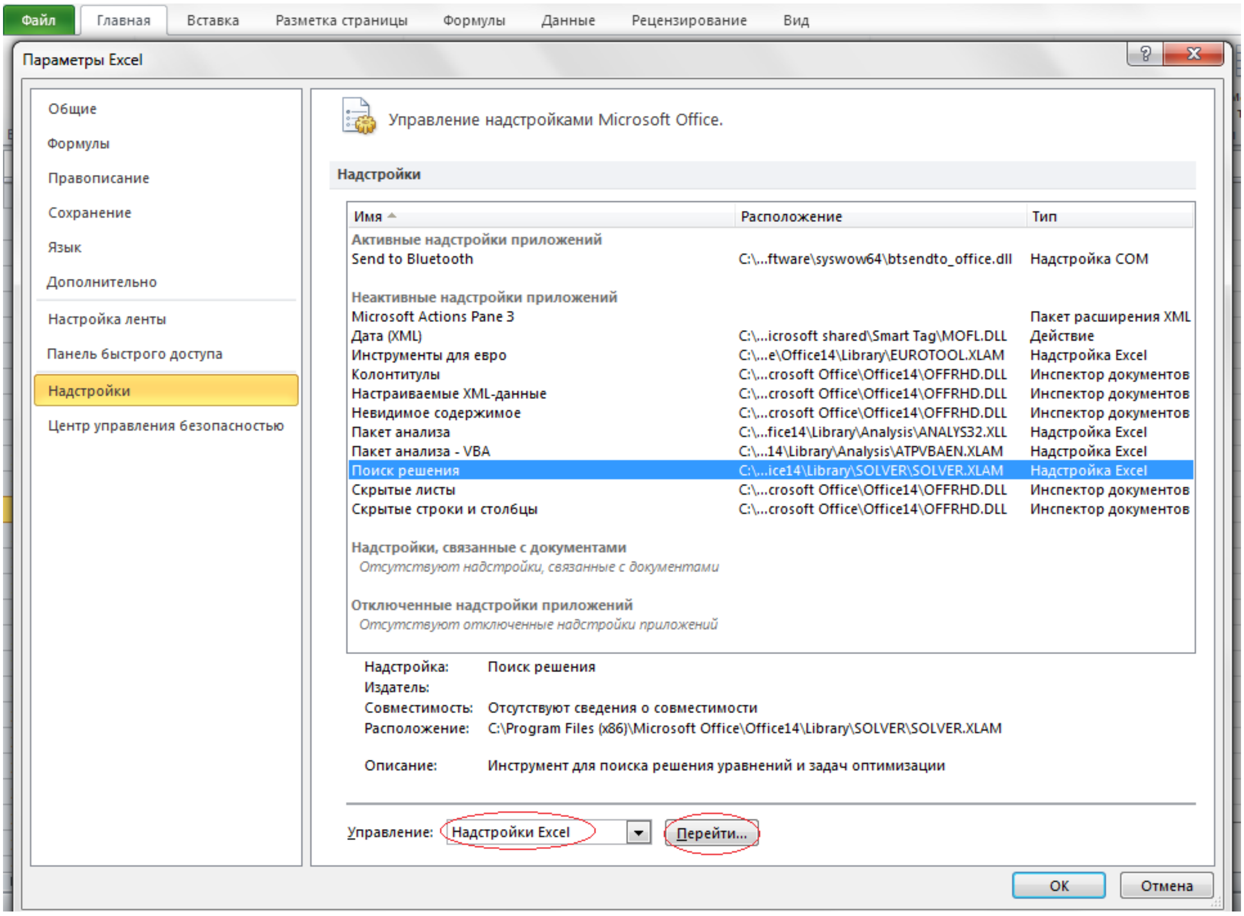


Рис. 12. Панель активизации Надстроек EXCEL

Появится панель «**Надстройки**» (рис.13), на которой следует поставить флажок у надстройки «**Поиск решения**». После этих действий строка надстройка «**Поиск решения**» переместится в список активных надстроек и на панели инструментов закладки «**Данные**» справа появится кнопка «**Поиск решения**» (рис. 13).

Процедура «Поиск решения» функционирует следующим образом. Задается некоторое начальное приближение в ячейке (называемой «ячейка переменной»), которая используется в ячейке с формулой, называемой оптимизируемой (целевой) ячейкой (выделено красным на рис. 14). Далее запускается итерационный процесс по уточнению значения в ячейке переменных для обеспечения некоторого условия в оптимизируемой ячейке. В процедуре «Поиск решения» кроме отыскания конкретного значения в оптимизируемой ячейке, можно назначить целью достижение там минимального или максимального значения.

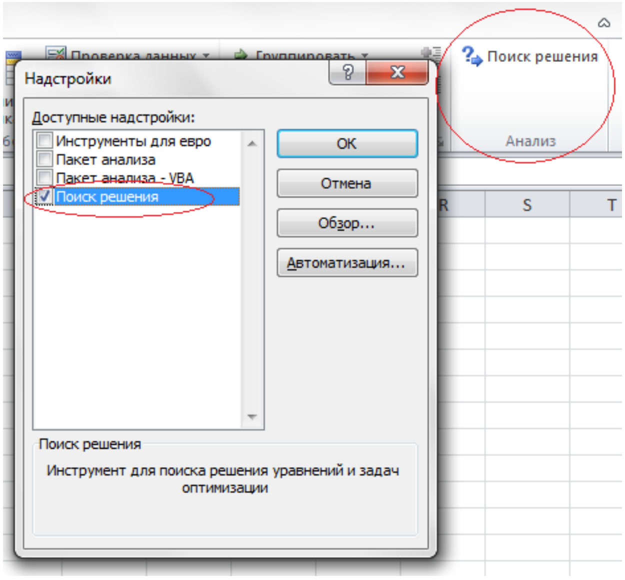


Рис. 13. Панель активизации процедуры «Поиск решения»

