

# ЛАБОРАТОРНАЯ РАБОТА №1

## ИССЛЕДОВАНИЕ ИТЕРАТИВНЫХ МЕТОДОВ ЧИСЛЕННОЙ ОПТИМИЗАЦИИ

### Цель работы

Ознакомление с методами поиска экстремума нелинейной выпуклой функции нескольких переменных методом градиентного спуска и методом имитации отжига.

### Описание методов

#### Метод градиентного спуска.

Задача состоит в отыскании минимума функции двух переменных  $f(x,y)$  (следует отметить, что если необходимо найти максимум некоторой функции  $F(x,y)$ , то эта задача сводится к поиску минимума функции  $f(x,y)=-F(x,y)$  ).

Большинство численных методов состоит в отыскании некоторой последовательности  $(x_0, y_0), (x_1, y_1), \dots, (x_k, y_k)$ , которая при  $k \rightarrow \infty$  (или при  $k \rightarrow k_M$ ) сходится к точке минимума  $(x^*, y^*)$ . Если при этом выполняется  $f(x_0, y_0) > f(x_1, y_1) > \dots > f(x_k, y_k)$ , то есть значения функции монотонно убывают при увеличении  $k$ , то такой метод называется методом спуска.

Известно, что вектор градиента функции

$$\overline{\text{grad}} f(x, y) = \left( \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right)$$

направлен в сторону наибольшего возрастания функции  $f(x,y)$ . Поэтому в качестве направления движения можно принять противоположное градиенту направление (антиградиент), т.е. координаты точек пересчитываются по формулам

$$\begin{aligned} x_{k+1} &= x_k - \alpha_k \frac{\partial f(x_k, y_k)}{\partial x}, \\ y_{k+1} &= y_k - \alpha_k \frac{\partial f(x_k, y_k)}{\partial y}. \end{aligned}$$

Выбор величины  $\alpha_k$ , с которой связана длина  $k$ -го шага, в общем случае является сложной задачей. Если  $\alpha_k$  мало, то движение будет слишком медленным и потребует значительного объема вычислений. Если  $\alpha_k$  велико, то существует возможность перескочить точку

минимума и выйти на противоположный склон функции. При этом возможно нарушение требования монотонного убывания последовательности  $f(x_k, y_k)$  и появляется опасность заикливания, то есть колебания последовательности  $(x_k, y_k)$  в некоторой окрестности точки минимума  $(x^*, y^*)$  без приближения к ней.

Существует несколько различных способов выбора  $\alpha_k$ . В данном примере рассматривается разновидность метода с дроблением шага. Для этого задается начальное приближение  $(x_0, y_0)$  и начальное значение  $\alpha_0$  (например,  $x_0=y_0=0$ ,  $\alpha_0=1$ ). Вычисление  $x_1, y_1$  и всех последующих  $x_{k+1}, y_{k+1}$  производится по вышеуказанным формулам. При этом если окажется, что  $f(x_{k+1}, y_{k+1}) > f(x_k, y_k)$ , то величина  $\alpha_k$  уменьшается в два раза и вычисление  $x_{k+1}, y_{k+1}$  повторяется от точки  $(x_k, y_k)$  с новым значением  $\alpha_k$ . Если же значение функции убывает, то величина  $\alpha_k = \alpha_{k-1}$ .

Критерием окончания счета принимается неравенство:

$$|\overline{\text{grad}} f(x, y)| = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2} < \varepsilon$$

либо одновременное выполнение двух неравенств

$$\left| \frac{\partial f}{\partial x} \right| < \frac{\sigma}{2}, \quad \left| \frac{\partial f}{\partial y} \right| < \frac{\sigma}{2}$$

### **Алгоритм имитации отжига**

Алгоритм имитации отжига (Simulated Annealing, SA) предложен в 1953г. Метрополисом (N. С. Metropolis). Алгоритм SA можно считать одним из немногих универсальных алгоритмов решения задач глобальной оптимизации. Алгоритм вдохновлён механизмом исправления дефектов в кристаллической решётке металла и некоторых других веществ. Указанные выше дефекты в кристаллической решётке обусловлены тем, что некоторые атомы занимают в решётке «неправильные» места. В силу нехватки кинетической энергии при нормальной температуре, указанные атомы не могут преодолеть потенциальный барьер и занять «правильные» положения в решётке. При этом в целом вся система атомов находится в состоянии локального энергетического минимума. Для вывода системы из этого минимума и перевода в состояние глобального энергетического минимума, соответствующего бездефектной кристаллической решётке, металл нагревают до высокой температуры, а затем медленно охлаждают. При этом «неправильные» атомы могут приобрести энергию, достаточную для преодоления потенциального барьера и занятия «правильных» положений в решётке. Вероятность преодоления потенциального барьера пропорциональна температуре металла, так что по мере его охлаждения вероятность такого перехода стремится к нулю, и кристаллическая решётка стабилизируется в окрестности нового (меньшего) локального или

глобального минимума.

