Лабораторная работа №5

Оптимизация многомерных функций с помощью эволюционной стратегии

5.1 Теоретическая часть

Цель работы: оптимизация функций многих переменных модификация методом эволюционной стратегии. Графическое отображение результатов оптимизации.

При выполнении лабораторной работы можно использовать следующие источники из прилагаемого списка литературы [1,3,7,11].

Общие сведения

Эволюционные стратегии (ЭС), также как и предыдущие парадигмы, основаны на эволюции популяции потенциальных решений, но, в отличие от них, здесь используется генетические операторы на уровне фенотипа, а не генотипа, как это делается в ГА. Разница в том, что ГА работают в пространстве генотипа – кодов решений, в то время как ЭС производят поиск в пространстве фенотипа – векторном пространстве вещественных чисел. В ЭС учитываются свойства хромосомы «в целом», в отличие от ГА, где при поиске решений исследуются отдельные гены. В природе один ген может одновременно влиять на несколько свойств организма. С другой стороны одно свойство особи может определяться несколькими генами. Естественная эволюция основана на исследовании совокупности генов, а не отдельного (изолированного) гена.

В эволюционных стратегиях целью является движение особей популяции по направлению к лучшей области ландшафта фитнесс-функции. ЭС изначально разработаны для решения многомерных оптимизационных задач, где пространство поиска – многомерное пространство вещественных чисел. Иногда при решении задачи накладываются некоторые ограничения, например, вида $g_i(x)>0$.

Ранние эволюционные стратегии (ЭС) основывались на популяции, состоящей из одной особи, и в них использовался только один генетический оператор – мутация. Здесь для представления особи (потенциального решения) была использована идея, не представленная в классическом генетическом алгоритме, которая заключается в следующем.

Здесь особь представляется парой действительных векторов

$$v = (\bar{x}, \bar{\sigma}), \tag{5.1}$$

где $\bar{x}$ - точка в пространстве решений и $\bar{\sigma}$ - вектор стандартных отклонений (вариабельность) от решения. В общем случае особь популяции определяется вектором потенциального решения и вектором «стратегических параметров» эволюции. Обычно это вектор стандартных отклонений (дисперсия), хотя допускаются (и иногда используются) и другие статистики.

Единственным генетическим оператором в классической ЭС [1] является оператор мутации, который выполняется путем сложения координат вектора-родителя со случайными числами, подчиняющимися закону нормального распределения, следующим образом:

$$\bar{x}^{t+1} = \bar{x}^{t} + N(0, \bar{\sigma}), \tag{5.2}$$

где $N(0,\bar{\sigma})$ - вектор независимых случайных чисел, генерируемых согласно распределению Гаусса (например, табличным способом) с нулевым средним значением и стандартным отклонением σ. Как видно из приведенной формулы величина мутации управляется нетрадиционным способом. Иногда эволюционный процесс используется для изменения и самих стратегических параметров σ, в этом случае величина мутации эволюционирует вместе с искомым потенциальным решением. Это соответствует адаптивному ГА с изменяемым шагом мутации.

Интуитивно ясно, что увеличение отклонения подобно увеличению шага поиска на поверхности ландшафта. Высокая вариабельность способствует расширению пространства поиска и эффективна при нахождении потенциальных зон (суб)оптимальных решений и соответствует высоким значениям коэффициента мутации. В тоже время малые значения вариабельности позволяют сфокусироваться на поиске решения в перспективной области. В данном случае стратегические параметры стохастически определяют величину шага поиска: большая вариабельность ведет к большим шагам. Отметим, что поскольку отклонения генерируются стохастически (по нормальному закону), то большая вариабельность может давать маленький шаг и наоборот. Известно, что 68,26% случайных чисел при нормальном распределении попадают в интервал, определяемый стандартным отклонением σ; 95% чисел попадают в интервал 1,96σ и т.д.

5.1. Двукратная эволюционная (1+1)- стратегия

Здесь потомок принимается в качестве нового члена популяции (он заменяет своего родителя), если значение фитнесс функции (ЦФ) на нем лучше, чем у его родителя и выполняются все ограничения. Иначе, (если значение фитнесс-функции на нем хуже, чем у родителей), потомок уничтожается и популяция остается неизменной.

Рассмотрим выполнение оператора мутации на конкретном примере следующей функции [2]

$$f(x_1, x_2) = 21{,}5 + x_1 \cdot \sin(4\pi x_1) + x_2 \cdot \sin(20\pi x_2)$$

$$\begin{array}{ll} -3.0 \le x_1 \le 12.1 & \bar{\mathrm{x}} = (\mathrm{x}_1, \mathrm{x}_2) \\ 4.1 \le \mathrm{x}_2 \le 5.8 & \bar{\sigma} = (\sigma_1, \sigma_2) \end{array}, \qquad (5.3)$$

в предположении поиска максимума.

