

Министерство образования и науки
Российской Федерации

Федеральное государственное бюджетное образовательное
учреждение высшего профессионального образования
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
СИСТЕМ УПРАВЛЕНИЯ И РАДИОЭЛЕКТРОНИКИ (ТУСУР)

КОЛИЧЕСТВЕННЫЕ МЕТОДЫ В ЭКОНОМИЧЕСКИХ ИССЛЕДОВАНИЯХ
ЧАСТЬ 1. КЛАССИЧЕСКИЕ МЕТОДЫ ОПТИМИЗАЦИИ
Методические указания к практическим и самостоятельным работам

2018

Содержание

Введение.....	3
1. Постановка задачи.....	3
2. Градиентные методы.....	5
2.1 Градиентный спуск.....	5
2.2 Наискорейший спуск.....	6
2.3 Покоординатный спуск.....	7
2.4 Метод Гаусса-Зейделя.....	7
2.5 Сходимость градиентных методов.....	8
3. Овражные методы.....	9
3.1 Первый овражный метод.....	9
3.2 Второй овражный метод.....	10
3.3 Выбор параметров овражных методов.....	11
Приложение 1.....	13
Приложение 2.....	20
Литература.....	24

Введение

Данная работа представляет первую часть комплекта учебно-методических пособий по дисциплине «Количественные методы в экономических исследованиях» и содержит некоторые методы **классической безусловной оптимизации**, имеющие широкое применение в различных областях науки и практики, в том числе в экономике. Достоинством этих методов является их относительная простота, в том числе для программной реализации. А также то обстоятельство, что они включены в качестве стандартных процедур в различные пакеты прикладных программ: EXCEL, MATHCAD, MATLAB и др.

Это учебно-методическое пособие предназначено для выполнения самостоятельных и лабораторных работ по указанной дисциплине.

1. Постановка задачи

Рассматривается задача минимизации (максимизации) функции $f(X)$, $X = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}$ при

отсутствии ограничений на вектор независимых переменных X (задача безусловной оптимизации). Один из способов безусловной оптимизации, обоснованный в рамках курса математического анализа, – это использование необходимых и достаточных условий экстремума функции [1].

Напомним некоторые определения. Частная производная функции $f(X)$ по переменной x_i обозначается $\frac{\partial f}{\partial x_i}$. Главная (линейная) часть приращения функции называется *дифференциалом* функции и обозначается $df(X)$. По определению,

$df(x_1, \dots, x_n) = \sum_{i=1}^n \frac{\partial f(x_1, \dots, x_n)}{\partial x_i} dx_i$, где dx_i - дифференциалы (приращения) независимых переменных.

Вектор в пространстве E^n , координатами которого являются частные производные функции $f(X)$ в некоторой точке X_0 , называется *градиентом* функции в данной точке и

обозначается $grad f(X_0)$ или $\nabla f(X_0)$. Таким образом, $\nabla f(X_0) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \Big|_{X_0} \\ \dots \\ \frac{\partial f}{\partial x_n} \Big|_{X_0} \end{bmatrix}$.

Вектор градиента функции в точке определяет величину и направление скорости наибольшего роста функции в данной точке, т.е. представляет собой наибольшую из всех

производных по направлению: $\frac{\partial f}{\partial l} = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \cos \alpha_i$. Здесь $\cos \alpha_i, i = \overline{1, n}$ - направляющие

косинусы заданного направления l .

Напомним, что уравнение $f(X) = const$ определяет *поверхность уровня* (поверхность постоянного значения целевой функции). Можно показать, что вектор градиента в каждой точке ортогонален соответствующей поверхности уровня.

Особое значение при исследовании функции на экстремум имеет матрица, составленная из вторых частных производных функции $f(X)$ (матрица Гессе):

$$A = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}, i = \overline{1, n}; j = \overline{1, n} \right].$$

С учетом введенных обозначений сформулируем необходимые и достаточные условия существования экстремума функции в точке.

Необходимые условия:

$$df(X_0) = 0, \text{ или } \left. \frac{\partial f}{\partial x_i} \right|_{X_0} = 0, i = \overline{1, n}, \text{ или } \nabla f(X_0) = 0 \quad (1)$$

Достаточные условия:

Пусть функция $f(X)$ дважды непрерывно дифференцируема в окрестности $G(X_0) \subset E^n$. В точке X_0 выполнены необходимые условия. Если, кроме того, *положительно (отрицательно)* определена квадратичная форма

$$A(dx_1, \dots, dx_n) = \sum_{i=1}^n \sum_{j=1}^n \left. \frac{\partial^2 f}{\partial x_i \partial x_j} \right|_{X_0} dx_i dx_j, \quad (2)$$

то функция $f(X)$ в точке X_0 имеет *минимум (максимум)*.

Если квадратичная форма является *знакопеременной*, то экстремума нет. Если квадратичная форма равна *нулю*, то требуются дополнительные исследования. Знаковая определенность квадратичной формы обычно исследуется по критерию Сильвестра. Для этого рассматриваются главные миноры матрицы Гессе. Если

$$|a_{11}| > 0, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \dots, \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} > 0, \text{ то квадратичная форма (2) положительно}$$

$$\text{определена. Если } |a_{11}| < 0, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \dots, (-1)^n \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} > 0, \text{ то (2) - отрицательно}$$

определена. В остальных случаях знаковая определенность отсутствует.

Условия (1) представляют собой систему уравнений, решение которой является самостоятельной, в общем случае достаточно трудоемкой, задачей. Исследование главных миноров на знак также является, как правило, нетривиальной проблемой. Поэтому широкое развитие получили различные приближенные процедуры поиска экстремумов.

Будем искать минимум целевой функции $f(X)$. Пусть удалось построить последовательность точек $X[1], X[2], \dots, X[k]$, такую, что $f(X[1]) > f(X[2]) > \dots > f(X[k])$. Можно надеяться, что при $k \rightarrow \infty$ удастся найти точку минимума, соответствующим образом выбрав условие останова. Последовательность точек $\{X[k]\}$ называется *релаксационной*. Способы построения таких последовательностей называются *методами спуска*. Если при построении релаксационной последовательности используются вторые производные целевой функции, то метод спуска называется методом *второго* порядка, при использовании только первых производных имеем метод *первого* порядка. Если в алгоритмах минимизации не используется необходимое условие существования экстремума, то методы называются *прямыми*. Если в процессе минимизации вычисление производных не производится, речь идет о *поисковых* методах.

Точка $X[k]$ называется точкой k -го (очередного) приближения. Процесс отыскания точки $X[k]$ называется k -той *итерацией*.

2. Градиентные методы [2]

Большинство методов спуска определяется следующей *итерационной* формулой: $X[k+1] = X[k] + t_k P_k$. Здесь: P_k - вектор, указывающий *направление* спуска на k -той итерации; t_k - скаляр, задающий *величину шага* вдоль выбранного направления. Методы, в которых направление спуска P_k определяется вектором градиента или его компонентами, будем называть *градиентными* методами.