Таким образом, положительная роль повышения температуры отжигаемого металла заключается в том, что температурные флуктуации позволяют системе покидать локальные минимумы энергии и смещаться в сторону более глубоких энергетических минимумов.

Применение данной схемы к оптимизации основано на том, что локальное (субоптимальное) решение, найденное в процессе решения задачи оптимизации, также можно рассматривать как дефектное решение. Улучшить это решение (приблизиться к глобальному оптимуму) можно путём его случайных флуктуаций, амплитуда которых уменьшается с ростом номера итераций. Принципиальным в алгоритме SA является то, что, в отличие от большинства других стохастических алгоритмов поисковой оптимизации, он допускает шаги, приводящие к увеличению значений фитнес-функции. Алгоритм SA относится к классу так называемых пороговых стохастических алгоритмов безусловной оптимизации.

Пороговый алгоритм в процессе поиска допускает ухудшение значений фитнес-функции до заданного порога, и этот порог в процессе итераций последовательно снижается до нуля. В алгоритме SA величина представляет собой случайную величину с математическим ожиданием, равным  $T$ , которому придается смысл «температуры» отжигаемого металла. Таким образом, в алгоритме SA переход от решения к решению допускается с вероятностью:

$$P(\Delta E, T) = \begin{cases} 1, & \text{если } \Delta E \leq 0 \\ e^{-\Delta E/T}, & \text{если } \Delta E > 0 \end{cases}$$

Последняя формула означает, что если переход от точки к точке приводит к уменьшению значения фитнес-функции, то этот переход осуществляется безусловно. В противном случае переход выполняем с вероятностью, которая убывает с ростом разности и уменьшением «температуры». Скорость сходимости алгоритма SA в значительной мере определяет вид функции уменьшения температуры. Поэтому известно большое число различных рекомендаций по выбору этой функции. Наиболее распространенными являются:

- Больцмановский отжиг

При достаточно больших  $T_0$  и количестве шагов, выбор такого семейства распределений гарантирует нахождение глобального минимума.

Закон Больцмана:

$$T = \frac{T_0}{\ln(1 + i)}$$

- Отжиг Коши (быстрый отжиг)

Нет недостатка медленного убывания температуры Больцмановского отжига.

Закон Коши:

$$T = \frac{T_0}{(1 + i)}$$

- Линейная зависимость

Линейный закон:

$$T_{i+1} = \alpha T_i.$$

### Порядок выполнения работы

1. Составить программу минимизации  $f(x,y)$  методом градиентного спуска и алгоритма имитации отжига .
2. Задать входные данные согласно номеру варианта.
3. Провести вычисления, определить минимальные значения исследуемой функции для обоих методов, визуализировать полученные данные.
4. Написать отчет, который должен содержать результаты пунктов 1-3, а также комментарий хода вычислений с объяснениями результатов.

### Методические указания

1. Для метода градиентного спуска минимизировать исследуемую функцию с точностью до  $\sigma=10^{-4}$ . В случае застревания в области локального минимума, попробовать доработать алгоритм и добиться достижения глобального минимума.
2. Для нахождения минимума функции методом имитации отжига согласно следующему примеру:

Реализовать алгоритм имитации отжига следующим образом:

1. Изначально задаются начальная температура ( $T_{max}$ ) и конечная температура ( $T_{min}$ ).

2. Случайно выбирается точка  $x_1$  на отрезке. Вычисляется значение функции в этой точке  $f(x_1)$ .

3. Пока  $T_i > T_{min}$

1) Случайно выбирается точка  $x_i$  на отрезке. Вычисляется значение функции в этой точке  $f(x_i)$ .

1\*) Точка  $x_i$  выбирается исходя из движения в сторону антиградиента от точки  $x_{i-1}$ .