Для определенности предположим, что в t-поколении текущая особь имеет вид:

$$(\bar{x}^t, \sigma) = ((5.3;4.9), (1.0;1.0)) \quad . \qquad (5.4)$$

Тогда потомок определяется следующим образом:

$$\left.\begin{aligned} x_1^{t+1} &= x_1^{t} + N(0;1.0) = 5.3 + 0.4 = 5.7 \\ x_2^{t+1} &= x_2^{t} + N(0;1.0) = 4.9 - 0.3 = 4.6 \end{aligned}\right\} \text{потомок} \quad , \tag{5.5}$$

где числа 0.4 и 0.3 получены случайным образом в соответствии с распределением Гаусса.

Поскольку $f(x^t) = f(5.3;4.9) = 18.383705 < 24.849532 = f(5.7;4.6) = f(x^{t+1})$ (значение ЦФ потомка лучше, чем у родителя), то полученный потомок заменяет родителя.

В целом алгоритм процесса эволюции двукратной (1+1)- эволюционной стратегии можно сформулировать следующим образом.

1. Выбрать множество P параметров X, необходимых для представления решения данной проблемы, и определить диапазон допустимых изменений каждого параметра:

$\{x_{1min}, x_{1max}\}, \{x_{2min}, x_{2max}\}, \ldots, \{x_{Pmin}, x_{Pmax}\}$,

установить номер поколения (итерации) t=0;

задать стандартное отклонение σ_i для каждого параметра, функцию f, для которой необходимо найти оптимум, и максимальное число поколений k.

2. Для каждого параметра случайным образом выбрать начальное значение из допустимого диапазона: множество этих значений составляет начальную популяцию (из одной особи) $X^t=(x_1, x_2, \ldots, x_P)$.

3. Вычислить значение оптимизируемой функции f для родительской особи $F^p=f(X^t)$.

4. Создать новую особь-потомка в соответствии с (5.2)

$$\bar{x}^* = \bar{x}^t + N(0, \bar{\sigma}).$$

5. Вычислить значение f для особи-потомка $F^o=f(X^*)$.

6. Сравнить значения функций f для родителя и потомка; если значение потомка F^o лучше, чем у родительской особи, то заменить родителя на потомка

$$\bar{x}^t = \bar{x}^*,$$

иначе оставить в популяции родителя.

7. Увеличить номер поколения t=t+1;

8. Если не достигнуто максимальное число поколений t<k, то переход на шаг 4, иначе выдать найденное решение X^t.

Несмотря на то, что фактически здесь популяция состоит из одной особи, рассмотренная стратегия называется двукратной ЭС. Причина в том, что здесь фактически происходит конкуренция потомка и родителя. Обычно вектор стандартных отклонений σ остается неизменным в течении всего процесса эволюции.

Поэтому, чтобы оптимизировать скорость сходимости этого процесса, И. Решенберг (основоположник ЭС) предложил правило успеха «1/5» .

Смысл его заключается в следующем - правило применяется после каждых *k* поколений процесса (где *k* – параметр этого метода):

$$\sigma^{t+1} = \begin{cases} c_d \cdot \sigma^t, если \varphi(k) < 1/5 \\ c_i \cdot \sigma^t, если \varphi(k) > 1/5 \\ \sigma^t, если \varphi(k) = 1/5 \end{cases}, \qquad (5.6)$$

где $\varphi(k)$ - отношение числа успешных мутаций к общему числу произведенных мутаций *k* (число успехов, деленное на *k*), которое называется коэффициентом успеха для оператора мутации в течении *k* последних поколений; величина $c_i>1$, $c_d<1$ – регулирует увеличение/уменьшение отклонения мутации.

Обычно на практике оптимальные значения полагают равными следующим величинам: $c_d=0.82$; $c_i=1/0.82=1.22$. Смысл этого правила в следующем:

- если коэффициент успеха $\varphi(k) > 1/5$, то отклонение σ^{t+1} увеличивается (мы идем более крупными шагами);
- если коэффициент успеха $\varphi(k) < 1/5$, то отклонение σ^{t+1} уменьшается (шаг поиска уменьшается).

Иногда рекомендуется устанавливать коэффициент мутации обратно пропорционально числу переменных в потенциальном решении (особи) и прямо пропорционально расстоянию от точки оптимального решения. Конечно, в реальных приложениях точное расположение оптимума неизвестно. Однако иногда может быть известна априорная информация об оптимуме (например, порядок величины). Даже ограниченная информация может быть полезна в процессе поиска в ЭС.

