МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное учреждение высшего образования
«Юго-Западный государственный университет»
(ЮЗГУ)

Кафедра программной инженерии

УТВЕРЖДАЮ
Проректор по учебной работе
_______________ О.Г. Локтионова
«15» 12 _______________ 2017 г.

МЕТОДЫ ОПТИМИЗАЦИИ

Методические указания к лабораторным работам
по дисциплине «Методы оптимизации»
для студентов направлений подготовки
09.03.04 «Программная инженерия»

Курск 2017

УДК 519.6

Составитель В.В. Свиридов

Рецензент:
кандидат технических наук, доцент кафедры программной инженерии ЮЗГУ *В.В. Апальков*

Методы оптимизации: методические указания к лабораторным работам по дисциплине «Методы оптимизации» для студентов направления подготовки 09.03.04 «Программная инженерия» / ЮгоЗап. гос. ун-т; сост.: В.В. Свиридов. Курск, 2017. 90 с.

Изложены основные теоретические положения методов оптимальных решений. Рассмотрены примеры выполнения заданий. Приведены варианты заданий, контрольные вопросы к защите лабораторных работ.

Методические указания соответствуют требованиям рабочей программы по дисциплине «Метод оптимизации».

Материал предназначен для студентов направлений подготовки 09.03.04 «Программная инженерия» очной и заочной форм обучения.

Текст печатается в авторской редакции

Подписано в печать . Формат 60 х 84 1/16.
Усл. печ. л. . Уч.- изд. л. . Тираж 100 экз. Заказ . Бесплатно.
Юго-Западный государственный университет.

305040, Курск, ул. 50 лет Октября, 94.

Содержание

Лабораторная работа 1. Методы нулевого порядка. Методы одномерной минимизации

Цель работы: приобретение практических навыков для решения задач одномерной минимизации численными методами.

Задание

Для данной функции $f(x)$, $x \in [a,b]$ своего варианта из списка индивидуальных заданий

1. Составить блок-схемы алгоритмов поиска точки экстремума заданной функции.
2. Построить график функции для выбора границ первоначального интервала.
3. По разработанным алгоритмам составить программы поиска минимума функции.
4. Найти координаты и значение функции в точке минимума всеми методами.
5. Найти точное значение координаты точки минимума, используя необходимые и достаточные условия экстремума.
6. Проанализировать полученные результаты и сделать выводы по достигнутой точности и количеству вычислений функции.
7. Дать письменные ответы на контрольные вопросы.

Краткие теоретические сведения

Постановка задачи: Требуется найти безусловный минимум функции одной переменной $y = f(x), x \in R$, то есть, такую точку $x^* \in R = [a,b]$, что

$$f(x^*) = \min_{x \in R} f(x) .$$

Поставленная задача может быть решена с помощью необходимых и достаточных условий безусловного экстремума. Однако во многих практических случаях найти производные от заданной функции не представляется возможным. Поэтому решение задач одномерной оптимизации численными методами является актуальным при изучении методов оптимизации.

К основным численным методам одномерной минимизации относят**:**

- метод равномерного поиска;
- метод деления отрезка пополам;
- метод дихотомии;
- метод золотого сечения;
- метод Фибоначчи;
- метод квадратичной интерполяции и др.

Алгоритм равномерного поиска точки минимума

Алгоритм поиска минимума функции сводится к выполнению следующих этапов.

1 этап. Задается начальный интервал неопределенности $L = [a;b]$ и N - количество вычислений функции, $\Delta = b - a$, $dx = \dfrac{\Delta}{N-1}$.

2 этап. Вычислить точки $x_i = a + i \cdot dx, i = 0,...,N-1.$, равноотстоящие друг от друга.

3 этап. Вычислить значения функции в N найденных точках $y_i = F(x_i), i = 0,...,N-1.$

4 этап. Среди точек $y_i, i = 1,..,N-2$, найти такую точку x^*, в которой функция принимает наименьшее значение $x^* = \arg\min\limits_{i \in 0,...,N-1} y_i$.

Погрешность нахождения точки минимума методом перебора не превосходит $dx = \dfrac{2\Delta}{N-1}$.

Программы Матлаб.

```
%Нахождение максимума унимодальной функции
% методом пассивного поиска
function [x,Fx,k]=Passiv_poisk_max(a,b,eps)
D=b-a;
N=ceil(2*D/eps)+1;
dx=D/(N-1); X=a:dx:b;
Y=f(X);
[x,Fx]=max(Y); k=N;
%Функция максимизации
 function z=f(x)
z=(2+x)./(1+x.^2);
end end

%Нахождение минимума унимодальной функции
% методом пассивного поиска
function [x,Fx,k]=Passiv_poisk_min(a,b,eps)
D=b-a;
N=ceil(2*D/eps)+1;
dx=D/(N-1); X=a:dx:b;
```

```
Y=f(X);
[x,Fx]=min(Y); k=N;
%Функция минимизации
function z=f(x)
z=(2+x)./(1+x.^2);
  end end
```

Метод деления интервала пополам

Метод относится к последовательным стратегиям и позволяет исключать из дальнейшего рассмотрения на каждой итерации в точности половину текущего интервала неопределенности. Алгоритм уменьшения интервала основан на анализе величин функции в трех точках, равномерно распределенных на текущем интервале (делящих его на четыре равные части).

Поиск заканчивается, если длина текущего интервала неопределенности меньше заданной величины.

Алгоритм поиска точки минимума методом деления интервала пополам
Алгоритм поиска минимума функции сводится к выполнению следующих этапов.

1 этап. Задается начальный интервал неопределенности $L = [a,b]$ и $\varepsilon > 0$ – требуемая точность.

2 этап. Задать $k = 0$.

3 этап. Вычислить среднюю точку $x = \frac{a+b}{2}$ и значение функции в ней $Fx = f(x)$, длину интервала неопределенности $\Delta = b - a$;

4 этап. Вычислить точки $y = a + \frac{\Delta}{4}$, $z = b - \frac{\Delta}{4}$, которые с x делят интервал $L = [a,b]$ на четыре равные части, и значения функции в этих точках $Fy = f(y)$, $Fz = f(z)$.

5 этап. Если $Fy < Fx$, исключить интервал $[x;b]$, приняв $b = x$. Средней точкой нового интервала становится точка $x = y$, $Fx = Fy$. Перейти на этап 7.

Иначе перейти на этап 6.

6 этап. Если, $Fz < Fc$ исключить интервал $[a;x]$, приняв a=x. Средней точкой нового интервала становится точка $x = z$, $Fx = Fz$.

Перейти на этап 7.

Если $F_z \geq F_c$, исключить интервалы, $[a;y)$, $(z;b]$. Приняв $b = z$, a=y. Средней точкой нового интервала останется точка x.

7 этап. Вычислить $\Delta = |b - a|$ и проверить условие окончания: Если $\Delta \leq \varepsilon$, то процесс поиска завершается и $x^* \in [a;b]$. В качестве приближенного решения принимают середину последнего интервала $x^* \approx x_c$.