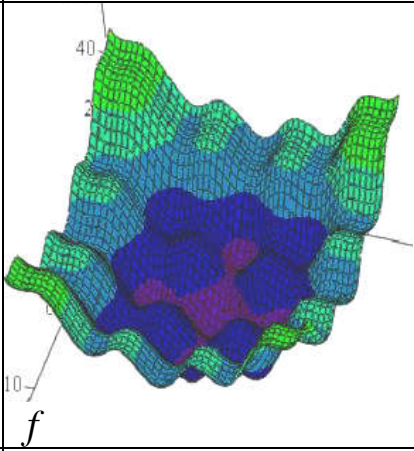
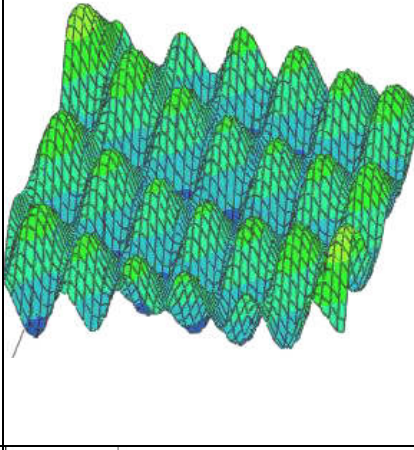
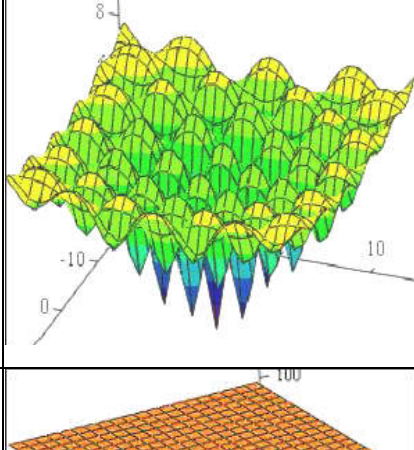
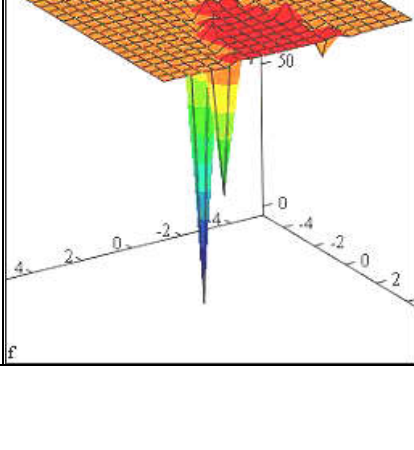
2) Определяется  $\Delta E$ .

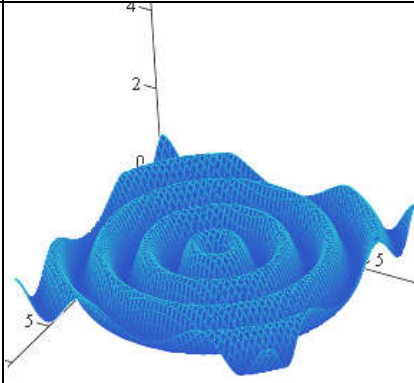
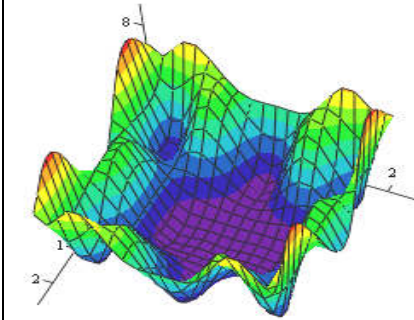
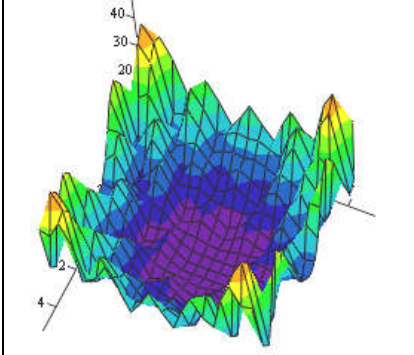
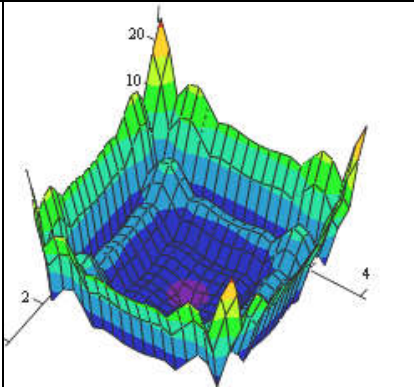
3) Если  $\Delta E \leq 0$ , то осуществляется переход в точку  $x_i$ .

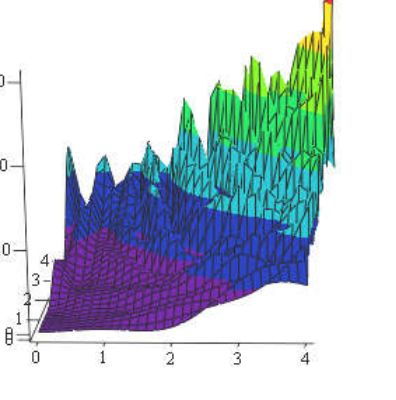
4) Если  $\Delta E > 0$ , то переход осуществляется с вероятностью  $e^{-\Delta E/T}$ .