Многократная эволюционная стратегия

По сравнению с двукратной многократная эволюция отличается не только размером популяции (N > 2), но и имеет некоторые дополнительные отличия:

- все особи в поколении имеют одинаковую вероятность выбора для мутации;
- имеется возможность введения оператора рекомбинации (например, однородного ОК в ГА, рассмотренного в разделе 4), где два случайно выбранных родителя производят потомка по следующей схеме:

$$\underbrace{\begin{array}{l}(\bar{x}^1,\bar{\sigma}^1)=((x_1{}^1,\ldots,x_n{}^1),(\sigma_1{}^1,\ldots,\sigma_n{}^1))\\(\bar{x}^2,\bar{\sigma}^2)=((x_1{}^2,\ldots,x_n{}^2),(\sigma_1{}^2,\ldots,\sigma_n{}^2))\end{array}}_{(\bar{x},\bar{\sigma})=((x_1{}^{q_1},\ldots,x_n{}^{q_n}),(\sigma_1{}^{q_1},\ldots,\sigma_n{}^{q_n}))}, \qquad (5.7)$$

где $q_i=1$ или $q_i=2$, $i=1,\ldots,n$ (т.е. каждая компонента потомка копируется из первого или второго родителя).

Имеется еще одно сходство между двукратными и многократными эволюционными стратегиями. При обоих видах ЭС производится только один потомок. В двукратных стратегиях потомок соревнуется со своим родителем. В многократной стратегии самая слабая особь уничтожается.

В современной литературе используются следующие обозначения:

(1+1)-ЭС – двукратная стратегия (1 родитель производит 1 потомка);

(μ+1)-ЭС − многократная стратегия (μ родителей производят 1 потомка);

(μ+λ)-ЭС, где μ-родителей производят λ-потомков и отбор μ лучших представителей производится среди объединенного множества
((μ+λ) особей) родителей и потомков;

(μ,λ)-ЭС, где μ особей родителей порождает λпотомков, причем λ>μ и процесс выбора лучших производится только на множестве потомков.

Следует подчеркнуть, что в обоих последних видах ЭС обычно число потомков существенно больше числа родителей λ>μ (иногда полагают λ/μ=7).

Укрупненный алгоритм решения задачи с помощью ЭС можно представить следующим образом.

Установка счетчика поколений t=0;
Инициализация параметров;
Инициализация популяции C(0) из μ особей;
for каждой особи $x_i(t) \in C(t)$ **do**
    оценка значения фитнесс-функции $f(x_i(t))$;
end
while условие останова не выполнено **do**
    for i=1,…, λ do
        случайный выбор $\rho \geq 2$ родительских особей;
        построение особи-потомка путем кроссинговера на генотипе и параметрах родительских особей;
        мутация генотипа и параметров особи-потомка;
        оценка значения фитнесс-функции потомка;
    end
    Отбор следующего поколения популяции C(t+1);

```
    t=t+1;
end
```

Здесь на этапе инициализации генерируются особи начальной популяции со значениями в пределах ограничений и задаются начальные значения параметров. Для оценки качества особи используется абсолютное значение фитнесс-функции. Далее выполняются генетические операторы отбора, кроссинговера и мутации, наиболее распространенные варианты которых представлены ниже. В качестве критерия останова может быть использован любой из рассмотренных ранее.

Тестовые примеры

Для данного вида задачи существует большое число тестовых примеров – Benchmark-ов. С некоторыми из них можно познакомиться, например, в приложении 1. Для данных тестов произведено большое число исследований на скорость алгоритма, количество эпох для достижения результата и пр. С результатами этих исследований можно ознакомиться в научной литературе, доступной в Internet.

Многочисленные исследования доказывают, что ЭС не менее эффективно, а часто гораздо лучше справляются с задачами оптимизации в многомерных пространствах, при этом более просты в реализации из-за отсутствия процедур кодирования и декодирования хромосом.

5.2 Практическая часть

Порядок выполнения лабораторной работы

1.Выбрать свой вариант задания согласно таблице приложения 1.

1. Создать программу, использующую ЭС для нахождения оптимума функции согласно таблице вариантов, приведенной в приложении А. Для всех Benchmark-ов оптимумом является минимум. Программу выполнить на встроенном языке пакета Matlab (или любом, доступным вам, языке программирования).
2. Для n=2 вывести на экран график данной функции с указанием найденного экстремума, точек популяции. Для вывода графиков использовать стандартные возможности пакета Matlab. Предусмотреть возможность пошагового просмотра процесса поиска решения.
3. Исследовать зависимость времени поиска, числа поколений (генераций), точности нахождения решения от основных параметров генетического алгоритма:
   - число особей в популяции
   - вероятность мутации.

   Критерий остановки вычислений – повторение лучшего результата заданное количество раз или достижение популяцией определенного возраста (например, 100 эпох).
4. Повторить процесс поиска решения для n=3, сравнить результаты, скорость работы программы.

Содержание отчета.