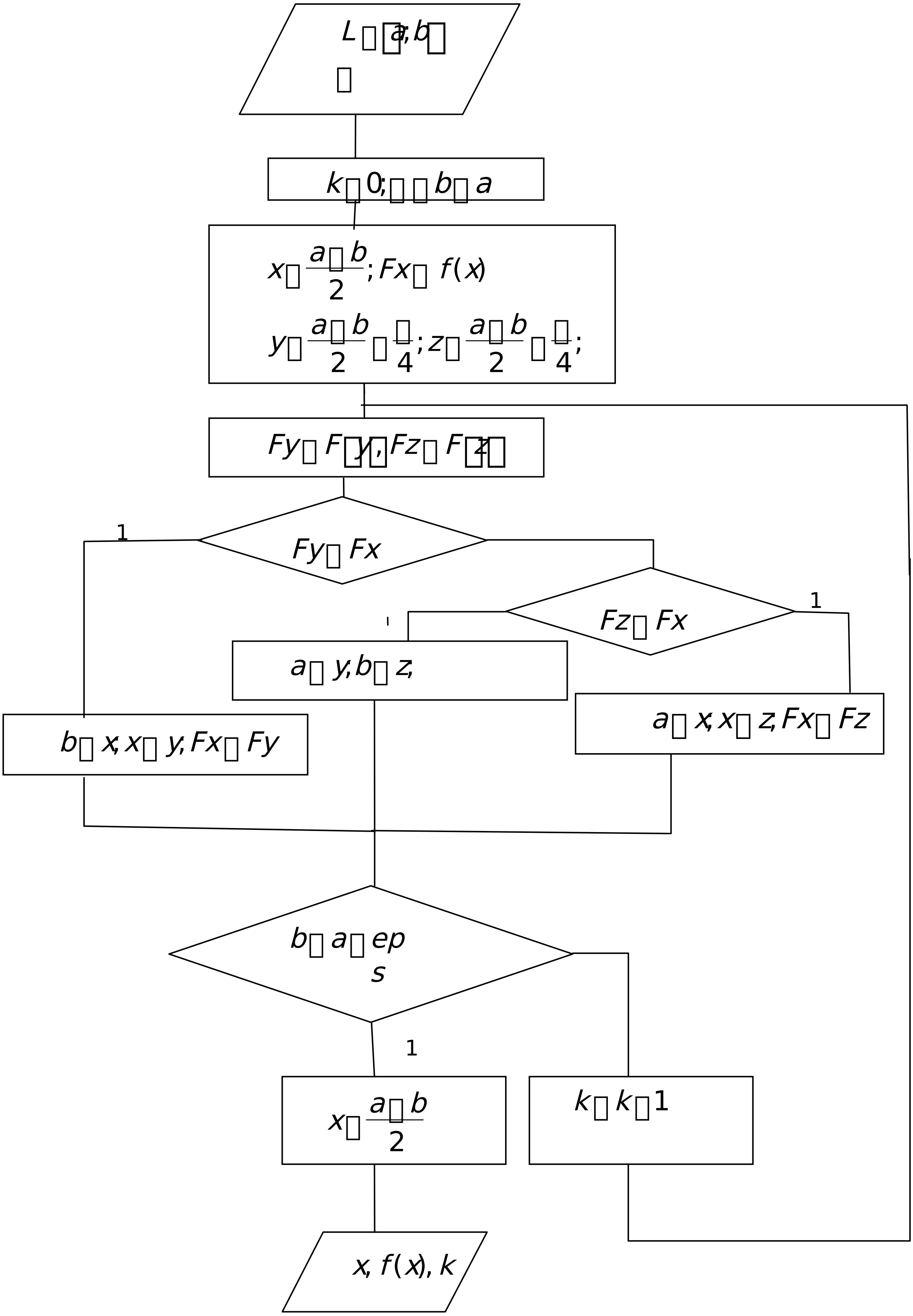
Если $\Delta > \varepsilon$, то принять $k = k + 1$ и перейти к этапу 4.

Для оценки сходимости метода используется характеристика относительного уменьшения начального интервала неопределенности

$R(N) = \frac{1}{2^{N/2}}$, где N количество вычислений функции.

```
% Метод деления отрезка пополам минимизация
function [x,Fx,k]=Minpopolam(a,b,eps)
x=(a+b)/2;k=0; while b-a>eps  k=k+1;
L=(b-a)/4; y=a+L;z=b-L; if (f(y)<f(x))
b=x;x=y; elseif f(z)<f(x) a=x;x=z; else
a=y;b=z; end end Fx=f(x);
%Функция минимизации
  function z=f(x)
z=(2+x)./(1+x.^2);
end end
```

```
%Метод деления отрезка пополам максимизация
function [x,Fx,k]=Maxpopolam(a,b,eps)
x=(a+b)/2;k=0; while b-a>eps  k=k+1; L=(b-a)/4;
y=a+L;z=b-L; if (f(y)>f(x)) b=x;x=y; elseif
f(z)>f(x) a=x;x=z; else a=y;b=z; end end Fx=f(x);
%Функция максимизации
function z=f(x)
z=(2+x)./(1+x.^2);   end
end
```


1
1
1

Метод дихотомии

Метод относится к последовательным стратегиям. Задается начальный интервал неопределенности и требуемая точность. Алгоритм опирается на анализ значений функции в двух точках. Для их нахождения текущий интервал неопределенности делится пополам и в обе стороны от середины откладывается по $\frac{\delta}{2}$, где δ малое положительное число. Поиск заканчивается, если длина текущего интервала неопределенности меньше заданной величины.

Алгоритм поиска точки минимума методом дихотомии

Алгоритм поиска минимума функции сводится к выполнению следующих этапов.

1 этап. Задается начальный интервал неопределенности $L_0 = [a;b]$ и ε – требуемая точность, $\delta > 0$ - малое положительное число.

2 этап. Задать $k = 0$.

3 этап. Вычислить $y = \frac{a+b-\delta}{2}$, $F(y)$, $z = \frac{a+b+\delta}{2}$, $F(z)$.

4 этап. Если $F(y) \le F(z)$, принять $b = z$ и перейти к этапу 5.

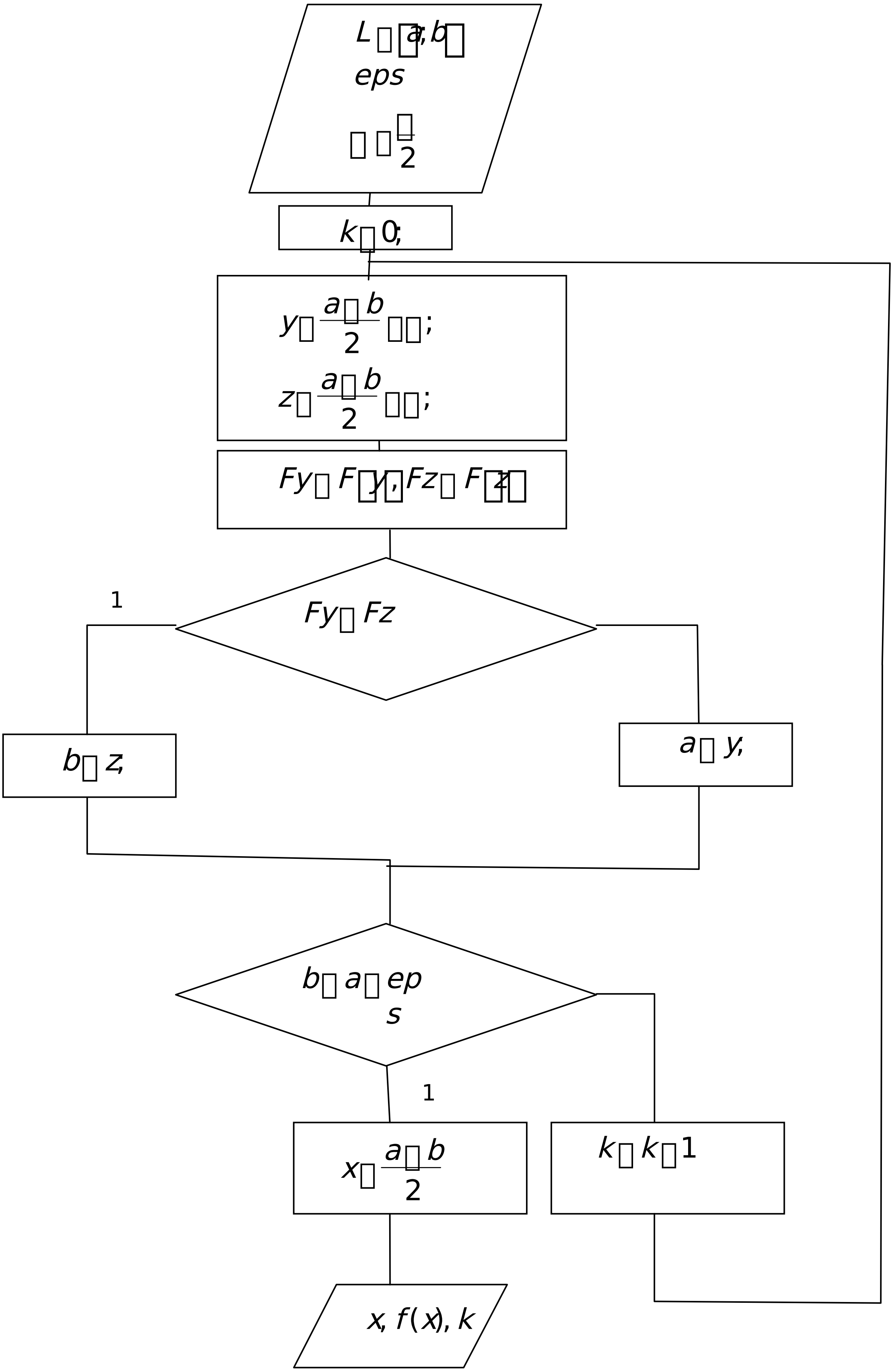
иначе, принять. $a = y$.