2.1. Градиентный спуск

Этот метод использует в качестве направления спуска на каждой итерации вектор антиградиента функции в точке очередного приближения. Поскольку вектор градиента функции указывает направление наибольшего возрастания в точке, то вектор антиградиента указывает направление наибольшего убывания функции в данной точке. Поэтому для отыскания минимума целесообразно спускаться именно в этом направлении.

Исходные данные: точка начального приближения $X[0]$; $t = const$ - величина шага; общее количество итераций N ; точность решения задачи ε .

Итерационная формула имеет вид: $X[k+1] = X[k] - t \cdot \nabla f(X[k])$.

Условие останова: $\|\nabla f(X[k])\| < \varepsilon$ или выполнено N итераций.

Главная проблема метода заключается в определении величины шага, с которым будет производиться спуск. Шаг может оказаться слишком большим, тогда процесс начнет расходиться. Если же шаг задать заведомо очень маленьким, то отыскание минимума приведет к чрезмерным временным затратам. Конкретные рекомендации можно дать только по отношению к некоторым классам функции.

Пример. Определить условия сходимости метода градиентного спуска для функции $f(x) = \alpha(x-c)^2, \alpha > 0, x$ - скаляр.

Очевидно, что $x_{min} = c$. Запишем итерационную формулу, учитывая, что

$$\nabla f = f'_x = 2\alpha(x-c):$$

$$x[k+1] = x[k] - t \cdot 2\alpha(x[k] - c);$$

$$x[k+1] - c = (x[k] - c) - t \cdot 2\alpha(x[k] - c) = (x[k] - c)(1 - 2\alpha t);$$

$$\frac{x[k+1] - c}{x[k] - c} = 1 - 2\alpha t.$$

Условие сходимости можно определить следующим образом: $\left| \frac{x[k+1] - c}{x[k] - c} \right| = |1 - 2\alpha t| < 1$.

Решив последнее неравенство, получим: при $0 < t < \frac{1}{\alpha}$ процесс сходится, при $t = \frac{1}{\alpha}$ -

«зацикливается», при $t > \frac{1}{\alpha}$ - расходится.

2.2. Наискорейший спуск

В отличие от предыдущего метода, требующего задавать величину шага спуска вручную перед началом работы, данный алгоритм использует автоматический выбор шага на каждой итерации.

Исходные данные: точка начального приближения $X[0]$; общее количество итераций N ; точность решения задачи ε .

Итерационная формула: $X[k+1] = X[k] - t_k \cdot \nabla f(X[k])$.

Шаг t_k определяется из условия:

$$\varphi(t_k) = \min_{t>0} \varphi(t) = \min_{t>0} f(X[k] - t \nabla f(X[k])).$$

При этом для отыскания t_k можно использовать необходимое условие (1): $\varphi'(t) = 0$ (3).

Условие останова: $\|\nabla f(X[k])\| < \varepsilon$ или выполнено N итераций.

Из условия выбора шага в методе наискорейшего спуска следует, что направления спуска на соседних итерациях ортогональны. Для доказательства достаточно в уравнении (3) записать $\varphi'(t)$ как производную сложной функции многих переменных.

2.3. Метод покоординатного спуска

Исходные данные: точка начального приближения $X[0]$, точность ε , максимальное количество итераций N .

Итерационная формула имеет вид:

$$X[kn + s + 1] = X[kn + s] - t_{kn+s} \cdot \left. \frac{\partial f}{\partial x_{s+1}} \right|_{X[kn+s]} \cdot e^{s+1}.$$

Здесь: n - размерность задачи (количество независимых переменных); $k=0,1,\dots$ - номер *внешней* итерации; $s=0,1,\dots,n-1$ - номер *внутренней* итерации; e^s - вектор, компоненты

которого удовлетворяют условию $e_j^s = \begin{cases} 1, & j = s \\ 0, & j \neq s \end{cases}$; t_{kn+s} - шаг. Если $t_{kn+s} = t = const$, то имеем

покоординатный спуск с постоянным шагом t , который должен быть задан перед началом работы метода. Суть метода заключается в том, что каждый раз спуск осуществляется только по *одной* переменной (внутренняя итерация). После того как пройдены *все* переменные (внешняя итерация), происходит возврат к первой.

Условие останова: $\|\nabla f(X[k])\| < \varepsilon$ или выполнено N итераций.

2.4. Метод Гаусса-Зейделя

Исходные данные: точка начального приближения $X[0]$, точность ε , максимальное количество итераций N .

Итерационная формула имеет вид:

$$X[kn + s + 1] = X[kn + s] - t_{kn+s} \cdot \left. \frac{\partial f}{\partial x_{s+1}} \right|_{X[kn+s]} \cdot e^{s+1}.$$

Здесь: n - размерность задачи; k - номер *внешней* итерации; $s=0,1,\dots,n-1$ - номер *внутренней* итерации; e^s - вектор, компоненты которого удовлетворяют условию

$e_j^s = \begin{cases} 1, & j = s \\ 0, & j \neq s \end{cases}$. Шаг t_{kn+s} определяется из условия

$$\varphi(t_{kn+s}) = \min_{t>0} f \left(X[kn+s] - t \frac{\partial f(X[kn+s])}{\partial x_{s+1}} \cdot e^{s+1} \right)$$

Условие останова: $\|\nabla f(X[k])\| < \varepsilon$ или выполнено N итераций.

Введем понятие *сепарабельной* функции как функции, представимой в виде:

$f(x_1, \dots, x_n) = \sum_{i=1}^n \psi_i(x_i)$, где ψ_i - некоторая функция одной переменной. Легко убедиться, что

метод Гаусса-Зейделя минимизирует сепарабельную функцию n переменных ровно за n шагов.

2.5. Сходимость градиентных методов

Доказано [3], что градиентные методы сходятся (при выполнении ряда условий) со скоростью геометрической прогрессии со знаменателем $q = \frac{M-m}{M+m}$. Здесь M и m - равномерные по x оценки наибольшего и наименьшего собственных чисел матрицы Гессе соответственно. Если $M \approx m$, то матрица Гессе является *хорошо обусловленной*. Поверхности уровня такой функции представляют собой семейство сфер. В этом случае $q \ll 1$, следовательно, градиентные методы сходятся хорошо. Если же $M \gg m$, то матрица Гессе *плохо обусловлена*. В этом случае поверхности уровня минимизируемой функции имеют структуру, сильно отличающуюся от сферической. Они оказываются вытянутыми вдоль некоторой поверхности, называемой «*дном оврага*». Функции, обладающие таким свойством, называются *овражными*. Приведем примеры овражных функций.

$$1. f(x_1, x_2) = \frac{1}{2}x_1^2 + 50x_2^2.$$

$$\nabla f = \begin{bmatrix} x_1 \\ 100x_2 \end{bmatrix}, \quad A = \begin{bmatrix} 1 & 0 \\ 0 & 100 \end{bmatrix}.$$

$$\det(A - \lambda E) = (1 - \lambda)(100 - \lambda) = 0. \quad m = 1, M = 100.$$

$$\frac{m}{M} = 0.01, q = \frac{99}{101} \approx 1.$$

$$2. f(x_1, x_2) = \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 + 50(x_1 - x_2)^2.$$

$$\nabla f = \begin{bmatrix} x_1 + 100(x_1 - x_2) \\ x_2 - 100(x_1 - x_2) \end{bmatrix}, \quad A = \begin{bmatrix} 101 & -100 \\ -100 & 101 \end{bmatrix}.$$

$$\det(A - \lambda E) = (101 - \lambda)^2 - 10000 = 0. \quad m = 1, M = 201.$$

$$\frac{m}{M} = \frac{1}{201}, q = \frac{200}{202} \approx 1.$$

Для лучшего представления особенностей овражных функций рекомендуем нарисовать линии уровня в приведенных примерах.