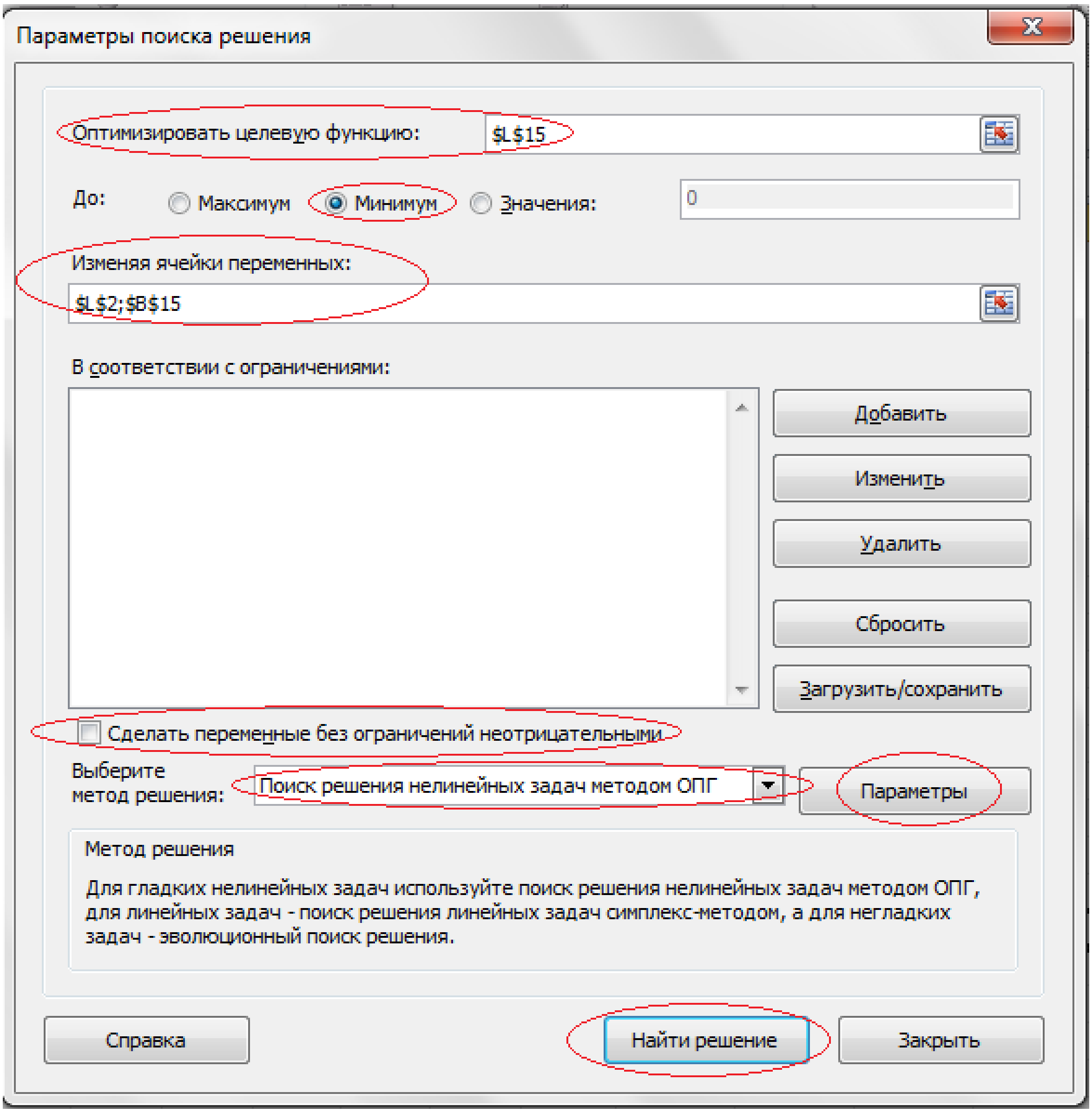


Рис. 14. Панель процедуры «Поиск решения»

В процедуре «Поиск решения» можно выбирать метод реализации итерационного процесса. Для анализа нелинейных функций целесообразно выбирать метод обобщенного понижающего градиента (ОПГ). Этот метод фактически является вариантом метода наискорейшего спуска, изложенного в разделе 2.5. Также в EXCEL-2010 можно выбрать один из методов прямого поиска, имеющий название «Эволюционный метод поиска» (рис. 15), применимый для негладких функций. Основы алгоритмов методов это категории изложены в разделе 2.4. Реализован также симплексный метод, который эффективен для специфической оптимизационной задачи, называемой «задача линейного программирования». Этот метод здесь не рассмотрен.

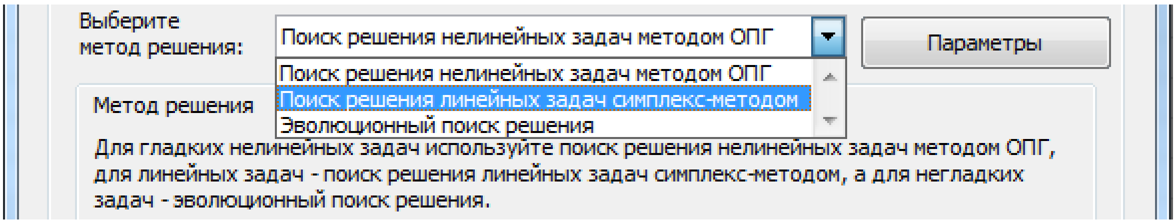
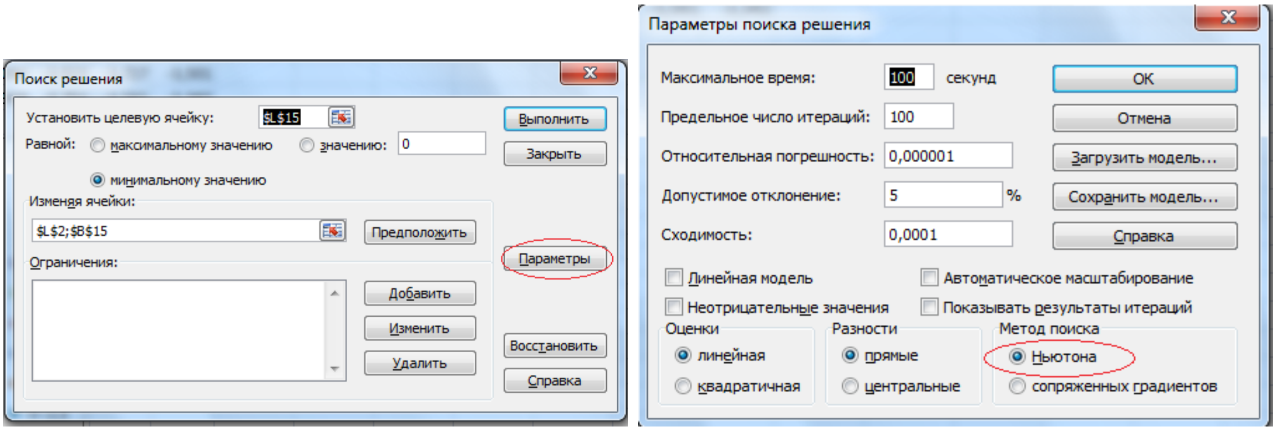


Рис. 15. Фрагмент панели процедуры «Поиск решения» с выбранным методом решения «Обобщенного понижающего градиента»

Кроме того в других версиях EXCEL реализованы другие методы уточнения решения. Так, в EXCEL-2007, у которого панель процедуры «Поиска решения» показана на рис. 16а, для уточнения решения нелинейной задачи можно выбрать (рис. 16б) метод второго порядка, а именно – метод Ньютона, кратко изложенный в разделе 2.6. Также можно выбрать метод сопряженных градиентов – один из вариантов градиентных методов (первого порядка).



а) б)

Рис. 16. Панель процедуры «Поиск решения» EXCEL-2007 (а) с выбранным методом Ньютона (б)

При решении некоторых задач (при наличии точек экстремума вблизи точек разрыва функции) потребуется задавать ограничения, На панелях процедур «Поиск решения» и в EXCEL-2010 (рис. 14) и в EXCEL-2007 (рис. 16) имеется возможность задавать такие ограничения. Кроме того в некоторых задачах целесообразно переменные без ограничений преобразовывать к неотрицательным значениям. В EXCEL-2010 эта функция установлена по умолчанию. В рассматриваемых нами задачах эти преобразования необходимо отключать (на рис. 14 отмечено красным) или задавать ограничения.

Итерационные процедуры, реализованные в процедурах «Поиск решения» можно настраивать (рис. 16б и рис.17), нажав на кнопку «**Параметры**» (см. рис. 14 и 16а).