5 этап. Вычислить $\Delta = b - a$ и проверить условие окончания:

Если $\Delta \le \varepsilon$, то процесс поиска завершается и $x^* \in [a;b]$. В качестве приближенного решения принимают середину последнего интервала $x^* = \frac{a+b}{2}$.

Если $\Delta > \varepsilon$, то принять $k = k+1$ и перейти к этапу 3.

Для оценки сходимости метода используется характеристика относительного уменьшения начального интервала неопределенности $R(N) = \frac{1}{2^{N/2}}$, где N количество вычислений функции.



Программы Матлаб:

```
function [ x,Fx,k ] =Dihot_max( a,b,eps,delt) %
Метод дихотомии –максимизация
k=0; while(b-a>eps)
k=k+1;
y=(a+b)/2-delt;
z=(a+b)/2+delt;
Fy=f(y);  Fz=f(z);
 if Fy>=Fz      b=z;
else      a=y;     end
end
x=(a+b)/2;Fx=f(x);
% максимизируемая  функция
                          function       z=f(x)
z=(2+x)./(1+x.^2);
   end
end
```

```
function [ x,Fx,k ] =Dihot_min( a,b,eps,delt)
% Метод дихотомии - минимизация
k=0; while(b-a>eps)
k=k+1;
y=(a+b)/2-delt;
z=(a+b)/2+delt;
Fy=f(y);  Fz=f(z);
 if Fy<=Fz      b=z;    else
a=y;     end  end
x=(a+b)/2;Fx=f(x); %
минимизируемая функция
                          function       z=f(x)
z=(2+x)./(1+x.^2);
   end
end
```

Метод золотого сечения

В методе золотого сечения в качестве двух внутренних точек выбираются точки золотого сечения.

Точка производит золотое сечение, если отношение длины всего отрезка к большей части равно отношению большей части к меньшей части. На отрезке $[a_0, b_0]$ имеются две симметричные относительно его концов точки y_0 и z_0:

$$\frac{b_0-a_0}{b_0-y_0}=\frac{b_0-y_0}{y_0-a_0}=\frac{b_0-a_0}{z_0-a_0}=\frac{z_0-a_0}{b_0-z_0}=\frac{1+\sqrt{5}}{2}\approx 1.618..$$

Точка y_0 производит золотое сечение отрезка $[a_0, z_0]$, а точка z_0 - отрезка $[y_0, b_0]$.

Метод относится к последовательным стратегиям. Задается начальный интервал неопределенности и требуемая точность. Алгоритм опирается на анализ значений функции в двух точках. В качестве точек вычисления функции выбираются точки золотого сечения. На каждой итерации, кроме первой, требуется только одно новое вычисление функции. Поиск заканчивается, если длина текущего интервала неопределенности меньше заданной величины.

Алгоритм поиска точки минимума методом золотого сечения

Алгоритм поиска минимума функции сводится к выполнению следующих этапов.

1 этап. Задается начальный интервал неопределенности $L_0 = [a_0, b_0]$ и $\varepsilon > 0$- требуемая точность.

2 этап. Задать $k = 0$.

3 этап. Вычислить $y_k = a_0 + \frac{3 - \sqrt{5}}{2}(b_0 - a_0)$, $z_k = a_0 + b_0 - y_k$.

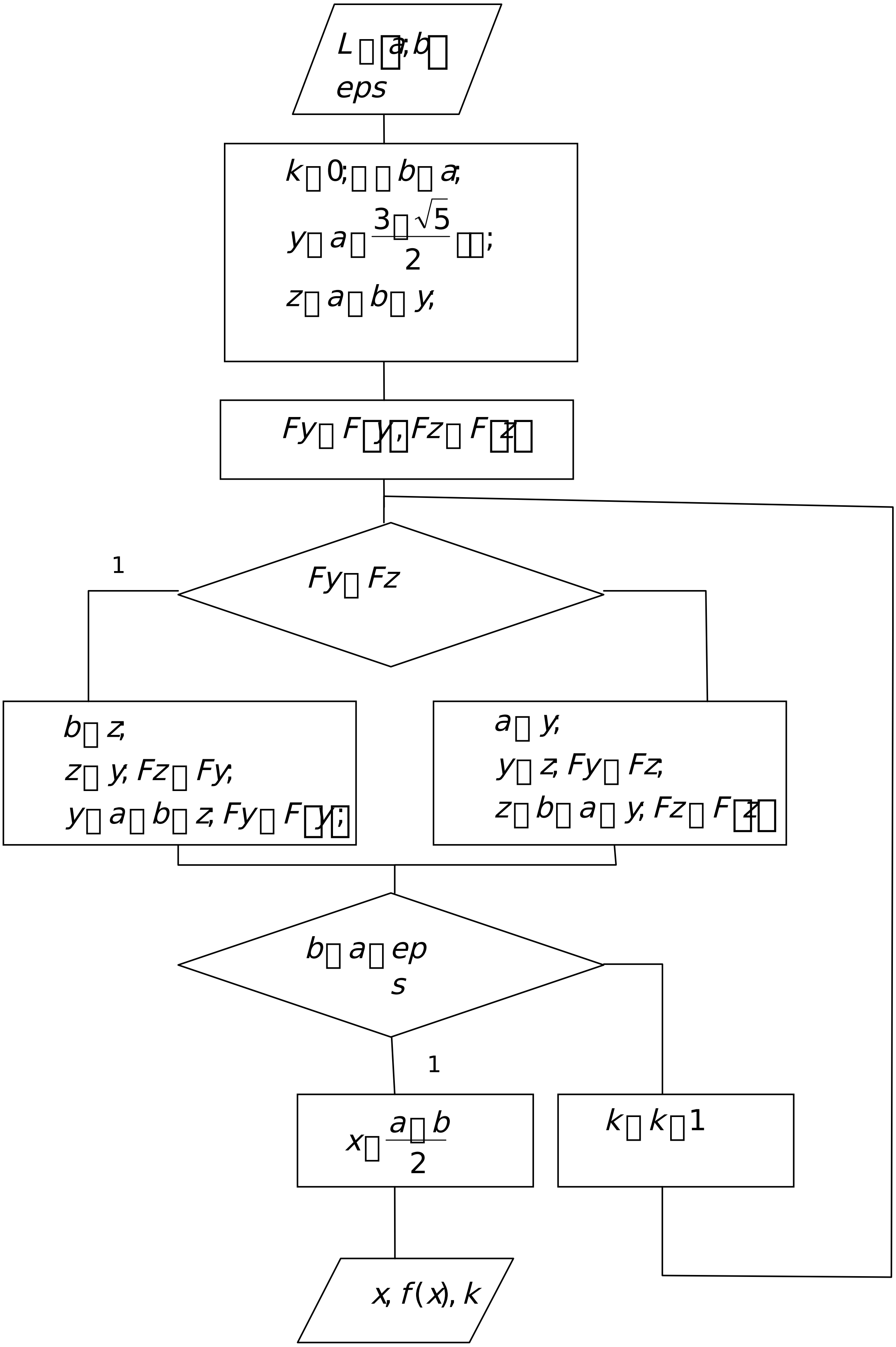
4 этап. Вычислить $F(y_k)$, $F(z_k)$.

5 этап. Если $F(y_k) \le F(z_k)$, то принять $a_{k+1} = a_k$, $b_{k+1} = z_k$ и $y_{k+1} = a_{k+1} + b_{k+1} - y_k$, $z_{k+1} = y_k$. Перейти к этапу 6.

Если $F(y_k) > F(z_k)$, принять $a_{k+1} = y_k$, $b_{k+1} = b_k$ и $y_{k+1} = z_k$, $z_{k+1} = a_{k+1} + b_{k+1} - z_k$.