1. Титульный лист установленной формы.
2. Условие задания с вариантом.
3. Распечатанный листинг программы.
4. Распечатка результатов выполнения программы (графиков).
5. Диаграммы исследованных зависимостей.

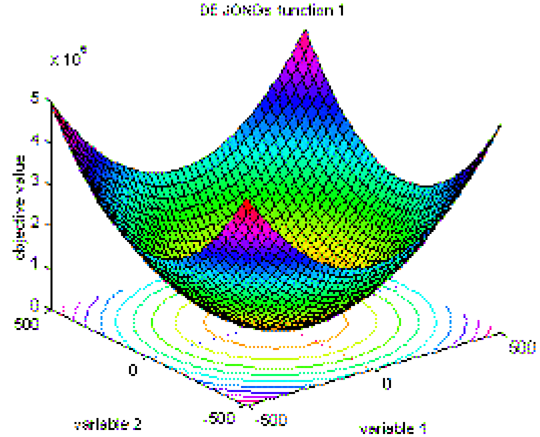
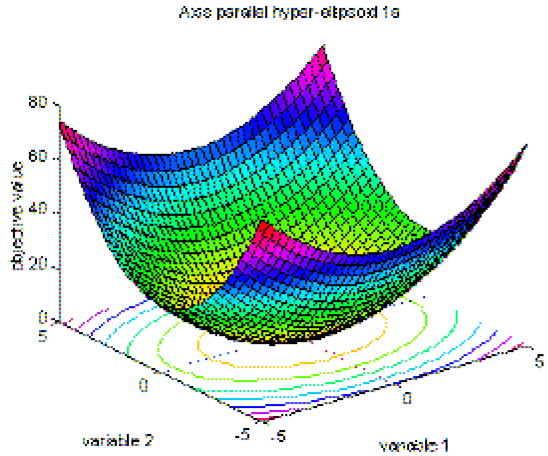
Контрольные вопросы

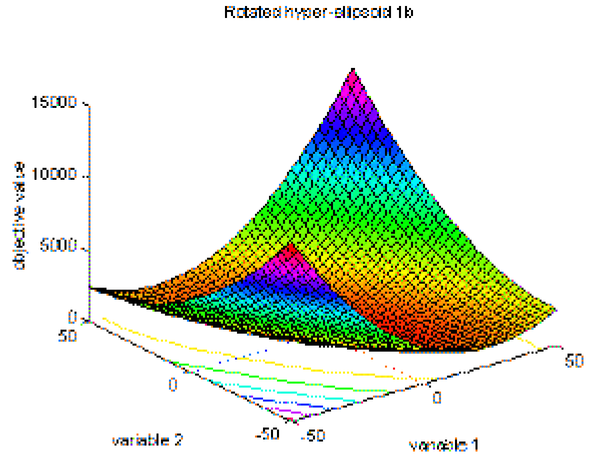
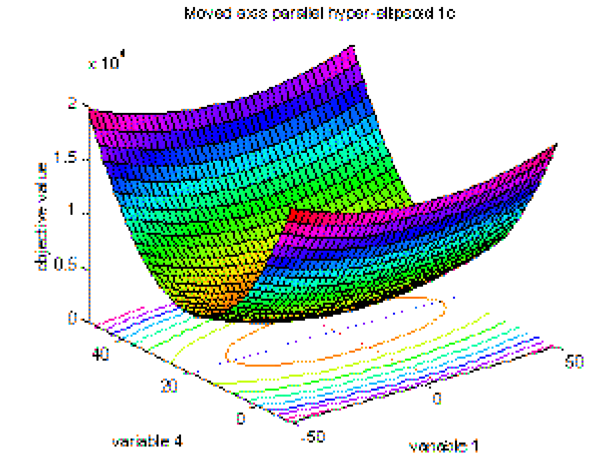
1. Как представляется потенциальное решение в ЭС?