Для минимизации овражных функций необходимо применять специально ориентированные методы, учитывающие указанную особенность матрицы Гессе. В частности, градиентные методы в данном случае сходятся очень медленно, т.к. коэффициент $q \approx 1$. Стандартные методы минимизации позволяют попасть лишь в одну из точек на дне оврага, где норма вектора градиента может быть существенно меньше, чем в других точках области определения функции. Однако полученная точка, как правило, расположена далеко от истинной точки минимума, в которой значение овражной функции во много раз меньше, чем почти всюду на дне оврага.

Величины приращения координат в результате шага градиентного или покоординатного спуска пропорциональны значениям соответствующих частных производных, которые в случае овражной функции существенно отличаются друг от друга. При этом условие сходимости (верхнюю границу величины шага) определяют «быстрые» переменные (большие производные), а скорость сходимости определяется «медленными» переменными (маленькие производные). В табл.1 и на рис.1а и 1б приведены примеры траекторий градиентного спуска для разных величин шага. Явно предпочтительнее в смысле скорости приближения к точке минимума выглядит траектория, имеющая колебательный характер и представленная на рис.1б.

В общем случае объективным признаком правильного выбора величины шага является сходящийся процесс с немонотонным изменением хотя бы одной «быстрой» переменной. Для проверки этого условия достаточно взглянуть на результаты нескольких первых шагов.

Заметим, что разбиение переменных на «быстрые» и «медленные» зависит от точки, из которой выполняется текущий шаг.

В табл.2 и на рис.2 приведен пример траектории наискорейшего спуска. Хорошо видны характерные особенности этого метода: ортогональность соседних перемещений и тот факт, что на каждом шаге траектория спуска пересекает дно оврага.

В табл. 3 приведены результаты работы методов градиентного и наискорейшего спуска для одной и той же функции и одинаковых исходных данных. Шаг градиентного спуска выбирался так, чтобы последовательность значений функции была **релаксационной**.

2 Овражные методы [2]

Низкая эффективность градиентных методов при минимизации овражных функций вызвала создание специально ориентированных на решение таких задач эвристических процедур, реализующих, как правило, разделение движения по «быстрым» и «медленным» переменным. От этих процедур не следует ожидать отыскание решения с высокой точностью. Их задача заключается в том, чтобы за относительно небольшое количество шагов хорошо улучшить точку приближения к решению задачи для последующего использования этой точки другим методом.

2.1. Первый овражный метод (метод Гельфанда-Цетлина)

Метод состоит из двух этапов: на первом этапе производится спуск на «дно оврага», на втором - движение вдоль «дна оврага» по направлению к точке минимума. При этом целесообразно разделить независимые переменные на *быстрые* и *медленные* и

использовать их для спуска на первом и втором этапах соответственно. Для разделения используются так называемые *параметры этапов*. Очевидно также, что движение ко «дну оврага» и вдоль «дна оврага» нужно осуществлять с разными шагами. Таким образом, **исходными данными** метода являются: $\varepsilon_1, \varepsilon_2$ - параметры первого и второго этапов соответственно; t_1, t_2 - шаги первого и второго этапов; N_1, N_2 - количество шагов на каждом из этапов. Кроме того, как обычно, должны быть заданы: точка начального приближения $X[0]$; общее количество итераций N ; точность решения задачи ε .

Итерационная формула: $X[k+1] = X[k] - t \cdot G[k]$, $G[k] = \begin{bmatrix} g_1[k] \\ \dots \\ g_n[k] \end{bmatrix}$.

Первый этап. $t = t_1$, $g_i[k] = \begin{cases} \left. \frac{\partial f}{\partial x_i} \right|_{X[k]}, \left. \frac{\partial f}{\partial x_i} \right|_{X[k]} \geq \varepsilon_1; \\ 0, \text{ в противном случае} \end{cases}$

Условие останова на первом этапе: $G[k] = 0$ или выполнено N_1 шагов.

Второй этап. $t = t_2$, $g_i[k] = \begin{cases} \left. \frac{\partial f}{\partial x_i} \right|_{X[k]}, \left. \frac{\partial f}{\partial x_i} \right|_{X[k]} \leq \varepsilon_2; \\ 0, \text{ в противном случае} \end{cases}$

Условие останова на втором этапе: $G[k] = 0$ или выполнено N_2 шагов.

Общее условие останова: $\|\nabla f(X[k])\| < \varepsilon$ или выполнено N итераций.

2.2. Второй овражный метод (метод большого овражного шага)

Исходные данные: точка начального приближения $X[0]$; $t = const$ - величина овражного шага; общее количество итераций N , точность решения задачи ε .

Суть метода заключается в следующем: в окрестности точки очередного приближения $X[k]$ выбирают точку $\tilde{X}[k]$. Спуском из точек $X[k]$ и $\tilde{X}[k]$ находят две точки на «дне оврага» X' и X'' , с использованием которых «дно оврага» аппроксимируют прямой линией. Вдоль полученной прямой делают шаг в направлении убывания функции и получают точку следующего приближения:

$$X[k+1] = X' - t \cdot (X'' - X') \cdot \frac{f(X'') - f(X')}{\|X'' - X'\|^2}.$$

Останов, как правило, по заданному количеству итераций. Однако на практике при работе с методом большого овражного шага процесс останавливают в том случае, когда несколько раз будет зафиксировано движение с резким увеличением целевой функции.

2.3. Выбор параметров овражных методов

Начнем с первого овражного метода. Задачей этапа 1 является спуск на дно оврага. Этот факт можно считать установленным, если на первых шагах этапа имеют место значительные изменения «быстрых» переменных, а на последних шагах их изменения практически отсутствуют. Этап 2 должен обеспечить движение к точке минимума вдоль дна оврага за счет ограниченной расходимости по «быстрым» переменным.

В табл.4 и на рис. 3 приведена траектория первого овражного метода для функции двух переменных, которые четко определены как «быстрая» и «медленная». В табл. 5 и на рис. 4 изображена траектория спуска для функции, обе переменные которой являются «быстрыми» и обе становятся «медленными» на дне оврага. Хорошо виден спуск на дно оврага на этапе 1 и движение вдоль дна оврага на этапе 2.