6 этап. Вычислить $\Delta = |a_{k+1} - b_{k+1}|$ и проверить условие окончания поиска. Если $\Delta \le \varepsilon$, процесс поиска прекращается и $x^* \in [a_{k+1}, b_{k+1}]$. В качестве приближенного решения принимают середину последнего интервала $x^* = \frac{a_{k+1} + b_{k+1}}{2}$.

Если $\Delta > \varepsilon$, принять $k = k + 1$ и перейти к этапу 4.



Программы Матлаб:

```
function [ x,Fx,k ] = Zol_sech_max( a,b,eps) %
Метод золотого сечения – максимизация
y=a+(3-sqrt(5))/2*(b-a);
z=a+b-y;
Fy=f(y);Fz=f(z); k=0;
while(b-a>eps)
k=k+1;  if Fy>=Fz
b=z;     z=y;     Fz=Fy;
y=a+b-z;
    Fy=f(y);

 else     a=y;     y=z;
Fy=Fz;     z=a+b-y;
Fz=f(z);  end  end
x=(a+b)/2;Fx=f(x); %
максимизируемая функция
                      function      z=f(x)
z=(2+x)./(1+x.^2);
   end
end

function [ x,Fx,k ] = Zol_sech_min( a,b,eps) %
Метод золотого сечения – минимизация
y=a+(3-sqrt(5))/2*(b-a);
z=a+b-y;
Fy=f(y);Fz=f(z); k=0;
while(b-a>eps)
k=k+1;
 if Fy<=Fz
b=z;     z=y;
y=a+b-z;
Fy=f(y);
else
    a=y;
y=z;
z=a+b-y;
Fz=f(z);
end  end
x=(a+b)/2;Fx=f(x);
% минимизируемая функция
                      function      z=f(x)
z=(2+x)./(1+x.^2);
```

```
    end
end
```

Алгоритм поиска точки минимума методом Фибоначчи

Алгоритм поиска минимума функции сводится к выполнению следующих этапов.

1 этап. Задается начальный интервал неопределенности $L_0 = [a_0, b_0]$, $l > 0$- допустимая длина конечного интервала, $\varepsilon > 0$ - константа различимости.

2 этап. Найти количество вычислений функции как наименьшее целое число, при котором удовлетворяется условие $Fi_N \geq \frac{|L_0|}{l}$ и числа Фибоначчи $Fi_0, Fi_1, \ldots, Fi_N$.

3 этап. Задать $k = 0$.

4 этап. Вычислить $y_k = a_0 + \frac{Fi_{N-2}}{Fi_N}(b_0 - a_0)$, $z_k = a_0 + \frac{Fi_{N-1}}{Fi_N}(b_0 - a_0)$.

5 этап. Вычислить $F(y_k)$, $F(z_k)$.

6 этап. Если $F(y_k) \leq F(z_k)$, то принять $a_{k+1} = a_k$, $b_{k+1} = z_k$, $z_{k+1} = y_k$, и $y_{k+1} = a_{k+1} + \frac{Fi_{N-k-3}}{Fi_{N-k-1}}(b_{k+1} - a_{k+1})$. Перейти к этапу 7.

Если $F(y_k) > F(z_k)$, принять $a_{k+1} = y_k$, $b_{k+1} = b_k$, $y_{k+1} = z_k$ и $z_{k+1} = a_{k+1} + \frac{Fi_{N-k-2}}{Fi_{N-k-1}}(b_{k+1} - a_{k+1})$.

7 этап. Если $k \neq N-3$, то принять $k = k+1$ и перейти к этапу 5. Если $k = N-3$, то всегда $y_{N-2} = z_{N-2} = \frac{(a_{N-2} + b_{N-2})}{2}$, то есть отсутствует точка нового вычисления функции. В этом случае следует принять $y_{N-1} = y_{N-2} = z_{N-2}$; $z_{N-1} = y_{N-1} + \varepsilon$. В точках y_{N-1}, z_{N-1} вычисляются значения функции и находятся границы конечного интервала неопределенности:

– если $F(y_{N-1}) \leq F(z_{N-1})$, то принять $a_{N-1} = a_{N-2}$, $b_{N-1} = z_{N-1}$;

– если $F(y_{N-1}) > F(z_{N-1})$, то принять $a_{N-1} = y_{N-1}$, $b_{N-1} = b_{N-2}$.

Процесс поиска завершается и $x^* \in [a_{N-1}, b_{N-1}]$. В качестве приближенного решения можно принять любую точку интервала, рекомендуется $x^* = \frac{a_{N-1} + b_{N-1}}{2}$.

Характеристика относительного уменьшения начального интервала неопределенности $R(N) = \frac{1}{Fi_N}$, где N количество вычислений функции.

Метод квадратичной интерполяции (метод Пауэлла)

Метод квадратичной интерполяции относится к последовательным стратегиям. Задается начальная точка и с помощью пробного шага находятся три точки так, чтобы они были как можно ближе к искомой точке минимума. В полученных точках вычисляются значения функции. Затем строится интерполяционный полином второй степени, проходящий через эти три точки. В качестве приближения точки минимума берется точка минимума полинома. Процесс поиска заканчивается, когда полученная точка отличается от наилучшей из трех опорных точек не более, чем на заданную величину.

Алгоритм поиска точки минимума с квадратичной интерполяцией

Алгоритм поиска минимума функции сводится к выполнению следующих этапов.

1 этап. Задать начальную точку x_1, величину шага $\Delta x > 0$, $\varepsilon_1, \varepsilon_2$ – малые положительные числа, характеризующие точность. 2 этап. Вычислить $x_2 = x_1 + \Delta x$.

3 этап. Вычислить $f_1 = F(x_1), f_2 = F(x_2)$.

4 этап. Если $F(x_1) > F(x_2)$, то принять $x_3 = x_1 + 2\Delta x$.

Если $F(x_1) \le F(x_2)$, то принять $x_3 = x_1 - \Delta x$.

5 этап. Вычислить $f_3 = F(x_3)$.

6 этап. Найти $F_{min} = \min(f_1, f_2, f_3), x_{min} = x_i : F(x_i) = F_{min}$.

7 этап. Вычислить точку минимума интерполяционного полинома, построенного по трем точкам:

$$\bar{x} = \frac{1}{2}\frac{f_1(x_2^2 - x_3^2) + f_2(x_3^2 - x_1^2) + f_3(x_1^2 - x_2^2)}{f_1(x_2 - x_3) + f_2(x_3 - x_1) + f_3(x_1 - x_2)}$$

и величину $F(\bar{x})$. Если знаменатель в формуле для $\bar{x}$ на некоторой итерации обращается в нуль, то результатом интерполяции является прямая линия. В этом случае рекомендуется принять $x_1 = x_{min}$ и перейти к шагу 2.

8 этап. Проверить выполнение условий окончания

$$\left|\frac{F_{min} - F(\bar{x})}{F(\bar{x})}\right| < \varepsilon_1, \quad \left|\frac{x_{min} - \bar{x}}{\bar{x}}\right| < \varepsilon_2.$$

Если оба условия выполнены, то процедура закончена и $x^* = \bar{x}$.

Если хотя бы одно из условий не выполнено и $\bar{x} \in [x_1, x_3]$, выбрать наилучшую точку (x_{min} или $\bar{x}$) и две точки по обе стороны от нее. Переобозначить эти точки в порядке возрастания и перейти к этапу 6.