5) Понижение температуры  $T_{i+1} = f(T_i)$

3. При исследовании алгоритма реализовать в п 3.1 оба варианта выполнения алгоритма, используя одну функцию понижения температуры из предложенных в теоретической справке. Сделать вывод о применимости алгоритма исходя из полученных экспериментальных данных.

Номер ва- рианта	Вид функции	График
1	<p>Функция Растригина:</p> $I(x, y) = 0,1x^2 + 0,1y^2 - 4 \cos(0,8x) - 4 \cos(0,8y) + 8,$ <p><math>x, y \in [-16; 16], \min = I(0,0) = 0</math></p>	
2	<p>Функция Растригина овражная с поворотом осей:</p> $I(x, y) = (0,1K_x \cdot A)^2 + (0,1K_y \cdot B)^2 - 4 \cos(0,8K_x \cdot A) - 4 \cos(0,8K_y \cdot B) + 8,$ <p>где <math>A = x \cos(a) - y \sin(a)</math>;  <math>B = x \sin(a) + y \cos(a)</math>; <math>a</math> – угол поворота,  <math>a = \frac{\pi}{2}</math>; <math>K_x, K_y</math> – растяжение/сжатие по <math>x, y</math>;  <math>K_x = 1,5, K_y = 0,8, x, y \in [-16; 16],</math>  <math>\min = I(0,0) = 0</math></p>	
3	<p>Функция Griewank:</p> $I(x, y) = \frac{-10}{0,005(x^2 + y^2) - \cos(x) \cdot \cos\left(\frac{y}{\sqrt{2}}\right) + 2} + 10,$ <p><math>x_1, x_2 \in [-16; 16], \min = I(0,0)</math></p>	
4	<p>Функция De Jong 2:</p> $F(x, y) = \frac{-100}{100(x^2 - y) + (1 - x)^2} + 100,$ <p><math>x_1, x_2 \in [5; 5], \min = I(1,1) = 0</math></p>	

Номер ва- рианта	Вид функции	График
5	<p>Функция «Сомбреро»:</p> $F(x, y) = \frac{1 - \sin^2(\sqrt{x^2 + y^2})}{1 + 0,001 \cdot (x^2 + y^2)},$ <p><math>x_1, x_2 \in [-10; 10]</math>, <math>\min = I(0,0) = 0</math></p>	
6	<p>Функция Катковника:</p> $I(x_1, x_2) = 0,5(x_1^2 + x_2^2) \times$ $\times [2 \cdot A + A \cos(1,5x_1) \cos(3,14x_2) +$ $+ A \cos(\sqrt{5}x_1) \cos(3,5x_2)],$ <p><math>A = 0,8</math>,  <math>x_1, x_2 \in [-2,5; 2,5]</math>, <math>\min = I(0,0) = 0</math></p>	
7	$I(x_1, x_2) = 0,5(x_1^2 + x_2^2) \times$ $\times [2 \cdot A + A \cos(1,5x_1) \cos(3,14x_2) +$ $+ A \cos(\sqrt{5}x_1) \cos(3,5x_2)]$ <p><math>A = 0,8</math>,  <math>x_1, x_2 \in [-5; 5]</math>, <math>\min = I(0,0) = 0</math></p>	
8	$I(x_1, x_2) = x_1^2  \sin 2x_1  + x_2^2  \sin 2x_2  -$ $- \frac{1}{(5x_1^2 + 5x_2^2 + 0,2)} + 5,$ <p><math>x_1, x_2 \in [-4; 4]</math>, <math>\min = I(0,0) = 0</math></p>	

Номер ва- рианта	Вид функции	График
9	$I(x_1, x_2) = 0,5(x_1^2 + x_1x_2 + x_2^2) \times$ $\times [1 + 0,5\cos(1,5x_1)\cos(3,2x_1x_2)\cos(3,14x_2) +$ $+ 0,5\cos(2,2x_1)\cos(4,8x_1x_2)\cos(3,5x_2)],$ $x_1, x_2 \in [-0; 4], \min = I(0,0) = 0$	
10	<p>Мультипликативная потенциальная функция:</p> $I(x_1, x_2) = -z(x_1)z(x_2),$ $z(x) = -\frac{1}{(x-1)^2 + 0,2} - \frac{1}{2(x-2)^2 + 0,15} -$ $-\frac{1}{3(x-3)^2 + 0,3},$ $x_1, x_2 \in [0; 4], \min = I(2,2) = -60,8$	