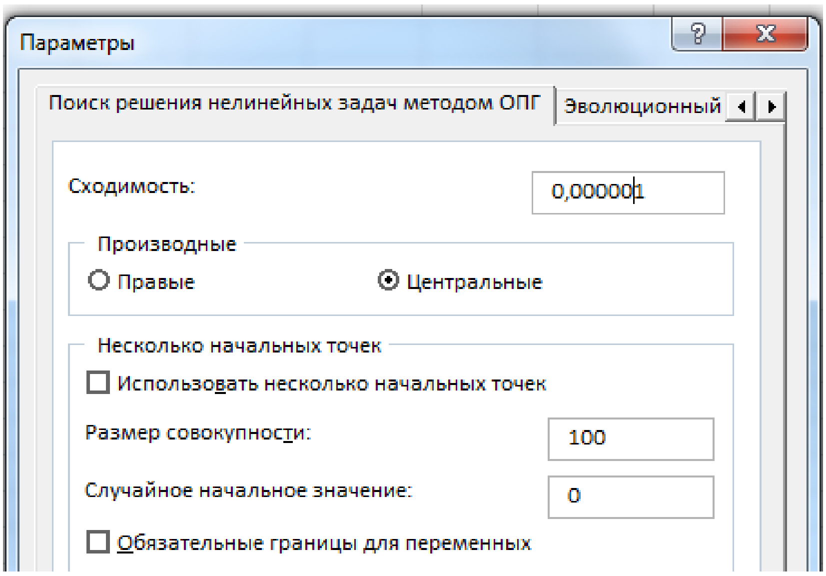


Рис. 17. Панель настройки параметров процедуры «Поиск решения» в EXCEL-2010

Формулирование задания дли поиска экстремума функции одной переменной подробнейшим образом описано в методических указаниях к лабораторной работе № ??? Здесь не будем повторяться. Краткое изложение приведено ниже – в примере отыскания минимума функции двух переменных.

3.2. Пример отыскания минимума функции двух переменных

В качестве примера рассмотрим функцию

$$z = -e^{\arccos\left(\frac{x^2-3x}{200}+\frac{y^2-5y}{320}\right)}. \qquad (2)$$

Требуется найти экстремум этой функции на множестве $\{x \in [-8;8],\ y \in [-10;10]\}$. Для этого необходимо протабулировать эту функцию, построить диаграмму двумерной поверхности, локализовать точку минимума (выбрать начальное приближение), с применением процедуры «Поиск решения» уточнить координаты точки, соответствующей минимальному значению функции.

Табуляция функции. Таблица значений анализируемой функции (2) приведена на рис. 18 для интервалов $x \in [-8;8]$ и $y \in [-10;10]$ с шагом 1. Причем приводим не самый красивый вариант таблицы, а самый рациональный, при реализации которого диаграмма строится простейшим образом и сразу с нужными значениями аргументов на двух осях. А именно, ячейка B2 оставлена пустой, а названия рядов значений аргументов *x* и *y* вынесены в ячейки A2 и B1. Важно отметить, что при построении такой таблицы необходимо использовать смешанную адресацию в формуле (отмечено красным), чтобы значения аргумента *x* всегда выбирались из второй строки (строка «заморожена» знаком $), а значения аргумента *y* всегда выбирались из столбца B (столбец «заморожен» знаком $).

C3 | =-EXP(ACOS((C2^2-3*C2)/200+($B3^2-5*$B3)/320))

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1		y																	
2	x		-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8
3		-10	-1,538	-1,843	-2,095	-2,312	-2,5	-2,659	-2,788	-2,887	-2,953	-2,987	-2,987	-2,953	-2,887	-2,788	-2,659	-2,5	-2,312
4		-9	-1,795	-2,08	-2,328	-2,547	-2,739	-2,903	-3,037	-3,14	-3,209	-3,244	-3,244	-3,209	-3,14	-3,037	-2,903	-2,739	-2,547
5		-8	-2,013	-2,293	-2,543	-2,768	-2,966	-3,136	-3,275	-3,382	-3,455	-3,492	-3,492	-3,455	-3,382	-3,275	-3,136	-2,966	-2,768
6		-7	-2,208	-2,488	-2,744	-2,974	-3,179	-3,355	-3,501	-3,612	-3,688	-3,727	-3,727	-3,688	-3,612	-3,501	-3,355	-3,179	-2,974
7		-6	-2,382	-2,667	-2,928	-3,166	-3,378	-3,561	-3,712	-3,829	-3,908	-3,948	-3,948	-3,908	-3,829	-3,712	-3,561	-3,378	-3,166
8		-5	-2,54	-2,829	-3,097	-3,342	-3,561	-3,751	-3,908	-4,029	-4,112	-4,154	-4,154	-4,112	-4,029	-3,908	-3,751	-3,561	-3,342
9		-4	-2,679	-2,974	-3,249	-3,501	-3,727	-3,923	-4,086	-4,212	-4,298	-4,341	-4,341	-4,298	-4,212	-4,086	-3,923	-3,727	-3,501
10		-3	-2,8	-3,101	-3,382	-3,641	-3,873	-4,076	-4,244	-4,374	-4,463	-4,508	-4,508	-4,463	-4,374	-4,244	-4,076	-3,873	-3,641
11		-2	-2,903	-3,209	-3,496	-3,761	-3,999	-4,207	-4,379	-4,513	-4,604	-4,651	-4,651	-4,604	-4,513	-4,379	-4,207	-3,999	-3,761
12		-1	-2,987	-3,297	-3,589	-3,858	-4,102	-4,314	-4,491	-4,628	-4,721	-4,769	-4,769	-4,721	-4,628	-4,491	-4,314	-4,102	-3,858
13		0	-3,05	-3,364	-3,66	-3,933	-4,18	-4,396	-4,576	-4,715	-4,81	-4,859	-4,859	-4,81	-4,715	-4,576	-4,396	-4,18	-3,933
14		1	-3,093	-3,409	-3,708	-3,984	-4,233	-4,451	-4,633	-4,775	-4,871	-4,92	-4,92	-4,871	-4,775	-4,633	-4,451	-4,233	-3,984
15		2	-3,114	-3,432	-3,732	-4,009	-4,26	-4,479	-4,662	-4,804	-4,902	-4,951	-4,951	-4,902	-4,804	-4,662	-4,479	-4,26	-4,009
16		3	-3,114	-3,432	-3,732	-4,009	-4,26	-4,479	-4,662	-4,804	-4,902	-4,951	-4,951	-4,902	-4,804	-4,662	-4,479	-4,26	-4,009
17		4	-3,093	-3,409	-3,708	-3,984	-4,233	-4,451	-4,633	-4,775	-4,871	-4,92	-4,92	-4,871	-4,775	-4,633	-4,451	-4,233	-3,984
18		5	-3,05	-3,364	-3,66	-3,933	-4,18	-4,396	-4,576	-4,715	-4,81	-4,859	-4,859	-4,81	-4,715	-4,576	-4,396	-4,18	-3,933
19		6	-2,987	-3,297	-3,589	-3,858	-4,102	-4,314	-4,491	-4,628	-4,721	-4,769	-4,769	-4,721	-4,628	-4,491	-4,314	-4,102	-3,858
20		7	-2,903	-3,209	-3,496	-3,761	-3,999	-4,207	-4,379	-4,513	-4,604	-4,651	-4,651	-4,604	-4,513	-4,379	-4,207	-3,999	-3,761
21		8	-2,8	-3,101	-3,382	-3,641	-3,873	-4,076	-4,244	-4,374	-4,463	-4,508	-4,508	-4,463	-4,374	-4,244	-4,076	-3,873	-3,641
22		9	-2,679	-2,974	-3,249	-3,501	-3,727	-3,923	-4,086	-4,212	-4,298	-4,341	-4,341	-4,298	-4,212	-4,086	-3,923	-3,727	-3,501
23		10	-2,54	-2,829	-3,097	-3,342	-3,561	-3,751	-3,908	-4,029	-4,112	-4,154	-4,154	-4,112	-4,029	-3,908	-3,751	-3,561	-3,342