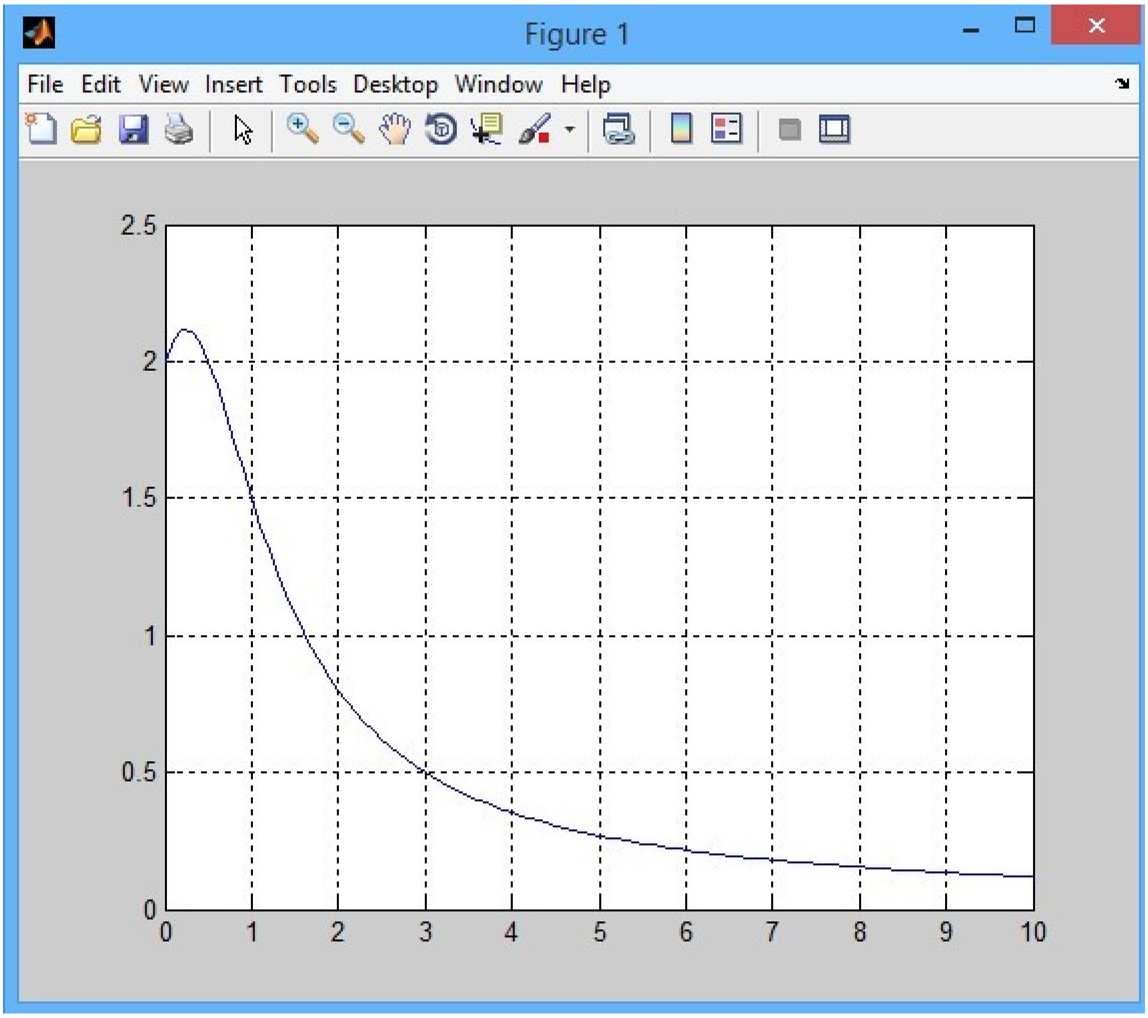
Если хотя бы одно из условий не выполнено и $\bar{x} \notin [x_1, x_3]$, то принять $x_1 = \bar{x}$ и перейти к этапу 2.

Пример выполнения

Найти с точностью eps = 0.00001 все локальные минимумы и максимумы функции $f(x) = \frac{2+x}{1+x^2}$ на отрезке $[0;10]$.

1. Для отделения точек экстремума интервалами унимодальности построим график данной функции на отрезке $[0;10]$

```
>> f=inline('(2+x)./(1+x.^2)');x=0:0.1:10;plot(x,f(x));hold on;grid;
```



Видим, что функция имеет единственный локальный экстремум, который является точкой максимума, расположенный на отрезке [0;1](можно выбрать любой отрезок унимодальности).

1) Применяем метод пассивного поиска.

```
>>[x,Fx,k]=Passiv_poisk_max(0,1,0.00001)
>>x =
   0.236070000000000
Fx =
   2.118033988741689
k =
200001
```

Итак: Точка локального максимума на отрезке [0;1] с абсолютной погрешностью 0.00001это x = 0.236070000000000 значение максимума равно

Fx =2.118033988741689 и потребовалось k = 200001 вычисление значения функции методом пассивного поиска.

2) Решим данную задачу методом деления отрезка пополам.

```
>>  [x,Fx,k]=Maxpopolam(0,1,0.00001) x
=
   0.236068725585938
Fx =
   2.118033988748772
k = 17
```

Итак: Точка локального максимума на отрезке [0;1] с абсолютной погрешностью 0.00001это x = 0.236068725585938, значение максимума равно

Fx =2.118033988748772 и потребовалось k = 17 итераций метода деления пополам

1) Используем М-функцию метода дихотомии для локальной макимизации:

```
>> [ x,Fx,k ] =Dihot_max( 0,1,0.00001,0.000001) x
=
0.236065438758850
Fx =
2.118033988736964
k = 17
```

Итак: Точка локального максимума на отрезке [0;1] с абсолютной погрешностью 0.00001 это x = 0.236065438758850, значение максимума равно

Fx =2.118033988736964и потребовалось k = 17 итераций метода дихотомии.

2) Используем метод золотого сечения, функции для максимизации и минимизации;

В данном случае применяем М-функцию максимизации:

```
>> [ x,Fx,k ] = Zol_sech_max( 0,1,0.00001) x
=
0.236066839071308
Fx =
2.118033988747295
k = 24
```

Итак: Точка локального максимума на отрезке [0;1] это x = 0.236066839071308, значение максимума равно

Fx = 2.118033988747295 и потребовалось k = 24 итераций метода золотого

Варианты индивидуальных заданий

№	F(X) =	Тип экстремума	Исходный интервал	Погрешность
1	$x^2 \square \sin(x)$	min	[-1; 0]	0.005
2	$0.1\square x \square \cos(x)$	max	[4; 9]	0.02
3	$\exp(x) \square x^2$	min	[-1; 0]	0.005
4	$\square x \square (x \square 2) / x^2$	min	[-2; 0]	0.01
5	$x\square \ln(\ln(x))$	min	[1.3; 3.0]	0.01
6	$0.2\square x \square \sin(2x)$	max	[0; 3]	0.02
7	$5/(x^2 \square 2x \square 5)$	max	[0.8; 2.0]	0.008
8	$\exp(x \square 1) \square 1/ x$	min	[1; 1.5]	0.01
9	$\exp(1/ x) \square \ln(x)$	min	[1; 3]	0.012
10	$x\square\exp(\square 0.5x)$	max	[0; 3]	0.02
11	$1 / \square\exp(\square x)\ x$	max	[-1; 0.5]	0.005
12	$2 \square x\square x^2$	min	[0; 2]	0.01
13	$x^4 \square 2x^2 \square 4x$	min	[-1; 0]	0.002
14	$x^2 \square x\square\exp(\square x)$	min	[0; 1]	0.005

15	$\exp(x)\,\square 1/(x\,\square\,2)$	min	[-1; 1]	0.01
16	$5x^2\exp(\square 0.5x)$	max	[2; 6]	0.02
17	$5/x\,\square\,x^2$	min	[0.5; 2]	0.01
18	$\exp(\square 2x)\,\square\,x^2/2$	min	[0; 1.5]	0.01
19	$2\,\square\,(\ln(x))^2\,\square\,x/2$	min	[0.5; 2]	0.005
20	$x^2\,\square 1/arctg(x)$	min	[0; 2]	0.005

Контрольные вопросы

1. В чем состоит необходимое и достаточное условие экстремума одномерной функции?
2. В чем заключается условие унимодальности функции и как это условие используется?
3. Понятие выпуклой функции.
4. Как найти экстремум функции?
5. Как ведет себя производная в области точки экстремума?
6. Верно ли утверждение, что всякая выпуклая непрерывная на отрезке функция является на этом отрезке унимодальной?
7. Как ведет себя касательная к выпуклой функции? Поведение ее в области экстремума?
8. Можно считать, что глобальный минимум является локальным? А наоборот?
9. В чем различие между пассивным и последовательным поиском?
10. Что называют интервалом неопределенности в задачах одномерной оптимизации?
11. В чем состоит метод дихотомии?
12. Какие трудности возникают в методе квадратичной аппроксимации?
13. Каким образом сравнивают эффективность методов прямого поиска?

Лабораторная работа 2. Методы нулевого порядка. Методы многомерной минимизации

Цель работы: изучить методы нулевого порядка многомерной оптимизации, алгоритм Нелдера-Мида.