Все действия алгоритма ориентированы на наличие оврага. Если овраг отсутствует, то результат этих действий может оказаться отличным от ожидаемого. Это имеет место в непосредственной близости от точки минимума и приводит к тому, что после выполнения очередного этапа 1 фиксируется большее (чем на предыдущем этапе 1) значение функции. Такое положение означает, что алгоритм уже выполнил свою задачу и получил точку, близкую к точке минимума. Работу алгоритма можно остановить.

Основная сложность при использовании первого овражного метода заключается в выборе его параметров. Неудачное определение параметров может привести к вырождению его в обычный градиентный или покоординатный спуск. Исходя из изложенного, можно предложить следующую последовательность действий при настройке первого овражного метода.

- а) Задать $\varepsilon_1 = 0$, ε_2 – максимально возможное значение.
- б) Выбором параметров t_1 , N_1 обеспечить спуск на дно оврага. При этом фиксируется разбиение переменных на «быстрые» и «медленные».
- в) Увеличением ε_1 попытаться остановить движение по «медленным» переменным.
- г) Выбором t_2 , N_2 обеспечить хорошее изменение «медленных» переменных, не обращая внимание на расходимость по «быстрым» переменным.
- д) Уменьшением ε_2 остановить расходимость по «быстрым» на нужном уровне.

Настройка метода овражного шага заключается в выборе единственного параметра – величины овражного шага t . Из приведенных выше рассуждений следует, что признаком хорошего выбора величины шага может служить значительное уменьшение функции на нескольких первых итерациях с последующим возрастанием функции. В табл. 6 и на рис.5 изображена траектория метода овражного шага. Для наглядности результаты первых трех

шагов и последующих шагов изображены на разных рисунках в разном масштабе. Видно, что увеличение функции (табл. 6) имеет место после того, как алгоритм «побывал» в непосредственной близости к точке минимума.

Приложение 1.

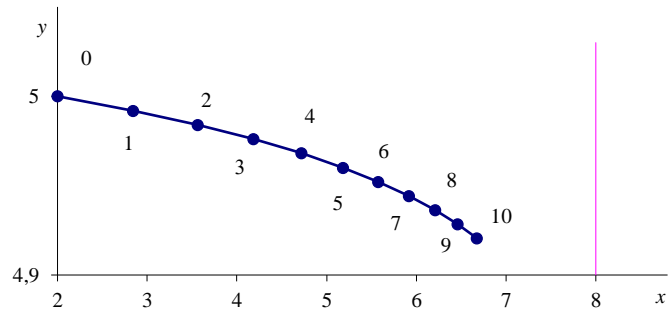
Рис. 1. Градиентный спуск (вариант А): 1а - $t=0,001$; 1б - $t=0,01225$.

Рис. 1а

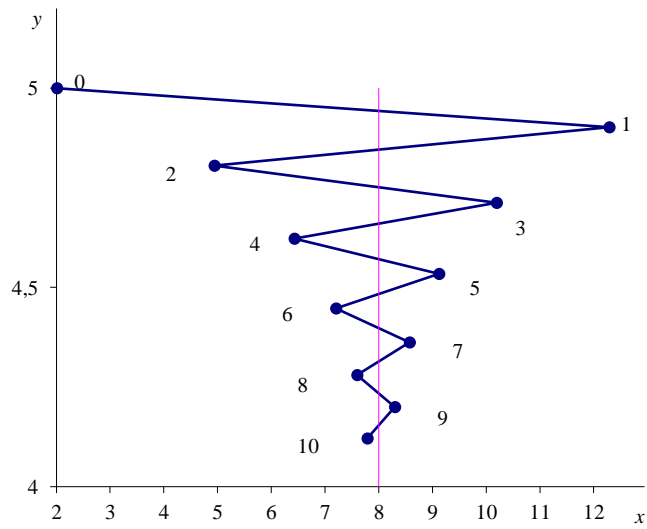


Рис. 1б

Рис. 2. Наискорейший спуск (вариант В): 2а - итерации 0-2; 2б - итерации 2-10.

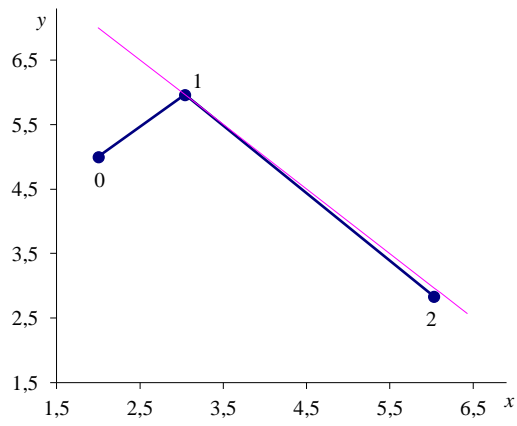


Рис.2а.

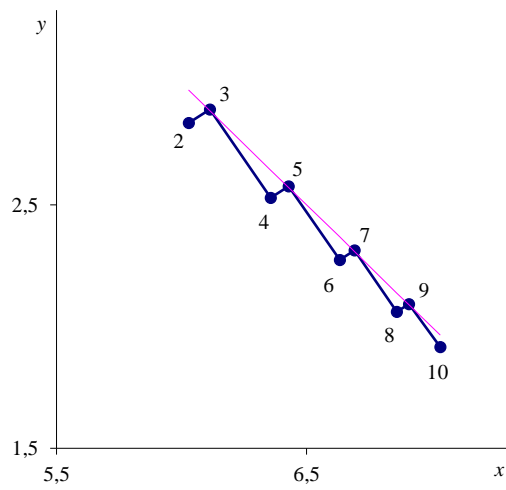


Рис.2б.

Рис. 3. Первые два цикла первого овражного метода для варианта А.

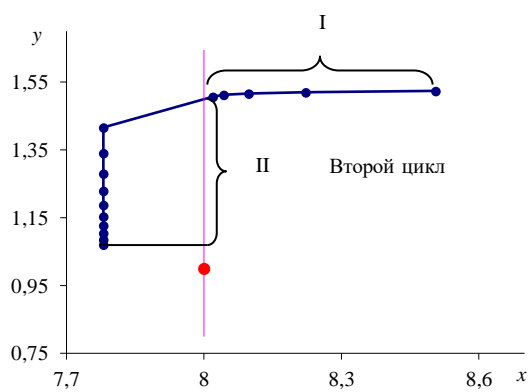
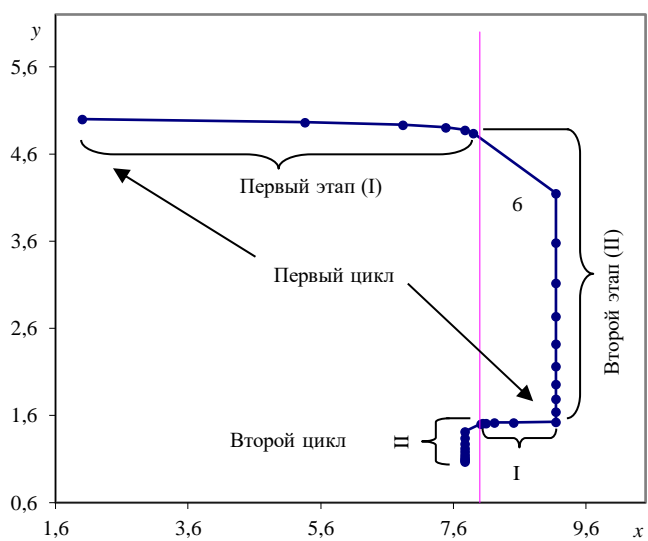


Рис. 4а. Первые два цикла первого овражного метода для варианта В.