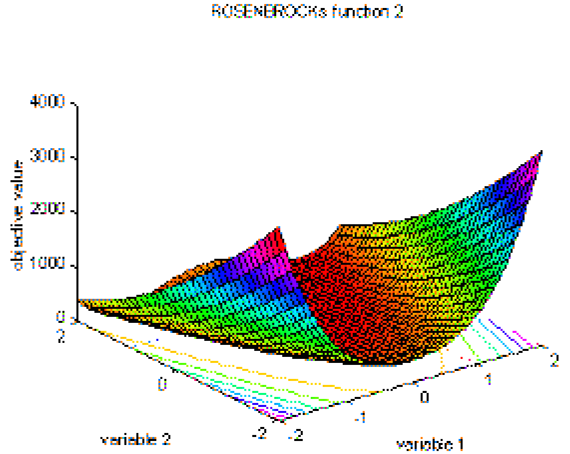
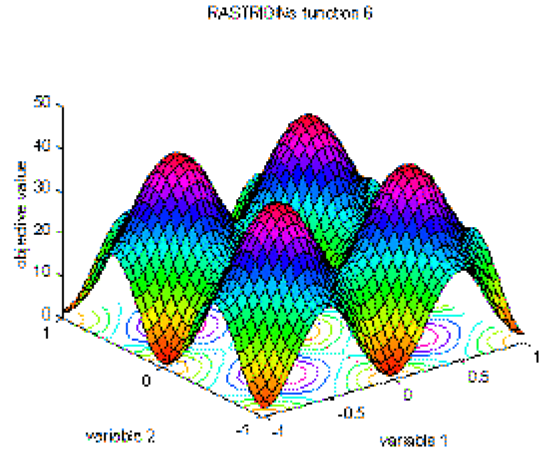
2. Какой генетический оператор применяется в ЭС?
3. Как выполняется мутация в ЭС?
4. Опишите двукратную ЭС.
5. Сформулируйте правило успеха в ЭС.
6. Что такое двукратная ЭС?
7. Что такое многократная ЭС?
8. Приведите основные параметры ЭС.
9. Что такое самоадаптация в ЭС?
10. Какие параметры, кроме отклонений, можно использовать в ЭС?
11. Как можно использовать углы вращения в ЭС?
12. Приведите основные стратегии самоадаптации.
13. Приведите основные типы операторов отбора в ЭС.
14. Что такое локальный и глобальный оператор кроссинговера?
15. Какие операторы рекомбинации могут быть использованы в ЭС?
16. Сформулируйте общий алгоритм решения задачи с использованием ЭС.
17. Как выполняется оператор мутации в ЭС?
18. Что такое направленная мутация?
19. Что общего между ГА и ЭС?
20. Каковы различия между ГА и ЭС?

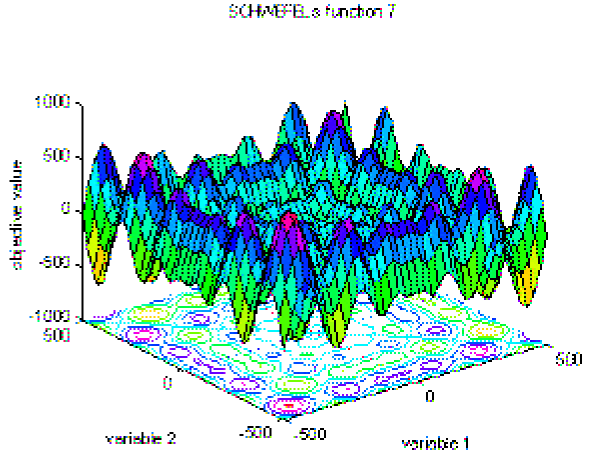
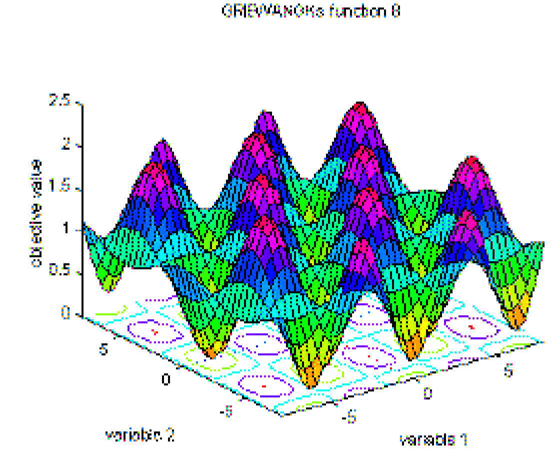
Приложение 1.

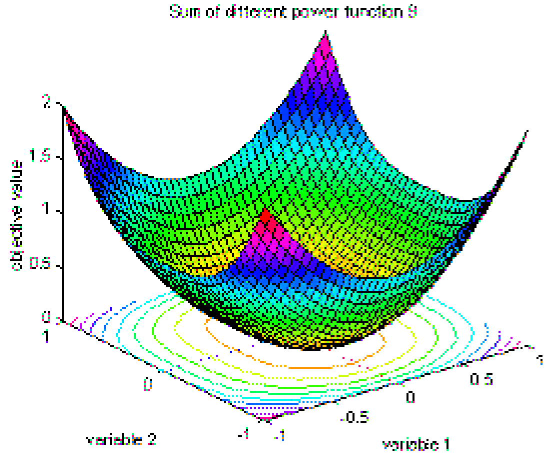
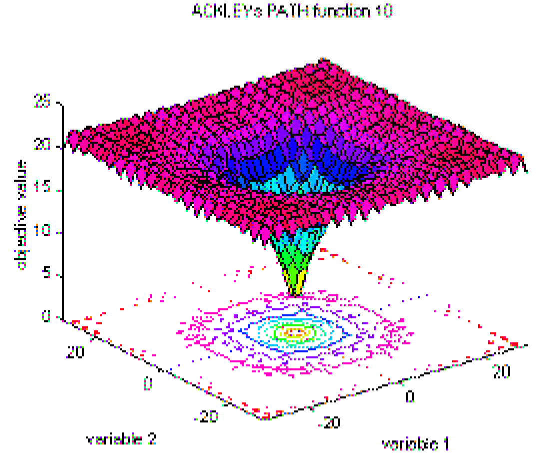
Индивидуальные задания на лабораторную работу №5.