Задание

Найти точку минимума заданной функции двух переменных $f(x_1, x_2)$ с помощью алгоритма Нелдера-Мида из начальной точки $x0 = (0,0)^T$ с точностью по значениям функции $\varepsilon = 0.1$, устанавливая следующие параметры алгоритма: $\alpha = 1, \beta = 2, \gamma = 0.5$

Краткие теоретические положения

Рассмотрим задачу минимизации без ограничений в n-мерном пространстве R^n: $f(x) \to_{x \in R^n} \min$

Метод оптимизации называется методом *нулевого порядка*, если он использует только значения оптимизируемой функции и не использует значения ее производной.

На практике часто применяется метод нулевого порядка – метод подвижного симплекса, который начинает свою работу с инициализации по своей начальной точке x_0 и величине ребра a, которая выполняется по формулам: $x_j = x_0 + a \cdot e_j$, $j = 1,...,n$.

Таким образом, получаем $n+1$ вершину $(x_0,...,x_n)$ в n мерном пространстве, являющиеся вершинами симплекса ненулевого объема.

Для найденных вершин симплекса находятся следующие вершины: x_h, x_l, x_s, где $x_h = \underset{x_i \in (x_0,...,x_n)}{\operatorname{argmax}} f(x_i)$ – вершина с наибольшим значением целевой функции среди вершин подвижного симплекса – худшая вершина, $x_l = \underset{x_i \in (x_0,...,x_n) \setminus (x_h)}{\operatorname{argmin}} f(x_i)$ – вершина с наименьшим значением целевой функции, обязательно отличная от x_h – лучшая вершина и вершина $x_s \underset{x_i \in (x_0,...,x_n) \setminus (x_h, x_l)}{\operatorname{argmax}} f(x_i)$ – вершина вторая по величине сверху после x_h, причем отличная от x_h, x_l.

Идея алгоритма подвижного симплекса заключается в замене худшей точки x_h на лучшую, получаемую расчетным путем. При этом применяются *операции отражения, сжатия и растяжения.*

При операции отражения мы отражаем с некоторым коэффициентом $\alpha \in [1,2]$ худшую вершину x_h относительно центра $\bar{x} = \frac{1}{n} \sum_{i=0,\dots,n,\, i \neq h} x_i$ вершин, отличных от x_h по формуле $x_r = \bar{x} + \alpha(\bar{x} - x_h)$. Параметр α является установочным параметром алгоритма и обычно принимается $\alpha = 1$.

Если оказывается, что $f_r = f(x_r)$ удовлетворяет неравенствам $f_l \le f_r \le f_s$, то мы заменяем худшую вершину x_h на вновь найденную $x_h := x_r$.

Если же $f_r < f_l$, то в направлении от x_h к x_l делается дополнительное продвижение по формуле растяжения $x_e = \bar{x} + \gamma(\bar{x} - x_h)$, где $\gamma \in [2,3], \gamma > \alpha$.

Если операция растяжения оказалась успешной, т.е. $f_e = f(x_e) < f_r$, то замену худшей точки выполняем по формуле $x_h := x_e$.

Если же шаг растяжения оказался неудачным, т.е. $f_e \ge f_r$, то для замены худшей точки используем присваивание: $x_h := x_r$.

Если же выполняется неравенство $f_s < f_r < f_h$, то делаем замену $x_h := x_r$ и выполняем сжатие к новой вершине x_h по формуле $x_r = \bar{x} + \beta(x_h - \bar{x})$, где $\beta \in [0,1]$, обычно принимается $\beta = \frac{1}{2}$.

В случае же $f_r \ge f_h$ сжатие к вершине x_h по указанной формуле производится без замены $x_h := x_r$. В обоих случаях далее анализируется полученная в результате сжатия точка x_c и значение функции $f_c = f(x_c)$ в ней.

Если оказывается, что $f_c < f_h$, то осуществляем замену худшей точки $x_h := x_c$, иначе производим процедуру общего сжатия всего симплекса к лучшей точке x_l по формуле $x_i = x_l + \frac{1}{2}(x_i - x_l)$, $i = 0,\dots,n$.

Алгоритм Нелдера-Мида завершает свою работу, если выполняется условие:

$$\left(\frac{1}{n+1} \sum_{i=0}^{n} \left(f_i - f(\bar{x})\right)^2 \right)^{1/2} \le \varepsilon,$$

где ε - заданная точность расчета и $\bar{x} = \frac{1}{n+1} \sum_{i=0,\dots,n} x_i$ – центр всего симплекса.

Блок-схема алгоритма изображена ниже.

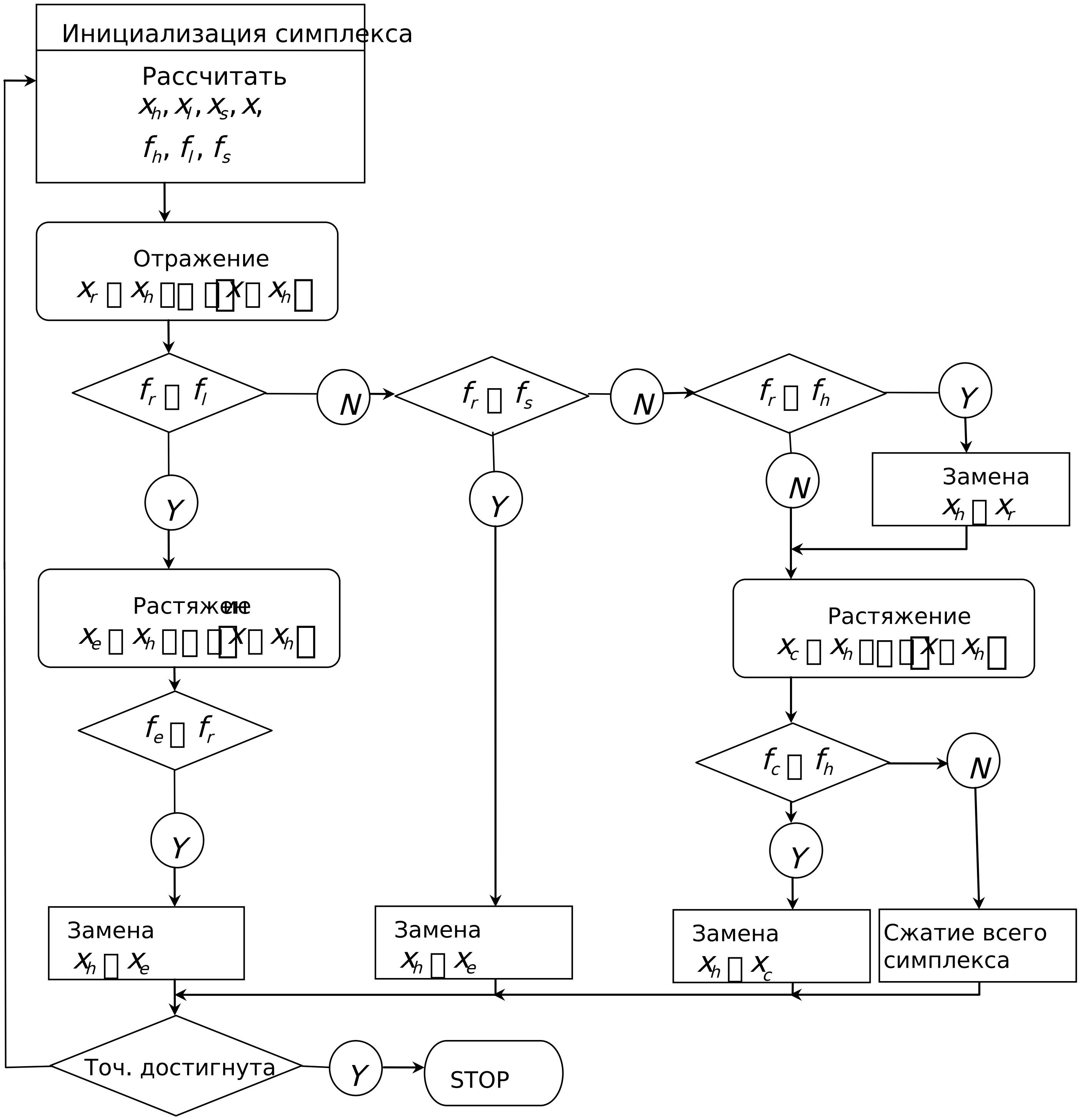


Рис. 2.1 Блок-схема алгоритма Нелдера-Мида.