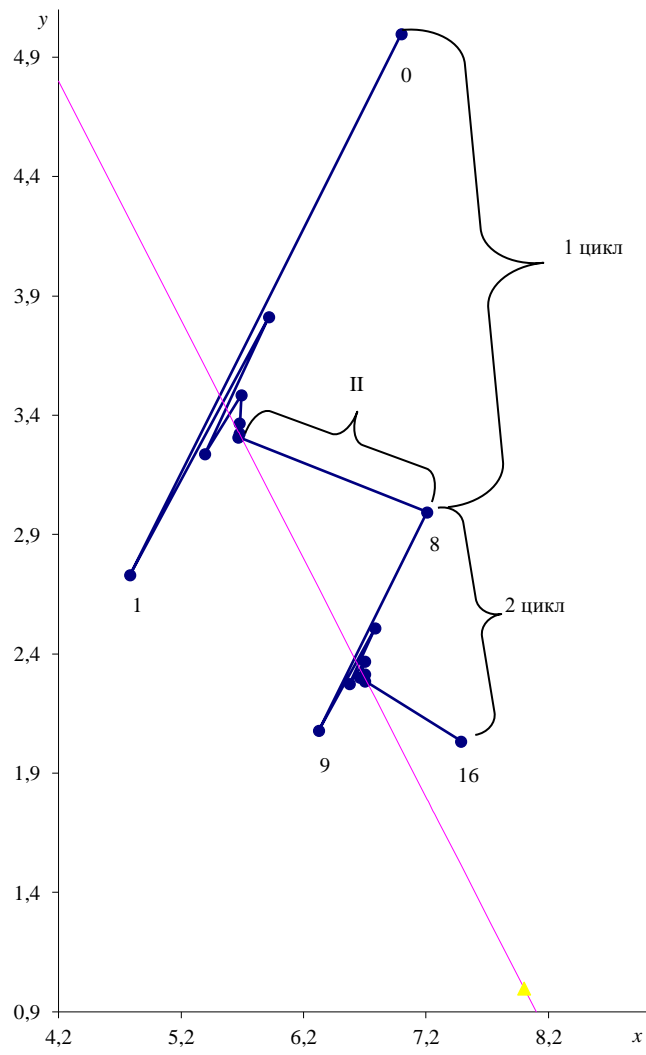


Рис.4б. Второй цикл первого овражного метода для варианта В.

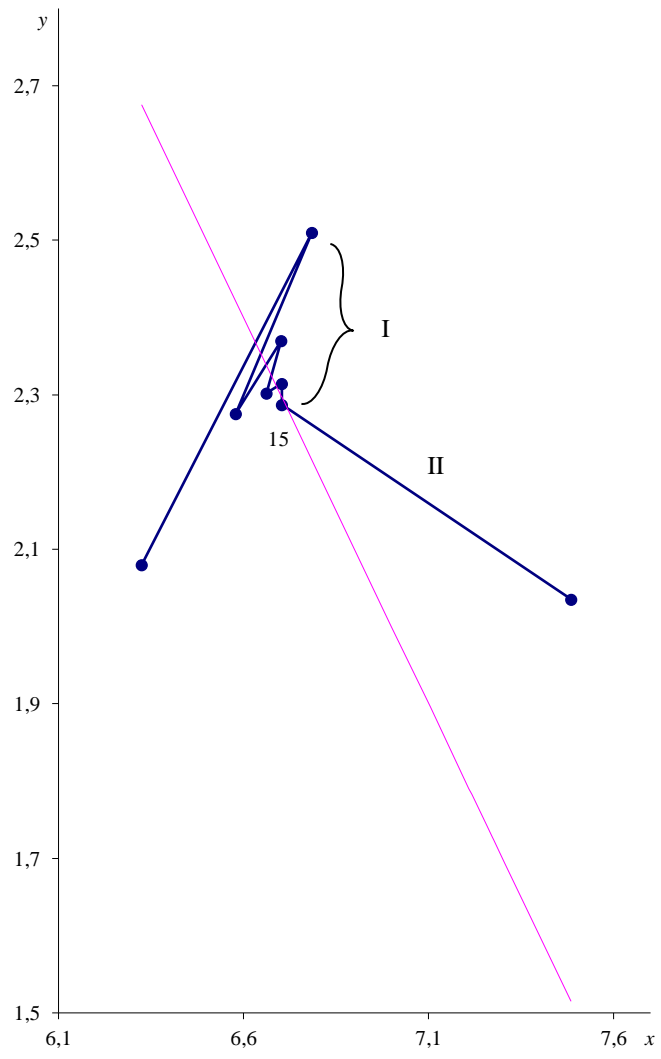


Рис. 5а. Второй овражный метод (вариант С): итерации 0-3.