Рис. 18. Табулирование функции двух переменных в EXCEL

Построение диаграммы (поверхности) и локализация начального приближения точки экстремума. На основе полученной таблицы построим поверхность. При таком оформлении таблицы, как указано выше, достаточно выделить диапазон данных (B2:S23) и выбрать соответствующий тип диаграммы (см. рис. 19 - закладка **Вставка**, среди вариантов **Другие диаграммы** выбрать **Поверхность** с линиями равного уровня). Появляется (рис. 20) поверхность с указанными значениями аргументов на осях. На поверхности (рис. 20) различным цветом выделены различные уровни значения функции и справа имеется **Легенда** с указанием, какому цвету какой уровень значений функции соответствует.

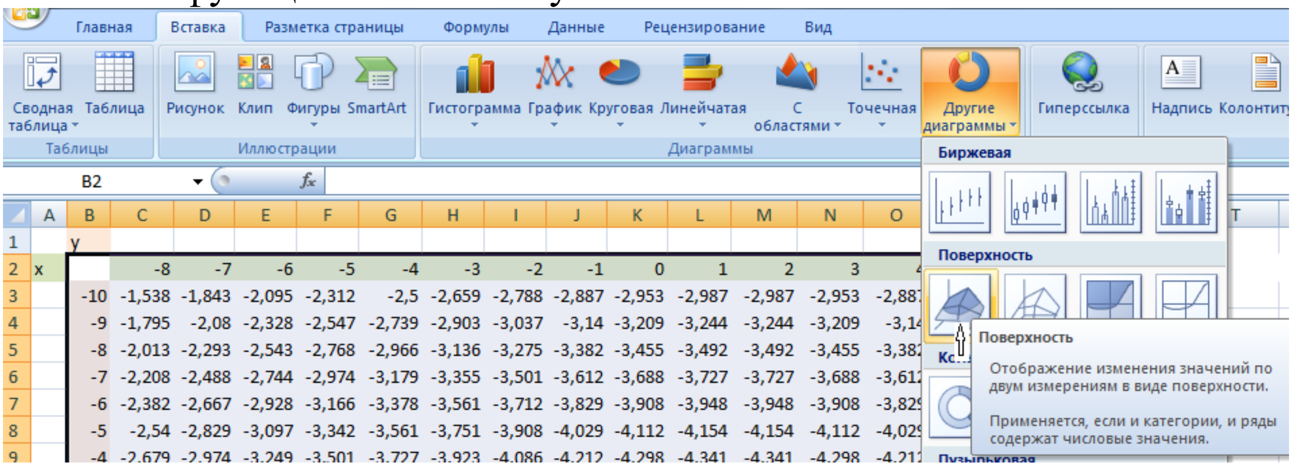


Рис. 19. Выбор типа диаграммы для построения поверхности

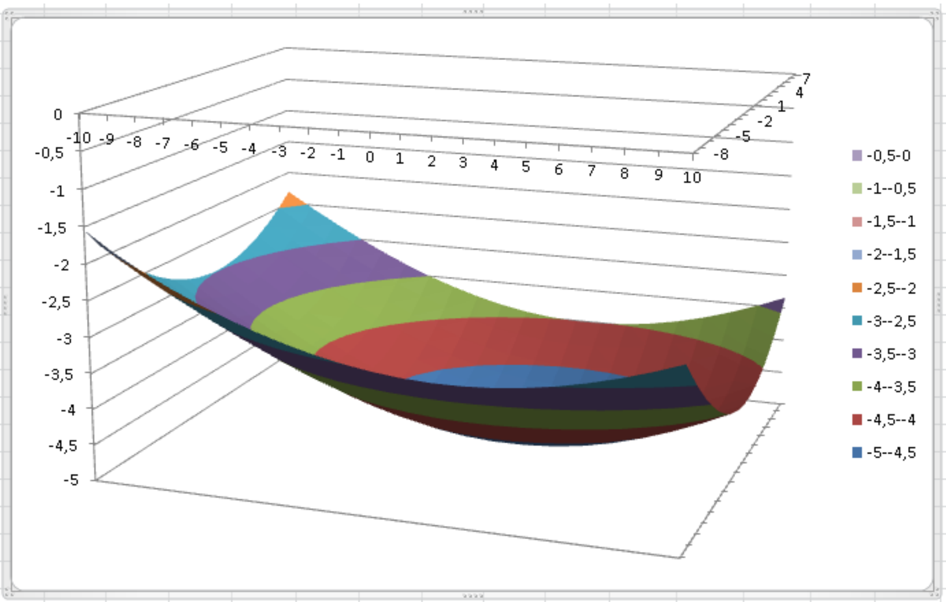


Рис. 20. Диаграмма поверхности, образуемая функцией двух переменных в EXCEL

Локализуем точку минимума. Из диаграммы (рис. 20) и таблицы (рис. 18) определяем, что минимум расположен вблизи точки x=1; y=2.

Уточнение положения минимума. С помощью надстройки «Поиск решения» найдем точку минимума (т. е уточним ее положение по сравнению с начальным приближением x=1; y=2). Для этого настроим соответствующее диалоговое окно следующим образом (рис. 21). Значение аргумента x=1 содержится в ячейке L2; значение аргумента y=2 находится я ячейке B15. Оптимизируемая ячейка L15, и в ней мы запрашиваем получение минимума функции за счет изменения содержимого ячеек L2 и B15.Все эти фрагменты на рис. 21 обведены красным. Обратим внимание на индикатор, преобразования отрицательных переменных в положительные (обведено красным) и на выбор метода решения задачи оптимизации (обведено красным).