Пример выполнения задания в пакете MathCad.

Исходные данные

$$f(x) := 2\cdot(x_1)^2 - 0.5\cdot x_1\cdot x_2 + 3\cdot(x_2)^4 - 5\cdot x_1 + 4\cdot x_2 + 3$$

$$x0 := (0 \quad 0)^T$$

$$a := 1$$

$$\alpha := 1 \qquad \beta := 2 \qquad \gamma := 0.5$$

$$\varepsilon := 0.1$$

Инициализация

$$X_1 := x0$$

$$X_2 := x0 + a\cdot(1 \quad 0)^T$$

$$X_3 := x0 + a\cdot(0 \quad 1)^T$$

Шаг1

$$F_1 := f(X_1) = 3 \qquad F_2 := f(X_2) = 0 \qquad F_3 := f(X_3) = 10$$

$$h := 3 \qquad l := 2 \qquad s := 1$$

$$fh := f(X_h) = 10 \qquad fl := f(X_l) = 0 \qquad fs := f(X_s) = 3$$

$$xs := \frac{\left(\sum_{i=1}^{3} X_i\right) - X_h}{2} \qquad xs = \begin{pmatrix} 0.5 \\ 0 \end{pmatrix}$$

Отражение

$$xr := xs + \alpha\cdot(xs - X_h) \qquad xr = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$fr := f(xr) \qquad fr = -0.5$$

Сравниваем fr и fl

$$fr = -0.5 \qquad fl = 0 \qquad fr < fl$$

Растяжение

$$xe := xs + \beta \cdot (xs - X_h) \qquad xe = \begin{pmatrix} 1.5 \\ -2 \end{pmatrix}$$

$$fe := f(xe) \qquad fe = 41.5$$

Сравниваем fr и fe

$fe < fr$

Растяжение не удачно

Замена худшей точки

$X_h := xr$

Проверка сходимости

$$Varf := \left[\frac{1}{3} \cdot \sum_{i=1}^{3} \left(f(X_i) - f\left(\frac{X_1 + X_2 + X_3}{3}\right)\right)^2\right]$$

$Varf = 4.529$

Сравнение с требуемой точностью

$Varf > 0.1$

Точность не достигнута

Шаг2

$$X_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad X_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad X_3 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$F_1 := f(X_1) = 3 \quad F_2 := f(X_2) = 0 \quad F_3 := f(X_3) = -0.5$$

$$h := 1 \qquad l := 2 \qquad s := 3$$

$$fh := f(X_h) = 3 \qquad fl := f(X_l) = 0 \qquad fs := f(X_s) = -0.5$$

$$xs := \frac{\left(\sum_{i=1}^{3} X_i\right) - X_h}{2} \qquad xs = \begin{pmatrix} 1 \\ -0.5 \end{pmatrix}$$

Отражение

$$xr := xs + \alpha \cdot (xs - X_h) \qquad xr = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

$$fr := f(xr) \qquad fr = 1$$

Сравниваем fr и fl

$fr = 1$ $\qquad fl = 0$ $\qquad fl < fr$

Сравниваем fr и fs

$fr = 1$ $\qquad fs = -0.5$ $\qquad fr > fs$

Сравниваем fr и fh

$fr = 1$ $\qquad fh = 3$ $\qquad fr < fh$

Замена худшей точки

$$X_h := xr \qquad fh := f(X_h) \qquad fh = 1$$

$$xs := \frac{\left(\sum_{i=1}^{3} X_i\right) - X_h}{2} \qquad xs = \begin{pmatrix} 1 \\ -0.5 \end{pmatrix}$$

Сжатие

$$xc := xs + \gamma \cdot (xs - X_h) \qquad xc = \begin{pmatrix} 0.5 \\ -0.25 \end{pmatrix} \qquad fc := f(xc)$$

Сравниваем fc и fh

$fc = 0.074$ $\qquad fh = 1$ $\qquad fc < fh$

Замена худшей точки

$X_h := xc \qquad fh := f(X_h) \qquad fh = 0.074$

Проверка сходимости

$$Varf := \left[\frac{1}{3}\cdot\sum_{i=1}^{3}\left(f(X_i) - f\left(\frac{X_1 + X_2 + X_3}{3}\right)\right)^2\right]$$

$Varf = 1.143$

Сравнение с требуемой точностью

$Varf > 0.1$

Точность не достигнута

Шаг3

$$X_1 = \begin{pmatrix} 0.5 \\ -0.25 \end{pmatrix} \quad X_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad X_3 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$F_1 := f(X_1) = 0.074 \quad F_2 := f(X_2) = 0 \quad F_3 := f(X_3) = -0.5$

$h := 1 \qquad l := 3 \qquad s := 2$

$fh := f(X_h) = 0.074 \quad fl := f(X_l) = -0.5 \quad fs := f(X_s) = 0$

$$xs := \frac{\left(\sum_{i=1}^{3} X_i\right) - X_h}{2} \qquad xs = \begin{pmatrix} 1 \\ -0.5 \end{pmatrix}$$

Отражение

$$xr := xs + \alpha\cdot(xs - X_h) \qquad xr = \begin{pmatrix} 1.5 \\ -0.75 \end{pmatrix}$$

$fr := f(xr) \qquad fr = -1.488$

Сравниваем fr и fl

$fr = -1.488 \qquad fl = -0.5 \qquad fr < fl$

Растяжение

$xe := xs + \beta\cdot(xs - X_h) \qquad xe = \begin{pmatrix} 2 \\ -1 \end{pmatrix} \qquad fe := f(xe)$

Сравниваем fr и fe

$fr = -1.488 \qquad fe = 1 \qquad fr < fe$

Растяжение не успешно

Замена худшей точки

$X_h := xr \qquad fh := f(X_h) \qquad fh = -1.488$

Проверка сходимости

$$Varf := \left[\frac{1}{3}\cdot\sum_{i=1}^{3}\left(f(X_i) - f\left(\frac{X_1 + X_2 + X_3}{3}\right)\right)^2\right]$$

$Varf = 1.579$

Сравнение с требуемой точностью

$Varf > 0.1$

Точность не достигнута

Шаг4

$X_1 = \begin{pmatrix} 1.5 \\ -0.75 \end{pmatrix} \; X_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad X_3 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

$F_1 := f(X_1) = -1.488 \quad F_2 := f(X_2) = 0 \qquad F_3 := f(X_3) = -0.5$

$h := 2 \qquad l := 1 \qquad s := 3$

$fh := f(X_h) = 0 \qquad fl := f(X_l) = -1.488 \quad fs := f(X_s) = -0.5$

$$xs := \frac{\left(\sum_{i=1}^{3} X_i\right) - X_h}{2} \qquad xs = \begin{pmatrix} 1.25 \\ -0.875 \end{pmatrix}$$

Отражение

$$xr := xs + \alpha \cdot (xs - X_h) \qquad xr = \begin{pmatrix} 1.5 \\ -1.75 \end{pmatrix}$$

$fr := f(xr) \qquad fr = 22.449$

Сравниваем fr и fl

$fr = 22.449 \qquad fl = -1.488 \qquad fr > fl$

Сравниваем fr и fs

$fr = 22.449 \qquad fs = -0.5 \qquad fr > fs$

Сравниваем fr и fh

$fr = 22.449 \qquad fh = 0 \qquad fr > fh$

Сжатие

$$xc := xs + \gamma \cdot (xs - X_h) \qquad xr = \begin{pmatrix} 1.5 \\ -1.75 \end{pmatrix}$$

$fc := f(xc) \qquad fc = 4.461$

Сравниваем fc и fh

$fc = 4.461 \qquad fh = 0 \qquad fc > fh$

Сжатие всего симплекса

$$X_h := \frac{X_h + X_l}{2} \qquad X_s := \frac{X_s + X_l}{2}$$

Проверка сходимости

$$Varf := \left[\frac{1}{3}\cdot\sum_{i=1}^{3}\left(f(X_i) - f\left(\frac{X_1 + X_2 + X_3}{3}\right)\right)^2\right]$$