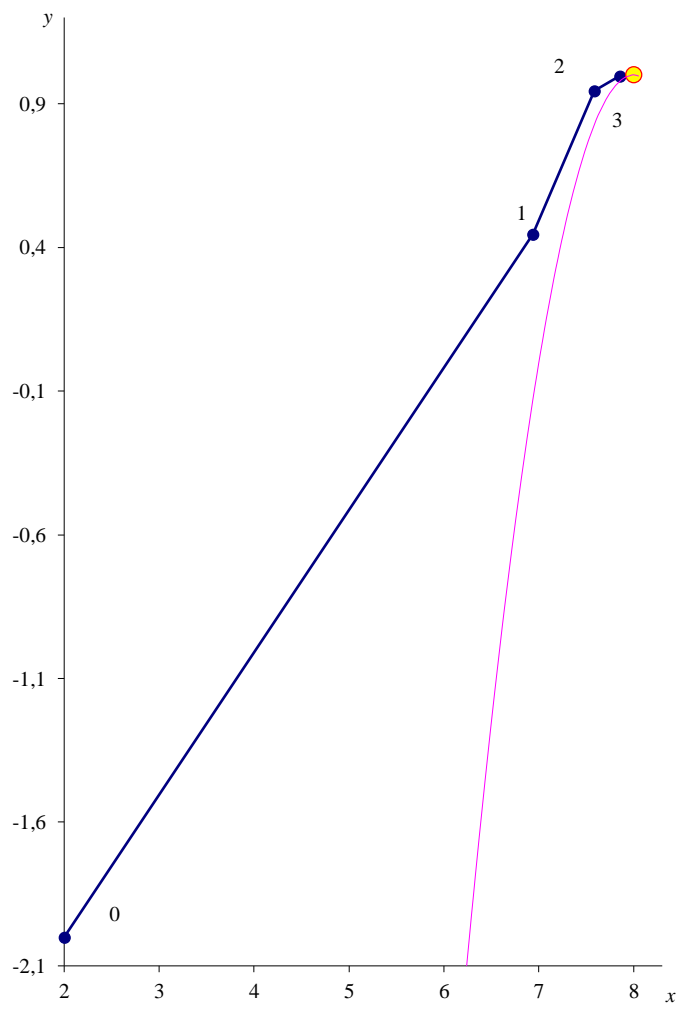
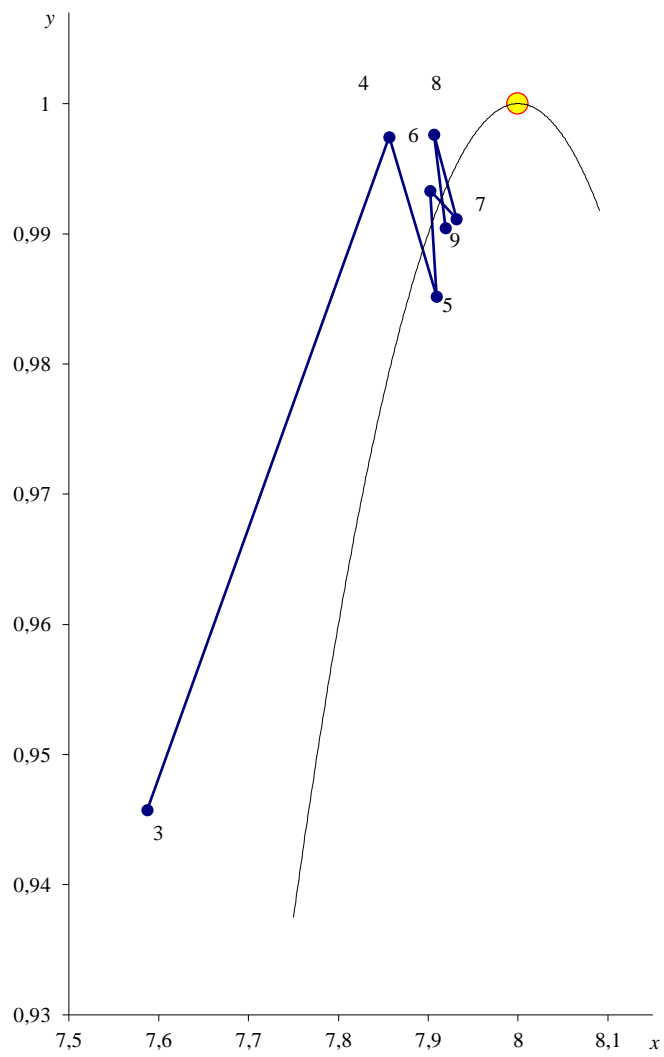


Рис. 5б. Второй овражный метод (вариант С): итерации 3-9.



Приложение 2

Табл.1. Траектория спуска метода градиентного спуска для функции

$$f(x, y) = 70(x - 8)^2 + (y - 1)^2 + 1 \quad (\text{вариант А}).$$

Точка минимума: $(8, 1)$.

Уравнение дна оврага: $x = 8$.

	Итерация	x	y	$f(x, y)$
$t=0,01225$	0	2	5	2537
	1	12.29	4.902	1304.51
	2	4.93265	4.8064	674.09
	3	10.19316	4.71314	351.48
	4	6.43189	4.62217	186.25
	5	9.1212	4.53343	101.48
	6	7.19835	4.44686	57.87
	7	8.57318	4.36241	35.30
	8	7.59017	4.28003	23.52
	9	8.29303	4.19967	17.25
	10	7.79049	4.12128	13.81
$t=0,001$	0	2	5	2537
	1	2.84	4.992	1880.73
	2	3.5624	4.984	1395.33
	3	4.1837	4.976	1036.32
	4	4.7179	4.9681	770.77
	5	5.1774	4.9602	574.36
	6	5.5726	4.9522	429.08
	7	5.9124	4.9443	321.61
	8	6.2047	4.9364	242.11
	9	6.456	4.9286	183.30
	10	6.6722	4.9207	139.79

Табл.2. Траектория спуска метода наискорейшего спуска для функции

$$f(x, y) = (x - 8)^2 + (y - 1)^2 + 70(x + y - 9)^2 + 1 \text{ (вариант В).}$$

Точка минимума: (8,1). Уравнение дна оврага: $y = 9 - x$.

Итерация	x	y	$f(x, y)$
0	2	5	333
1	3.0373	5.9663	50.29
2	6.0289	2.836	9.53
3	6.1131	2.8921	8.14
4	6.3578	2.5297	6.92
5	6.4278	2.5765	5.96
6	6.6326	2.2737	5.11
7	6.6909	2.3127	4.44
8	6.8617	2.0603	3.85
9	6.9103	2.0927	3.38
10	7.0339	1.9172	2.94

Табл.3. Сравнение траекторий спуска для функции

$$f(x, y) = (x - 8)^2 + (y - 1)^2 + 70(y + (x - 8)^2 - 1)^2 + 1 \text{ (вариант С).}$$

Точка минимума: (8,1). Уравнение дна оврага: $y = 1 - (x - 8)^2$.

Итерац ии	Градиентный спуск $t = 0.00938$			Наискорейший спуск		
	x	y	$f(x, y)$	x	y	$f(x, y)$
0	6	-2	84	6	-2	84
1	6.52903	-2.12569	77.71	6.25183	-2.05984	13.42
2	6.16016	-1.99351	24.08	6.47234	-1.17808	9.77
3	6.35279	-2.03930	20.39	6.52142	-1.19037	7.99
4	6.21484	-1.99079	15.82	6.60843	-0.83794	6.99
5	6.31009	-2.01092	14.61	6.64278	-0.84639	6.25
6	6.24441	-1.98490	13.65	6.70074	-0.61173	5.69
7	6.29252	-1.99206	13.28	6.72944	-0.61878	5.24
8	6.26138	-1.97639	13.03	6.77285	-0.44316	4.86
9	6.28583	-1.97690	12.90	6.79800	-0.44933	4.55
10	6.27170	-1.96626	12.82	6.83397	-0.30483	4.27

Табл.4. Траектория спуска первого овражного метода для функции

$$f(x, y) = 70(x - 8)^2 + (y - 1)^2 + 1 \quad (\text{вариант А}).$$

Точка минимума: $(8, 1)$.Уравнение дна оврага: $x = 8$.