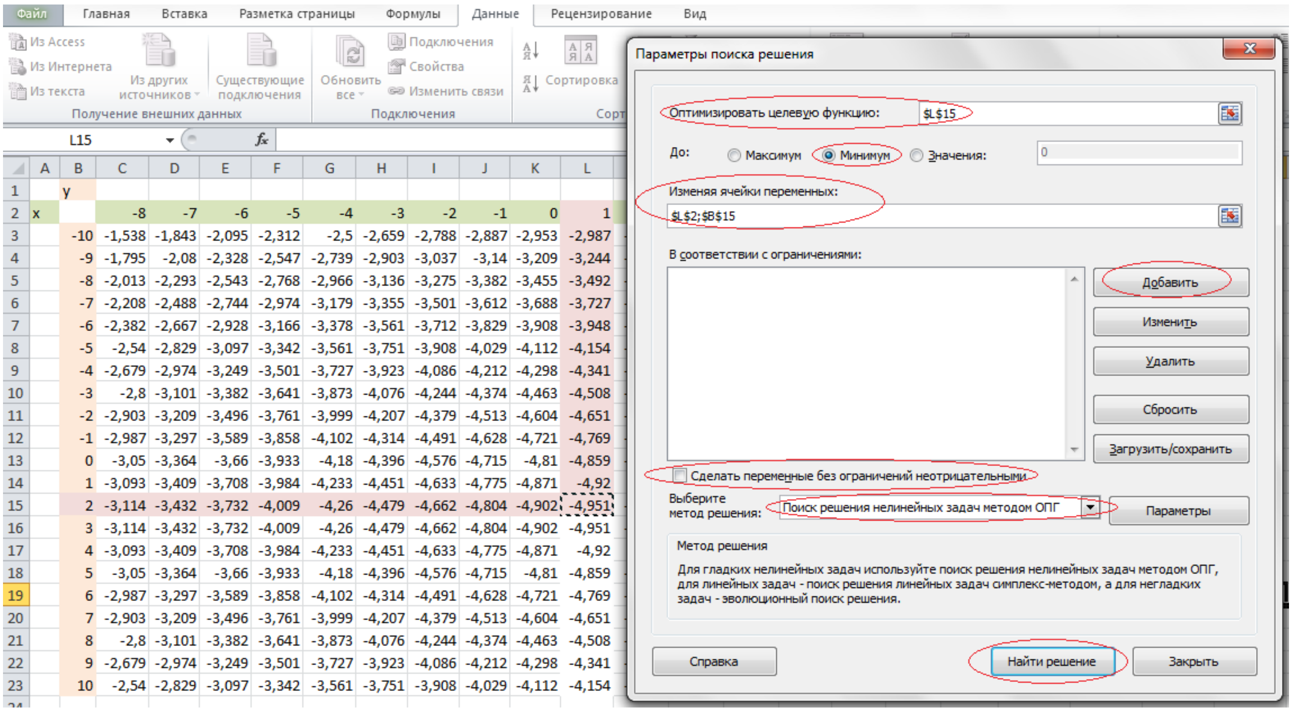


Рис. 21. Настройка панели процедуры «Поиск решения» для нахождения минимума функции двух переменных

После нажатия кнопки **Найти решение** появится (рис. 22) панель с резолюцией, что решение найдено. На этой панели можно указать необходимость вывести отчет (отмечено красным). В результате выполнения получили искомую точку минимума: x=1,5; y=2,5 (рис. 23).). Значение аргумента в ячейке L2 вместо исходного x=1 изменилось на x=1,5. Значение аргумента ячейке B15 вместо исходного y=2 изменилось на *y*=2,5. В оптимизируемой ячейке L15 вместо исходного значения -4,951 появилось меньшее значение -4,961, которое является минимальным значением функции. Косвенным подтверждением правильности решения является тот факт, что форма диаграммы поверхности (рис. 23) практически не изменилась.

На рис. 24 приведен фрагмент отчета о решении задачи.

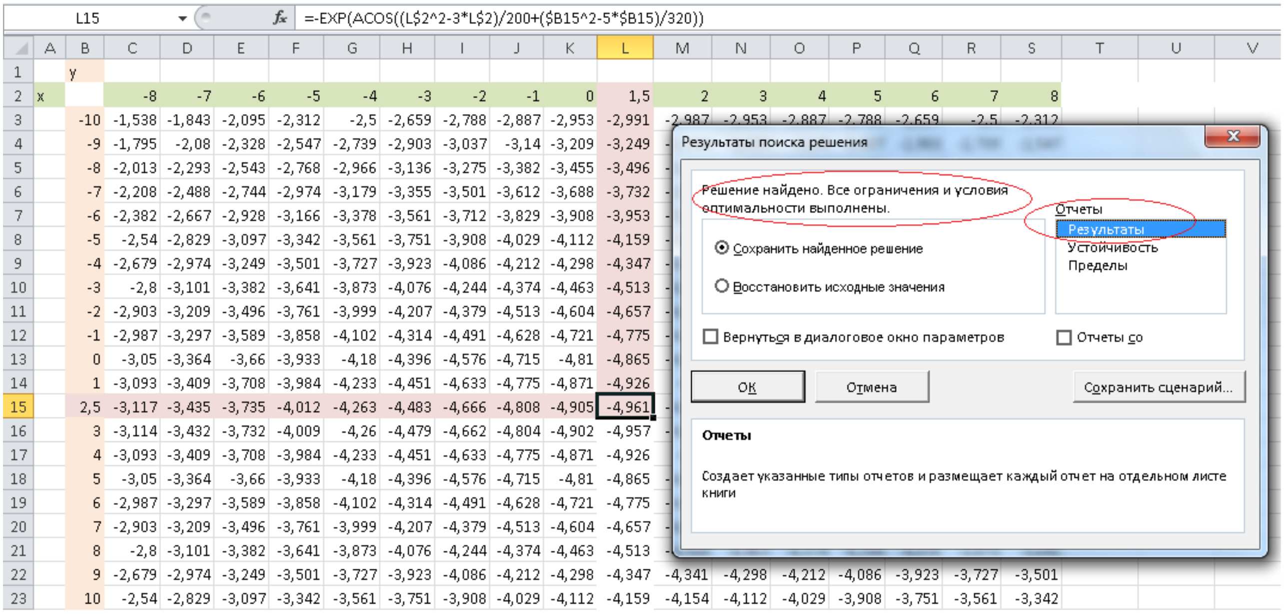


Рис. 22. Результаты поиска решения с найденным минимумом

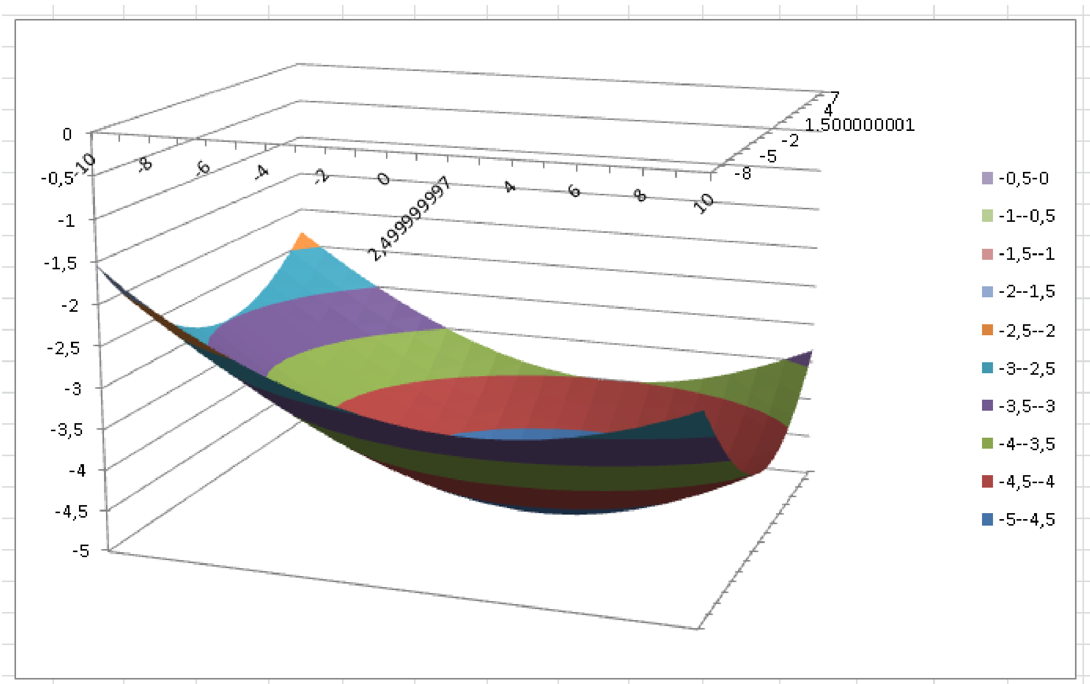


Рис. 23. Диаграмма (поверхность) с найденным минимумом

Параметры поиска решения

Максимальное время Без пределов, Число итераций Без пределов, Precision 0,000001
Сходимость 0,0001, Размер совокупности 100, Случайное начальное значение 0,
Центральные производные Максимальное число подзадач Без пределов,
Максимальное число целочисленных решений Без пределов,
Целочисленное отклонение 1%