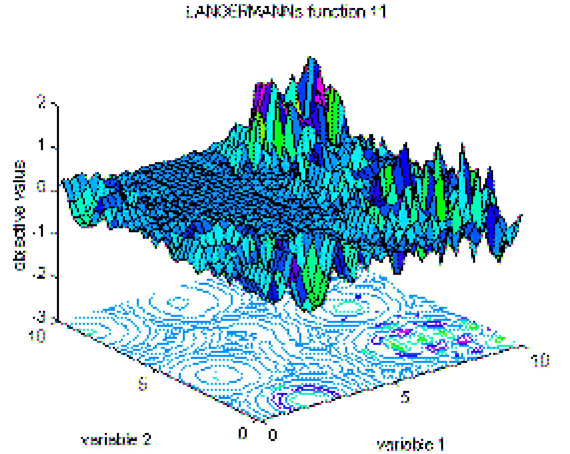
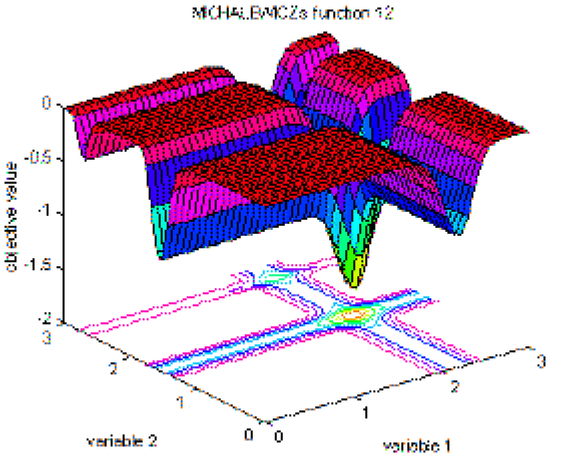
№ вв.	Название	Оптимум	Вид функции	График функции
1	De Jong's function 1	global minimum f(x)=0; x(i)=0, I=1:n.	$f_1(x)=\sum_{i=1}^{n} x_i^2 \qquad -5.12 \le x_i \le 5.12$ f1(x)=sum(x(i)^2), i=1:n;	
2	Axis parallel hyper-ellipsoid function	global minimum f(x)=0; x(i)= 0, i=1:n.	$f_{1a}(x)=\sum_{i=1}^{n} i\cdot x_i^2 \qquad -5.12 \le x_i \le 5.12$ f1a(x)=sum(i·x(i)^2), i=1:n;	

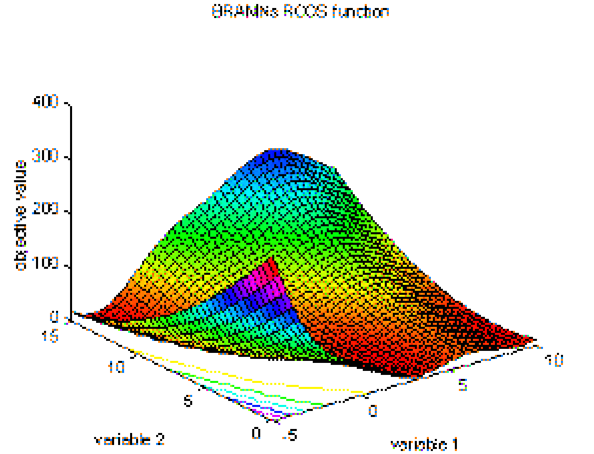
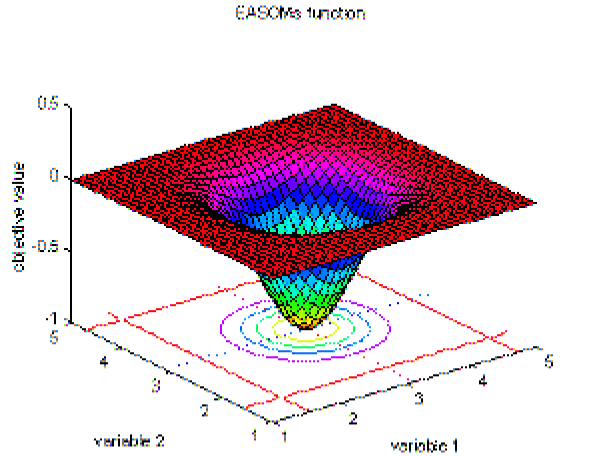
3	Rotated hyper-ellipsoid function	global minimum f(x)=0; x(i)=0, i=1:n	$f_{1b}(x)=\sum_{i=1}^{n}\left(\sum_{j=1}^{i}x_j\right)^2$ $-65.536 \le x_i \le 65.536$ f1b(x)=sum(sum(x(j)^2), j=1:i), i=1:n; .	
4	Moved axis parallel hyper-ellipsoid function	global minimum f(x)=0; x(i)= 5*i, i=1:n	$f_{1c}(x)=\sum_{i=1}^{n}5i\cdot x_i^2$ $-5.12 \le x_i \le 5.12$ f1c(x)=sum(5*i·x(i)^2), i=1:n;	