		Итерац ии	x	y	$f(x, y)$
Первый цикл	Первый этап $t_1 = 0,004$ $\varepsilon_1 = 0,5$ $N_1 = 5$	0	2	5	2537
		1	5.36	4.968	504.62
		2	6.8384	4.93626	110.95
		3	7.4889	4.90477	34.53
		4	7.77511	4.87353	19.54
	Второй этап $t_2 = 0,15$ $\varepsilon_2 = 20$ $N_2 = 10$	5	7.90105	4.84254	16.45
		6	9.14781	4.15058	103.15
		7	9.14781	3.58372	99.899
		8	9.14781	3.11865	97.71
		9	9.14781	2.7373	96.24
		10	9.14781	2.42458	95.25
		11	9.14781	2.16816	94.59
		12	9.14781	1.95789	94.14
		13	9.14781	1.79547	93.84
		14	9.14781	1.64408	93.64
Второй цикл	Первый этап $t_1 = 0,004$ $\varepsilon_1 = 0,5$ $N_1 = 5$	15	9.14781	1.528149	93.50
		16	8.50504	1.52392	19.13
		17	8.2222	1.51973	4.73
		18	8.09778	1.51557	1.94
		19	8.04302	1.51145	1.39
	Второй этап $t_2 = 0,15$ $\varepsilon_2 = 20$ $N_2 = 10$	20	8.0193	1.50736	1.28
		21	7.78042	1.41603	4.55
		22	7.78042	1.34115	4.49
		23	7.78042	1.27974	4.45
		24	7.78042	1.22939	4.43
		25	7.78042	1.1881	4.41
		26	7.78042	1.15424	4.40
		27	7.78042	1.126664	4.39
		28	7.78042	1.10371	4.386
		29	7.78042	1.08504	4.382
		30	7.78042	1.06974	4.380

Табл.5. Траектория спуска первого овражного метода для функции

$$f(x, y) = (x-8)^2 + (y-1)^2 + 70(x+y-9)^2 + 1 \text{ (вариант В).}$$

Точка минимума: $(8, 1)$. Уравнение дна оврага: $y = 9 - x$.

		Итерац и	x	y	f(x, y)
Первый цикл	Первый этап $t_1 = 0,0053$ $\varepsilon_1 = 0,5$ $N_1 = 7$	0	7	5	648
		1	4.7846	2.7316	168.45
		2	5.91966	3.81422	50.95
		3	5.39717	3.23985	20.01
		4	5.69409	3.48544	14.75
		5	5.67532	3.32588	12.79
		6	5.67681	3.3674	12.14
	7	5.66884	3.30943	11.80	
	Второй этап $t_2 = 0,2; \varepsilon_2 = 70$ $N_2 = 5$	8	7.20952	2.99388	106.97
Второй цикл	Первый этап $t_1 = 0,0053$ $\varepsilon_1 = 0,5$ $N_1 = 7$	9	6.32498	2.07982	29.79
		10	6.78437	2.51001	10.83
		11	6.57882	2.27557	6.13
		12	6.70193	2.37009	4.92
		13	6.66225	2.30213	4.57
		14	6.70286	2.31476	4.43
		15	6.70286	2.28775	4.35
		Второй этап $t_2 = 0,2; \varepsilon_2 = 70$ $N_2 = 5$	16	7.48467	2.0356
Третий цикл	Первый этап $t_1 = 0,0053$ $\varepsilon_1 = 0,5$ $N_1 = 7$	17	7.10409	1.63858	6.86
		18	7.30452	1.82275	3.30
		19	7.21746	1.71959	2.41
		20	7.27246	1.75867	2.17
		21	7.25707	1.72753	2.10
		22	7.27637	1.73124	2.06
		23	7.27637	1.71784	2.04
		Второй этап $t_2 = 0,2; \varepsilon_2 = 70$ $N_2 = 5$	24	7.72783	1.59271
		25	-1.13863	-7.6197	22216.5

Табл.6. Траектория спуска второго овражного метода для функции

$$f(x, y) = (x-8)^2 + (y-1)^2 + 70(y + (x-8)^2 - 1)^2 + 1 \text{ (вариант С).}$$

Точка минимума: $(8, 1)$.

Уравнение дна оврага: $y = 1 - (x-8)^2$.

Шаг метода: $t = 0.4$

Итерации	x	y	$f(x, y)$
0	2	-2	76276
1	6.93885	0.44528	25.28
2	7.58732	0.94577	2.11636
3	7.85623	0.99744	1.04363
4	7.909	0.98519	1.01149
5	7.90177	0.99331	1.01030
6	7.93153	0.99114	1.00598
7	7.90631	0.99764	1.01167
8	7.91925	0.99047	1.00724
9	7.90533	0.99599	1.01069
10	7.92173	0.99102	1.00678

Литература

1. Фихтенгольц Г.М. Курс дифференциального и интегрального исчисления. – М.: Наука, 1975. – 608 с.
2. Моисеев.Н.Н., Иванилов Ю.П., Столярова Е.М. Методы оптимизации. – М.: Наука, 1975.
3. Пантелеев А.В., Летова Т.А. Методы оптимизации в примерах и задачах. – М.: Высшая школа, 2002. – 544 с.

Приложение

Табл.1. Траектория спуска метода градиентного спуска для функции

$$f(x, y) = 70(x - 8)^2 + (y - 1)^2 + 1 \quad (\text{вариант А}).$$

Точка минимума: $(8, 1)$.

Уравнение дна оврага: $x = 8$.

	Итераци	x	y	$f(x, y)$
	я			
$t=0,01225$	0	2	5	2537
	1	12.29	4.902	1304.51
	2	4.93265	4.8064	674.09
	3	10.1931	4.71314	351.48
	4	6.43189	4.62217	186.25
	5	9.1212	4.53343	101.48
	6	7.19835	4.44686	57.87
	7	8.57318	4.36241	35.30
	8	7.59017	4.28003	23.52
	9	8.29303	4.19967	17.25
	10	7.79049	4.12128	13.81
$t=0,001$	0	2	5	2537
	1	2.84	4.992	1880.73
	2	3.5624	4.984	1395.33
	3	4.1837	4.976	1036.32
	4	4.7179	4.9681	770.77
	5	5.1774	4.9602	574.36
	6	5.5726	4.9522	429.08
	7	5.9124	4.9443	321.61
	8	6.2047	4.9364	242.11
	9	6.456	4.9286	183.30
	10	6.6722	4.9207	139.79

Табл.2. Траектория спуска метода наискорейшего спуска для функции

$$f(x, y) = (x-8)^2 + (y-1)^2 + 70(x+y-9)^2 + 1 \text{ (вариант В).}$$

Точка минимума: $(8, 1)$. Уравнение дна оврага: $y = 9 - x$.