Ячейка целевой функции (Минимум)

Ячейка	Имя	Исходное значение	Окончательное значение
L15		-4,960876504	-4,960876504

Ячейки переменных

Ячейка	Имя	Исходное значение	Окончательное значение	Целочисленное
L2	x	1,500000001	1,500000001	Продолжить
B15	y	2,499999997	2,499999997	Продолжить

Ограничения

НЕТ

Рис. 24. Фрагмент отчета о результатах поиска решений

Иногда при реализации итерационной процедуры для разрывной функции возможно попадание очередного приближения в область разрыва. Для недопущения такого нежелательного «аварийного» завершения итерационного процесса, следует назначать ограничения. В рассматриваемом примере такого «срыва в пропасть» не произошло. Тем не менее, для этой функции покажем (рис. 25), как настраивается панель процедуры «Поиск решения» при необходимости назначить ограничения содержимого ячеек L2 и B15 в заданных пределах. Для этого на панели настройки процедуры «Поиск решения» на отыскание минимума (рис. 21) следует нажать кнопку «Добавить» в разделе «Ограничения». На появившейся панели задания ограничений (рис. 25) заполнить соответствующие поля.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1		y																	
2	x		-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8
3		-10	-1,538	-1,843	-2,095	-2,312	-2,5	-2,659	-2,788	-2,887	-2,953	-2,987							
4		-9	-1,795	-2,08	-2,328	-2,547	-2,739	-2,903	-3,037	-3,14	-3,209	-3,244							
5		-8	-2,013	-2,293	-2,543	-2,768	-2,966	-3,136	-3,275	-3,382	-3,455	-3,492							
6		-7	-2,208	-2,488	-2,744	-2,974	-3,179	-3,355	-3,501	-3,612	-3,688	-3,727							
7		-6	-2,382	-2,667	-2,928	-3,166	-3,378	-3,561	-3,712	-3,829	-3,908	-3,948							
8		-5	-2,54	-2,829	-3,097	-3,342	-3,561	-3,751	-3,908	-4,029	-4,112	-4,154							
9		-4	-2,679	-2,974	-3,249	-3,501	-3,727	-3,923	-4,086	-4,212	-4,298	-4,341							
10		-3	-2,8	-3,101	-3,382	-3,641	-3,873	-4,076	-4,244	-4,374	-4,463	-4,508							

Добавление ограничения
Ссылка на ячейки: L2 >= Ограничение: 0
ОК Добавить Отмена

Рис. 25. Задание ограничение на значение аргумента при уточнении положения минимума разрывной функции (4)

В результате (рис. 26) на панели процедуры «Поиск решения» появится ограничения (выделено красным).

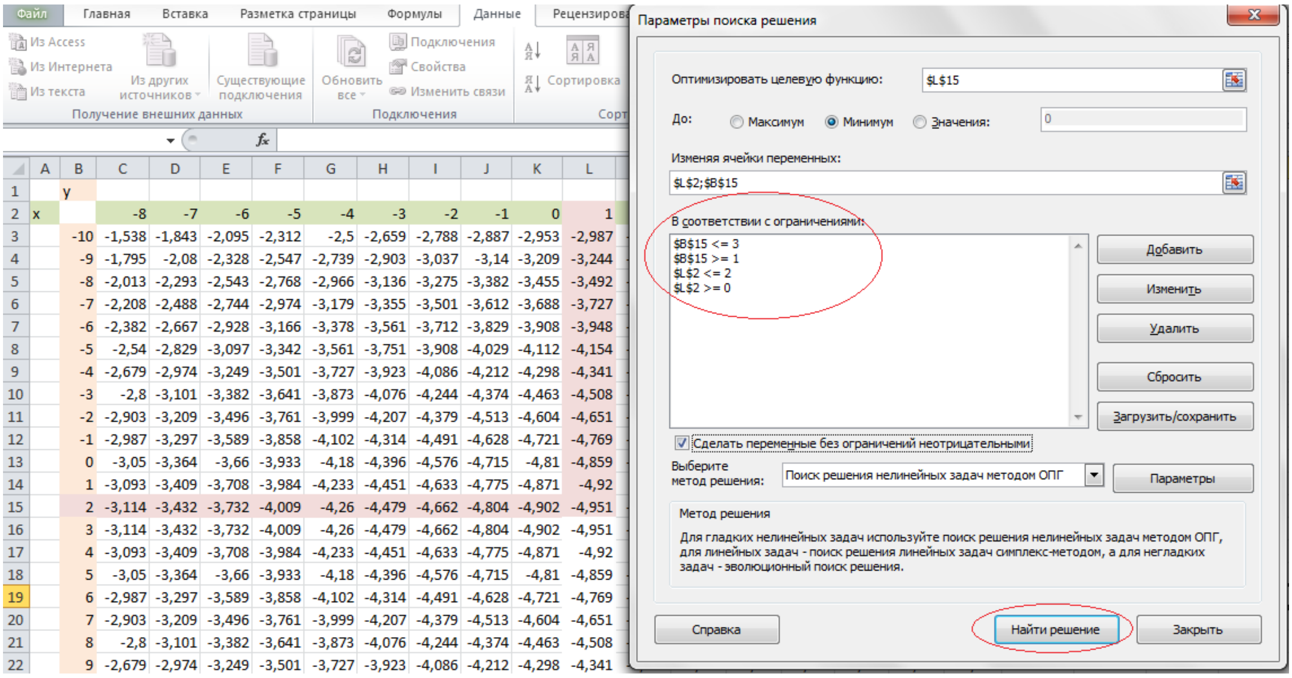


Рис. 26. Задание на уточнение положения минимума функции с применением процедуры «Поиск решения» при наличии ограничения значений аргументов

Появляется панель (рис. 27) с резолюцией о том, что решение найдено и все ограничения удовлетворены. Результаты решения такие же, как и на рис. 22. Если затребовать отчет о выводе результатов (на рис. 27 обведено красным), то на отдельной странице EXCEL-файла получим отчет, отличающийся от отчета, приведенного на рис. 24 только наличием ограничений и резолюцией об их выполнении (рис. 28).