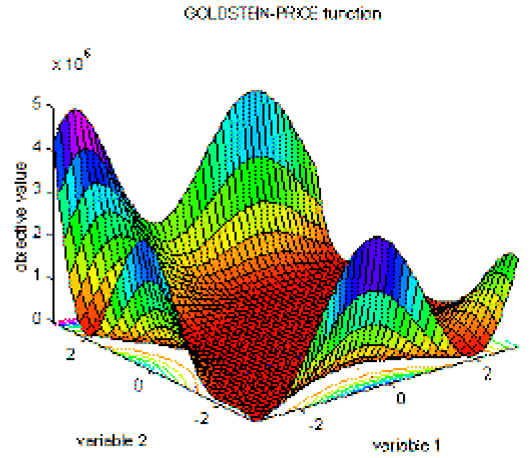
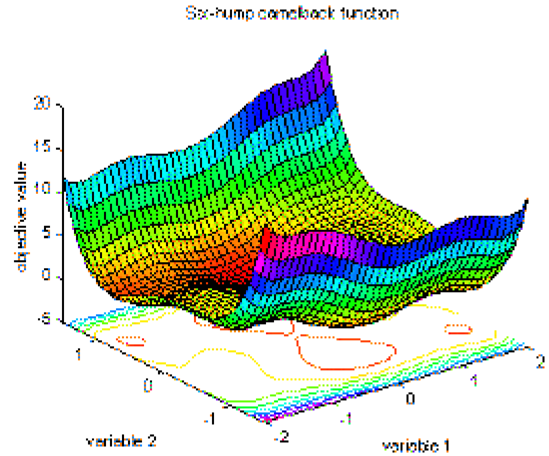
5	Rosenbrock's valley (De Jong's function 2)	global minimum f(x)=0; x(i)=1, i=1:n.	$f_2(x) = \sum_{i=1}^{n-1} 100 \cdot \left(x_{i+1} - x_i^2\right)^2 + \left(1 - x_i\right)^2 \qquad -2.048 \le x_i \le 2.048$ f2(x)=sum(100·(x(i+1)-x(i)^2)^2+(1-x(i))^2), i=1:n-1;	
6	Rastrigin's function 6	global minimum f(x)=0; x(i)=0, i=1:n.	$f_6(x) = 10 \cdot n + \sum_{i=1}^{n} \left(x_i^2 - 10 \cdot \cos\left(2 \cdot \pi \cdot x_i\right)\right) \qquad -5.12 \le x_i \le 5.12$ f6(x)=10·n+sum(x(i)^2-10·cos(2·pi·x(i))), i=1:n;	

7	Schwefel's function 7	global minimum f(x)=n·418.9829; x(i)=420.9687, i=1:n.	$f_7(x)=\sum_{i=1}^{n} -x_i \cdot \sin\left(\sqrt{\lvert x_i \rvert}\right) \qquad -500 \le x_i \le 500$ f7(x)=sum(-x(i)·sin(sqrt(abs(x(i))))), i=1:n;	
8	Griewangk's function 8	global minimum f(x)=0; x(i)=0, i=1:n	$f_8(x)=\sum_{i=1}^{n}\frac{x_i^2}{4000}-\prod_{i=1}^{n}\cos\left(\frac{x_i}{\sqrt{i}}\right)+1 \qquad -600 \le x_i \le 600$ f8(x)=sum(x(i)^2/4000)-prod(cos(x(i)/sqrt(i)))+1, i=1:n;	