Итерация	x	y	$f(x, y)$
0	2	5	333
1	3.0373	5.9663	50.29
2	6.0289	2.836	9.53
3	6.1131	2.8921	8.14
4	6.3578	2.5297	6.92
5	6.4278	2.5765	5.96
6	6.6326	2.2737	5.11
7	6.6909	2.3127	4.44
8	6.8617	2.0603	3.85
9	6.9103	2.0927	3.38
10	7.0339	1.9172	2.94

Табл.3. Сравнение траекторий спуска для функции

$$f(x, y) = (x - 8)^2 + (y - 1)^2 + 70(y + (x - 8)^2 - 1)^2 + 1 \text{ (вариант С).}$$

Точка минимума: $(8, 1)$. Уравнение дна оврага: $y = 1 - (x - 8)^2$.

Итерац ии	Градиентный спуск $t = 0.00938$			Наискорейший спуск		
	x	y	$f(x, y)$	x	y	$f(x, y)$
0	6	-2	84	6	-2	84
1	6.529 03	- 2.1256 9	77.71	6.25183	-2.05984	13.42
2	6.160 16	- 1.9935 1	24.08	6.47234	-1.17808	9.77
3	6.352 79	- 2.0393 0	20.39	6.52142	-1.19037	7.99
4	6.214 84	- 1.9907 9	15.82	6.60843	-0.83794	6.99
5	6.310 09	- 2.0109 2	14.61	6.64278	-0.84639	6.25
6	6.244 41	- 1.9849 0	13.65	6.70074	-0.61173	5.69
7	6.292 52	- 1.9920 6	13.28	6.72944	-0.61878	5.24
8	6.261 38	- 1.9763 9	13.03	6.77285	-0.44316	4.86
9	6.285 83	- 1.9769 0	12.90	6.79800	-0.44933	4.55

10	6.271 70	- 1.9662 6	12.82	6.83397	-0.30483	4.27
----	-------------	------------------	-------	---------	----------	------

Табл.4. Траектория спуска первого овражного метода для функции

$$f(x, y) = 70(x-8)^2 + (y-1)^2 + 1 \quad (\text{вариант А}).$$

Точка минимума: $(8, 1)$.

Уравнение дна оврага: $x = 8$.

		Итерац ии	x	y	$f(x, y)$
Первый цикл	Первый этап $t_1 = 0,004$ $\varepsilon_1 = 0,5$ $N_1 = 5$	0	2	5	2537
		1	5.36	4.968	504.62
		2	6.8384	4.93626	110.95
		3	7.4889	4.90477	34.53
		4	7.77511	4.87353	19.54
	5	7.90105	4.84254	16.45	
	Второй этап $t_2 = 0,15$ $\varepsilon_2 = 20$ $N_2 = 10$	6	9.14781	4.15058	103.15
		7	9.14781	3.58372	99.899
		8	9.14781	3.11865	97.71
		9	9.14781	2.7373	96.24
		10	9.14781	2.42458	95.25
		11	9.14781	2.16816	94.59
		12	9.14781	1.95789	94.14
		13	9.14781	1.79547	93.84
		14	9.14781	1.64408	93.64
15		9.14781	1.52814	93.50	
Второй цикл	Первый этап $t_1 = 0,004$ $\varepsilon_1 = 0,5$ $N_1 = 5$	16	8.50504	1.52392	19.13
		17	8.2222	1.51973	4.73
		18	8.09778	1.51557	1.94
		19	8.04302	1.51145	1.39
		20	8.0193	1.50736	1.28
	Второй этап $t_2 = 0,15$ $\varepsilon_2 = 20$ $N_2 = 10$	21	7.78042	1.41603	4.55
		22	7.78042	1.34115	4.49
		23	7.78042	1.27974	4.45
		24	7.78042	1.22939	4.43
		25	7.78042	1.1881	4.41
		26	7.78042	1.15424	4.40
		27	7.78042	1.12666	4.39
		28	7.78042	1.10371	4.386
		29	7.78042	1.08504	4.382
		30	7.78042	1.06974	4.380

Табл.5. Траектория спуска первого овражного метода для функции

$$f(x, y) = (x-8)^2 + (y-1)^2 + 70(x+y-9)^2 + 1 \quad (\text{вариант В}).$$

Точка минимума: $(8, 1)$. Уравнение дна оврага: $y = 9 - x$.

		Итерац и	x	y	$f(x, y)$
Первый цикл	Первый этап $t_1 = 0,0053$ $\varepsilon_1 = 0,5$ $N_1 = 7$	0	7	5	648
		1	4.7846	2.7316	168.45
		2	5.91966	3.81422	50.95
		3	5.39717	3.23985	20.01
		4	5.69409	3.48544	14.75
		5	5.67532	3.32588	12.79
		6	5.67681	3.3674	12.14
	7	5.66884	3.30943	11.80	
	Второй этап $t_2 = 0,2; \varepsilon_2 = 70$ $N_2 = 5$	8	7.20952	2.99388	106.97
Второй цикл	Первый этап $t_1 = 0,0053$ $\varepsilon_1 = 0,5$ $N_1 = 7$	9	6.32498	2.07982	29.79
		10	6.78437	2.51001	10.83
		11	6.57882	2.27557	6.13
		12	6.70193	2.37009	4.92
		13	6.66225	2.30213	4.57
		14	6.70286	2.31476	4.43
		15	6.70286	2.28775	4.35
		Второй этап $t_2 = 0,2; \varepsilon_2 = 70$ $N_2 = 5$	16	7.48467	2.0356
Третий цикл	Первый этап $t_1 = 0,0053$ $\varepsilon_1 = 0,5$ $N_1 = 7$	17	7.10409	1.63858	6.86
		18	7.30452	1.82275	3.30
		19	7.21746	1.71959	2.41
		20	7.27246	1.75867	2.17
		21	7.25707	1.72753	2.10
		22	7.27637	1.73124	2.06
		23	7.27637	1.71784	2.04
		Второй этап $t_2 = 0,2; \varepsilon_2 = 70$ $N_2 = 5$	24	7.72783	1.59271
		25	-1.13863	-7.6197	22216.5

Табл.6. Траектория спуска второго овражного метода для функции

$$f(x, y) = (x - 8)^2 + (y - 1)^2 + 70(y + (x - 8)^2 - 1)^2 + 1 \text{ (вариант С).}$$

Точка минимума: $(8, 1)$.

Уравнение дна оврага: $y = 1 - (x - 8)^2$.

Шаг метода: $t = 0.4$

Итерации	x	y	$f(x, y)$
0	2	-2	76276
1	6.93885	0.44528	25.28
2	7.58732	0.94577	2.11636
3	7.85623	0.99744	1.04363
4	7.909	0.98519	1.01149
5	7.90177	0.99331	1.01030
6	7.93153	0.99114	1.00598
7	7.90631	0.99764	1.01167
8	7.91925	0.99047	1.00724
9	7.90533	0.99599	1.01069
10	7.92173	0.99102	1.00678

Примечание. Приводимые ниже рисунки иллюстрируют данные соответствующих таблиц, как указано в тексте. Пунктирной линией на каждом рисунке изображено «дно оврага». На некоторых рисунках отдельно отмечена точка минимума.