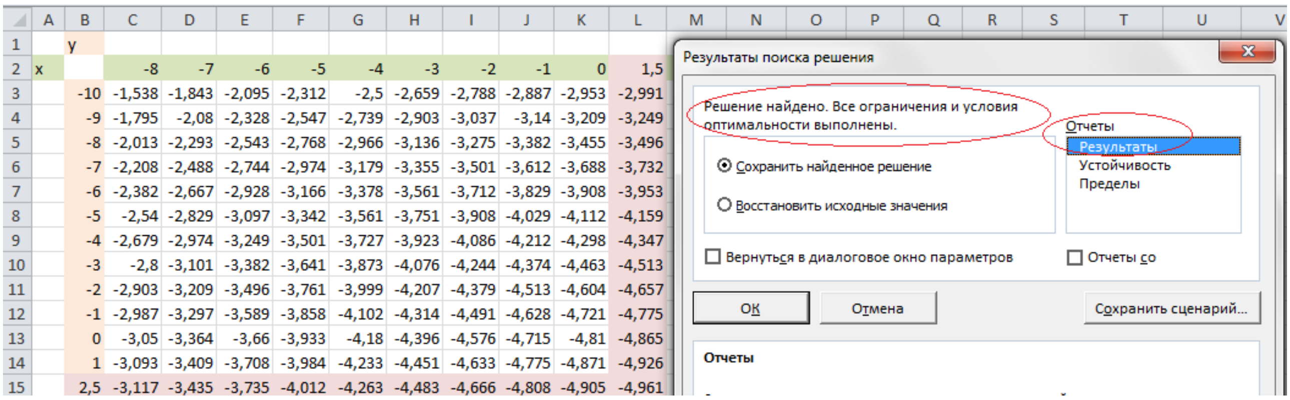


Рис. 27. Результаты поиска задачи с ограничениями решения

Ограничения

Ячейка	Имя	Значение ячейки	Формула	Состояние	Допуск
B15	y	2,499999997	B15<=3	Без привязки	0,500000003
B15	y	2,499999997	B15>=1	Без привязки	1,499999997
L2	x	1,500000001	L2<=2	Без привязки	0,499999999
L2	x	1,500000001	L2>=0	Без привязки	1,500000001

Рис. 28. Фрагмент отчета о результатах удовлетворения ограничений

3. ПРИМЕР НАХОЖДЕНИЯ ЭКСТРЕМУМА ФУНКЦИИ ДВУХ ПЕРЕМЕННЫХ В ПРОГРАММЕ MATHCAD

Определим экстремум функции двух переменных (2) с помощью программы MATHCAD.

Построение диаграммы (поверхности) и локализация начального приближения точки экстремума. Для этого запрограммируем функцию (2) и построим диаграмму поверхности, образуемой функцией в координатах *xy* (рис. 29), выбрав тип диаграммы (**График поверхности**). После построения диаграммы поверхности, ее можно поворачивать для того, чтобы точнее определить координаты точки минимума. По графику определяем приближенные значения координат точки минимума *x*=1; *y*=1.5 (на рис. 30 эти параметры названы Xmin Ymin и выделены красным). Напоминаем, что в программе MATHCAD разделителем между целой и дробной частью числа является точка.

Програмирование функции двух переменных и построение диаграммы

$$z(x,y) := -e^{\operatorname{acos}\left(\frac{x^2-3x}{200}+\frac{y^2-5y}{320}\right)}$$

Рис. 29. Задание функции двух переменных и построение ее графика

Програмирование функции двух переменных и построение диаграммы

$$z(x,y) := -e^{\operatorname{acos}\left(\frac{x^2-3x}{200}+\frac{y^2-5y}{320}\right)}$$

$Xmin := 1$ $Ymin := 1.5$

Уточнение положения точки минимума функции z(x,y)

$TOL = 1 \times 10^{-3}$

$TOL := 10^{-7}$

$TOL = 1 \times 10^{-7}$

$Min := minimize(z, Xmin, Ymin)$

$$Min = \begin{pmatrix} 1.5 \\ 2.5 \end{pmatrix}$$

$Xmin := Min_0 = 1.5$

$Ymin := Min_1 = 2.5$

$z(Xmin, Ymin) = -4.961$

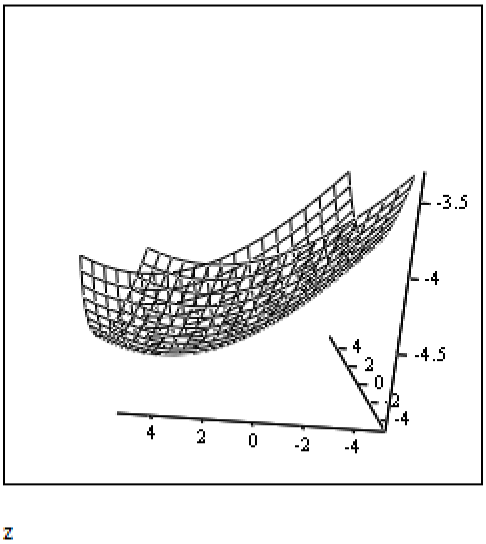


Рис. 30. Уточнение положения точки минимума функции двух переменных

Этап уточнения положения точек экстремума. В программе MATHCAD имеются две отдельные процедуры для нахождения точек экстремумов. Для функции двух переменных эти процедуры записываются следующим образом: «**minimize(f,x,y)**» уточняет положение минимума; «**maximize(f,x,y)**» уточняет положение максимума. Аргументами этих процедур являются: имя функции и начальные приближения аргументов. В результате итерационного процесса будет найден вектор, компонентами которого являются уточненные значения аргументов.). На рис. 30 показано применение процедуры «**minimize(z,Xmin,Ymin)**» для уточнения положения точки минимума функции (2). В результате уточнения стало Xmin=1.5, Ymin=2.5. Значение функции стало равным -4.961. То есть результаты совпали с результатами, полученными в программе EXCEL.

Точность решения этой задачи определяется параметром «TOL». По умолчанию этот параметр в программе MATHCAD равен 10^{-3} (на рис. 30 первый раз показано то значение параметра «TOL», какое оно установлено по умолчанию). Его можно переопределить. Так на рис. 30 второй раз показано, как переопределяется это значение, а третий раз проверяется, каким стало значение параметра «TOL».

Иногда при реализации итерационной процедуры для разрывной функции возможно попадание очередного приближения в область разрыва. Для недопущения такого нежелательного «аварийного» завершения итерационного процесса, следует назначать ограничения. В этом случае ограничения и процедура нахождения точки экстремума должны быть рассмотрены как система уравнений. Чтобы MATHCAD анализировал несколько уравнений как систему, эти уравнениия должны быть записаны после параметра **Given**. В рассматриваемом примере такого «срыва в пропасть» не произошло. Тем не менее, для этой функции покажем (рис. 31), как формулируется задание для избежания такого «аварийного» завершения итерационного процесса.