9	Sum of different power function 9	global minimum f(x)=0; x(i)=0, i=1:n.	$f_9(x) = \sum_{i=1}^{n} \lvert x_i \rvert^{(i+1)} \qquad -1 \le x_i \le 1$ f9(x)=sum(abs(x(i))^(i+1)), i=1:n;	
10	Ackley's Path function 10	global minimum f(x)=0; x(i)=0, i=1:n.	$f_{10}(x) = -a \cdot e^{-b\cdot\sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}}} - e^{\frac{\sum_{i=1}^{n}\cos(c\cdot x_i)}{n}} + a + e^1 \qquad -1 \le x_i \le 1$ f10(x)=-a·exp(-b·sqrt(1/n·sum(x(i)^2)))-exp(1/n·sum(cos(c·x(i))))+a+exp(1); a=20; b=0.2; c=2·pi; i=1:n;	

11	Langermann's function 11	global minimum f(x)=-1.4 (for m=5); x(i)=???, i=1:n.	$f_{11}(x) = -\sum_{i=1}^{m} c_i \left(e^{-\frac{\|\bar{x}-A(i)\|^2}{\pi}} \cdot \cos\left(\pi \cdot \|\bar{x}-A(i)\|^2\right)\right)$ $i=1:m,\ 2 \le m \le 10,\ 0 \le x_i \le 10$ f11(x)=-sum(c(i)·(exp(-1/pi·sum((x-A(i))^2))·cos(pi·sum((x-A(i))^2)))), i=1:m, m=5; A(i),C(i)<>0, m=5	
12	Michalewicz's function 12	global minimum f(x)=-4.687 (n=5); x(i)=???, i=1:n. f(x)=-9.66 (n=10); x(i)=???, i=1:n.	$f_{12}(x) = -\sum_{i=1}^{n} \sin(x_i) \cdot \left(\sin\left(\frac{i \cdot x_i^2}{\pi} \right) \right)^{2\cdot m}$ $i=1:n,\ m=10,\ 0 \le x_i \le \pi$ f12(x)=-sum(sin(x(i))·(sin(i·x(i)^2/pi))^(2·m)), i=1:n, m=10; проверить для n=5,10	

13	Branins's rcos function	global minimum f(x1,x2)=0.397887; (x1,x2)=(-pi,12.275), (pi,2.275), (9.42478,2.475).	$f_{Bran}(x_1,x_2)=a\cdot(x_2-b\cdot x_1^2+c\cdot x_1-d)^2+e\cdot(1-f)\cdot\cos(x_1)+e \quad -5\le x_1\le 10,\ 0\le x_2\le 15$ $a=1,\quad b=\frac{5.1}{4\cdot\pi^2},\quad c=\frac{5}{\pi},\quad d=6,\quad e=10,\quad f=\frac{1}{8\cdot\pi}$ fBran(x1,x2)=a·(x2-b·x1^2+c·x1-d)^2+e·(1-f)·cos(x1)+e; a=1, b=5.1/(4·pi^2), c=5/pi, d=6, e=10, f=1/(8·pi);	
14	Easom's function	global minimum f(x1,x2)=-1; (x1,x2)=(pi,pi).	$f_{Easo}(x_1,x_2)=-\cos(x_1)\cdot\cos(x_2)\cdot e^{-\left((x_1-\pi)^2+(x_2-\pi)^2\right)} \quad -100\le x_i\le 100,\ i=1:2$ fEaso(x1,x2)=-cos(x1)·cos(x2)·exp(-((x1-pi)^2+(x2-pi)^2));	

15	Goldstein-Price's function	global minimum f(x1,x2)=3; (x1,x2)=(0,-1).	$f_{Gold}(x_1,x_2) = \left(1+(x_1+x_2+1)^2\cdot(19-14x_1+3x_1^2-14x_2+6x_1x_2+3x_2^2)\right)$ $\cdot\left(30+(2x_1-3x_2)^2\cdot(18-32x_1+12x_1^2+48x_2-36x_1x_2+27x_2^2)\right)$ $-2\le x_i\le 2,\ i=1:2$ fGold(x1,x2)=[1+(x1+x2+1)^2·(19-14·x1+3·x1^2-14·x2+6·x1·x2+3·x2^2)]·[30+(2·x1-3·x2)^2·(18-32·x1+12·x1^2+48·x2-36·x1·x2+27·x2^2)]	
16	Six-hump camel back function	global minimum f(x1,x2)=-1.0316; (x1,x2)=(-0.0898,0.7126), (0.0898,-0.7126).	$f_{Sixh}(x_1,x_2) = \left(4-2.1x_1^2+x_1^{4/3}\right)\cdot x_1^2+x_1x_2+\left(-4+4x_2^2\right)\cdot x_2^2 \quad -3\le x_1\le 3,\ -2\le x_1\le 2$ fSixh(x1,x2)=(4-2.1·x1^2+x1^4/3)·x1^2+x1·x2+(-4+4·x2^2)·x2^2;	

Для остальных вариантов берется строчка, соответствующая остатку от деления номера варианта на 16 (для 17-1, 18-2 и т.д.)