Програмирование функции двух переменных и построение диаграммы

$$z(x,y) := -e^{\operatorname{acos}\left(\frac{x^2-3x}{200}+\frac{y^2-5y}{320}\right)}$$

$Xmin := 1 \qquad Ymin := 1.5$

Уточнение положения точки минимума функции z(x,y)

$TOL := 10^{-7}$

Given

$-8 \le Xmin \le 8$

$-10 \le Ymin \le 10$

$Min := minimize(z, Xmin, Ymin)$

$$Min = \begin{pmatrix} 1.5 \\ 2.5 \end{pmatrix}$$

$Xmin := Min_0 = 1.5$

$Ymin := Min_1 = 2.5$

$z(Xmin, Ymin) = -4.961$

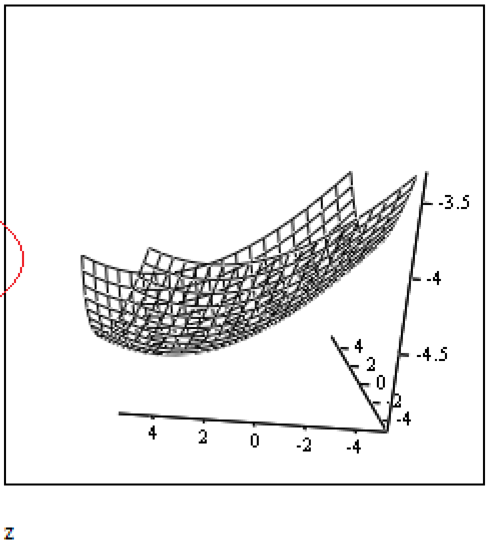


Рис. 31. Уточнение положения точки минимума функции двух переменных с указанием ограничений значений аргументов

4. ЗАДАНИЕ К РАБОТЕ

В соответствии с приведенными теоретическими сведениями и приведенными примерами определить экстремумы заданной функции двух переменных.

5. ПОРЯДОК ВЫПОЛНЕНИЯ РАБОТЫ

1. Уяснить основные положения теории, изложенной в разделе 2: формулировку задачи нахождения экстремумов функций двух аргументов и основы численных методов решения этих задач.

2. С помощью программы EXCEL найти экстремум функции двух переменных (функцию выдает преподаватель непосредственно перед началом занятия):

- Провести табулирование функции *z(x,y)* на заданных интервалах определения аргументов. Шаг табуляции *h* выбирать таким, чтобы было от 10 до 20 точек по каждому аргументу. Основная задача этого и следующего этапа - диаграмма поверхности должна получиться информативной и наглядным. Оформить таблицу (рамки, названия столбцов и т.п.).
- На основе табличных данных построить диаграмму функции *z(x,y)*.
- По таблице значений и по диаграмме поверхности определить приближенные значения координат экстремума функции *z(x,y)* (начальное приближение). Выделить в таблице цветом строку и столбец с начальным приближением (оттенки цвета должны быть очень светлые).
- С помощью процедуры «Поиск решения» определить уточненные значения экстремума функции *z(x,y)*. Точность реализации этого этапа можно настроить, используя меню «Параметры». Результат записать с точностью не менее 5 знаков после запятой.

3. В программе MATHCAD найти экстремум этой же функции *z(x,y)* двух переменной.:

- Запрограммировать функцию *z(x,y)* и построить диаграмму поверхности.
- По диаграмме поверхности определить начальные приближения экстремума функции *z(x,y)*.
- Использовать процедуры «minimize» и «maximize» для отыскания минимума или максимума.

4. Сравнить результаты, полученные в EXCEL и MATHCAD, и сформулировать выводы об эффективности EXCEL и MATHCAD при решении задач нахождения корней и экстремумов функции.

6. ФОРМА ОТЧЕТНОСТИ

Отчёт о лабораторной работе не оформляется. Преподавателю предъявляются результаты работы на мониторе компьютера. Устно формулируются выводы по результатам работы.

7. ВОПРОСЫ ДЛЯ САМОКОНТРОЛЯ

1. Что такое экстремум функции одной и двух переменных?
2. Как взаимосвязаны значения функции, ее первой и второй производных в точках экстремума?
3. Какие экстремумы бывают?
4. Какие этапы решения задачи отыскания экстремумов можно выделить?
5. Какова цель этапа локализации (отделения) различных экстремумов и зачем этот этап необходимо выполнять?
6. Что такое «начальное приближение» точки экстремума и как его выбирать?
7. На каком принципе основаны итерационные методы уточнения положения точек экстремумов?
8. Объяснить основы простейшего итерационного метода уточнения положения экстремума функции одной переменной – метода половинного деления.
9. Какие группы методов итерационного уточнения выделяются и по какому критерию?
10. Объяснить действия, выполняемые процедурой «Поиск решения», имеющейся в программе EXCEL.
11. Объяснить действия, выполняемые процедурами «minimize» и «maximize», имеющимися в программе MATHCAD.

СПИСОК РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ

1. Трифонов, А.Г. Постановка задачи оптимизации и численные методы ее решения"/А.Г.Трифонов, А.Г. [ЭР] Математика\Optimization Toolbox
2. Трубников, С.В. Вычислительная математика: учеб. пособие / С.В.Трубников, Б.В. Порошин. – Брянск: БГТУ, 2005. – 396 с.
3. Порошин, Б.В. Вычислительная математика: Сборник заданий /Б.В.Порошин. – Брянск: БГТУ, 2007. – 104 с.

4. Веденеева, Е.А. Функции и формулы Excel 2007/ Е.А.Веденеева. – СПб.: Питер, 2008. – 384 с.

5. Понятный самоучитель Excel 2010. / В.. Волков – С. П.: Питер, - 2010. – 252 с. – режим доступа: rsload.net›knigi/8325…samouchitel-excel-2010.html

6. Индейкин, В.В. Табличный редактор Microsoft Excel: учебное пособие/ В.В. Индейкин. М., 2011. – 75 с.

7. Самоучитель Microsoft Office Excel 2010 - М.: ID COMPANY, 2010 - режим доступа: aleka.org.ua›soft…samouchitel…excel-2010.html

8. Охорзин, В.А. Прикладная математика в системе MATHCAD: учебное пособие. – СПб.: Лань, 2009. – 352 с.

9. Кирьянов, Д.В. MATHCAD 14/ Д.В.Кирьянов. СПб.: ИРМ, 2007. – 704 с.

10. Шушкевич, Г.Ч. Введение в MATHCAD: учебное пособие/ Г.Ч. Шушкевич, С.В.Шушкевич. М., 2012. – 138 с.

11. Воскобойников, Ю.Е.Основы вычислений и программирования в пакете MATHCAD/ Ю.Е.Воскобойников, А.Ф.Задорожный, Л.А.Литвинов, Ю.Г.Черный. – Новосибирск, Новосибирский архитектурно-строительный университет, 2012 – 212 с.

12. Программирование в среде MathCAD: учеб.-метод. пособие для бакалавров инженерных и физических специальностей / сост. В. К. Толстых. – Донецк: ДонНУ, 2010. – 128 с.

Информатика (информационные технологии). Нахождение экстремума функции двух переменных средствами программ EXCEL и MATHCAD: [Текст]+[Электронный ресурс]: метод. указания к выполнению лабораторной работы № ?? для студентов очной формы всех технических специальностей (квалификация «бакалавр»).

МИХАИЛ ВИКТОРОВИЧ ЗЕРНИН

Научный редактор Л.И.Пугач
Редактор издательства Л.Н. Мажугина
Компьютерный набор М.В.Зернин

Темплан 2014 г., п. .

Подписано в печать Формат 60×84 1/16 Бумага офсетная. Офсетная печать Усл. печ. л. 2,5 Уч.-изд. л. 2,5 . Тираж экз. Заказ Бесплатно

Издательство Брянского государственного технического университета. 241035, Брянск, бульвар им. 50-летия Октября, 7, тел. 58-82-49
Лаборатория оперативной полиграфии БГТУ, ул. Институтская, 16