$Varf = 0.136$

Сравнение с требуемой точностью

$Varf > 0.1$

Точность не достигнута

Шаг5

$X_1 = \begin{pmatrix} 1.5 \\ -0.75 \end{pmatrix}$ $X_2 = \begin{pmatrix} 1.25 \\ -0.375 \end{pmatrix}$ $X_3 = \begin{pmatrix} 1.25 \\ -0.875 \end{pmatrix}$

$F_1 := f(X_1) = -1.488$ $F_2 := f(X_2) = -1.331$ $F_3 := f(X_3) = -1.32$

$h := 3$ $l := 1$ $s := 2$

$fh := f(X_h) = -1.32$ $fl := f(X_l) = -1.488$ $fs := f(X_s) = -1.331$

$$xs := \frac{\left(\sum_{i=1}^{3} X_i\right) - X_h}{2} \qquad xs = \begin{pmatrix} 1.375 \\ -0.563 \end{pmatrix}$$

Отражение

$xr := xs + \alpha\cdot(xs - X_h)$ $xr = \begin{pmatrix} 1.5 \\ -0.25 \end{pmatrix}$

$fr := f(xr)$ $fr = -0.801$

Сравниваем fr и fl

$fr = -0.801$ $fl = -1.488$ $fr > fl$

Сравниваем fr и fs

$fr = -0.801$ $fs = -1.331$ $fr > fs$

Сравниваем fr и fh

$fr = -0.801 \qquad fh = -1.32 \qquad fr > fh$

Сжатие

$xc := xs + \gamma \cdot (xs - X_h) \qquad xc = \begin{pmatrix} 1.438 \\ -0.406 \end{pmatrix} \qquad fc := f(xc)$

Сравниваем fc и fh

$fc = 0.074 \qquad fh = -1.32 \qquad fc > fh$

Сжатие всего симплекса

$$X_h := \frac{X_h + X_l}{2} \qquad X_s := \frac{X_s + X_l}{2}$$

+

Проверка сходимости

$$Varf := \left[\frac{1}{3} \cdot \sum_{i=1}^{3} \left(f(X_i) - f\left(\frac{X_1 + X_2 + X_3}{3}\right)\right)^2\right]$$

$Varf = 0.018$

Сравнение с требуемой точностью

$Varf > 0.1$

Точность достигнута

$X_1 = \begin{pmatrix} 1.5 \\ -0.75 \end{pmatrix} \qquad X_2 = \begin{pmatrix} 1.375 \\ -0.563 \end{pmatrix} \qquad X_3 = \begin{pmatrix} 1.375 \\ -0.813 \end{pmatrix}$

$f(X_1) = -1.488 \qquad f(X_2) = -1.657 \qquad f(X_3) = -1.478$

Ответ

$xmin := X_2 = \begin{pmatrix} 1.375 \\ -0.563 \end{pmatrix}$

$fmin := f(xmin) = -1.657$

Проверка по встроенной функции

$$x := (0 \quad 0)^T$$

$$xmin1 := Minimize(f, x)$$

$$xmin1 = \begin{pmatrix} 1.168 \\ -0.658 \end{pmatrix}$$

$$fmin1 := f(xmin1) = -1.797$$

$$\frac{|fmin - fmin1| \cdot 100}{|fmin1|} = 7.803$$

Вывод: Минимальное значение функции найдено с относительной погрешностью 8%

Индивидуальное задание.

№	f □x1,x2 □	№	f □x1,x2 □
1	2x12 □0.2xx1 2 □ 4x22 □ □3x1 6x2 □ 2	2	$3x_1^2$ □0.7x $x_{1\,2}$ □$5x_2^2$ □$6x_1$ □$4x_2$ □ 2
3	3x12 □0.2xx1 2 □ 4x22 □ □3x1 6x2 □ 2	4	$4x_1^2$ □0.7x $x_{1\,2}$ □$5x_2^2$ □$6x_1$ □$4x_2$ □ 2
5	3x12 □0.2xx1 2 □5x22 □ □3x1 6x2 □ 2	6	3x12 □0.7x x1 2 □5x22 □ □8x1 4x2 □ 2
7	3x12 □0.2xx1 2 □5x22 □ □3x1 2x2 □ 2	8	3x12 □1.7x x1 2 □ □ □ □5x22 6x1 4x2 2
9	3x12 □0.2x x1 2 □5x22 □ □3x1 2x2 □ 2	10	3x12 □1.7x x1 2 □4x22 □ □ □6x1 4x2 2
11	3x12 □0.2xx1 2 □5x22 □ □3x1 2x2 □ 2	12	$3x_1^2$ □1.3x $x_{1\,2}$ □$5x_2^2$ □$6x_1$ □$4x_2$ □ 2
13	3x12 □0.2xx1 2 □5x22 □ □3x1 4x2 □ 2	14	$4x_1^2$ □1.3x $x_{1\,2}$ □$5x_2^2$ □$6x_1$ □$4x_2$ □ 2
15	3x12 □0.4xx1 2 □5x22 □ □3x1 4x2 □ 2	16	$3x_1^2$ □1.2x $x_{1\,2}$ □$5x_2^2$ □$6x_1$ □$4x_2$ □ 2
17	3x12 □0.5xx1 2 □5x22 □ □3x1 4x2 □ 2	18	$3x_1^2$ □1.3x $x_{1\,2}$ □$5x_2^2$ □$3x_1$ □$4x_2$ □ 2
19	3x12 □0.7xx1 2 □5x22 □ □3x1 4x2 □ 2	20	2x12 □1.3xx1 2 □5x22 □ □6x1 4x2 □ 2
21	2x12 □0.5xx1 2 □5x22 □ □3x1 4x2 □ 2	22	3x12 □1.3x x1 2 □6x22 □ □ □6x1 4x2 2

Контрольные вопросы

1. Какие методы безусловной оптимизации называются методами нулевого порядка.
2. Дать определение точкам x_h, x_k, x_s в составе симплекса.
3. В чем заключается операция отражения, растяжения, сжатия.

4. В каком случае осуществляется сжатие всего симплекса к лучшей точке.

5. В чем состоит идея алгоритма Нелдера-Мида.

Лабораторная работа 3. Методы первого порядка. Метод градиентного спуска с дроблением шага

Цель работы: изучить метод градиентного спуска с дроблением шага.

Задание

Для функции двух переменных, определенной во всем пространстве, выполнить процедуру безусловной минимизации из заданной начальной точки градиентным методом с дроблением шага и достижением точности $\varepsilon \le 0.1$ по каждой компоненте градиента.

Краткие теоретические сведения

Задача многомерной безусловной оптимизации формулируется в виде: min $f(x)$, $x \in X$,

где $x=\{x^{(1)}, x^{(2)},\ldots, x^{(n)}\}$ – точка в n-мерном пространстве $X = IR^n$, то есть целевая функция $f(x)=f(x^{(1)},\ldots,f(x^{(n)})$ – функция n аргументов.

Численные методы отыскания минимума, как правило, состоят в построении последовательности точек $\{x_k\}$, удовлетворяющих условию $f(x_0)>f(x_1)>\ldots>f(x_n)>\ldots$. Методы построения таких последовательностей называются методами спуска. В этих методах точки последовательности $\{x_k\}$ вычисляются по формуле:

$$x_{k+1} = x_k + \alpha_k p_k,\ k=0,1,2,\ldots ,$$

где p_k – направление спуска, α_k – длина шага в этом направлении.

Различные методы спуска отличаются друг от друга способами выбора направления спуска p_k и длины шага α_k вдоль этого направления. Алгоритмы безусловной минимизации принято делить на классы, в зависимости от максимального порядка производных минимизируемой функции, вычисление которых предполагается.

Так, методы, использующие только значения самой целевой функции, относят к методам нулевого порядка (иногда их называют также методами прямого поиска); если, кроме того, требуется вычисление первых производных минимизируемой функции, то мы имеем дело с методами первого порядка; если же дополнительно используются вторые производные, то это методы второго порядка и т. д.

Градиентные методы. Общая схема градиентного спуска

Как известно, градиент функции в некоторой точке x_k направлен в сторону наискорейшего локального возрастания функции и перпендикулярен линии уровня (поверхность постоянного значения функции f(x), проходящей через точку x_k). Вектор, противоположный градиенту $f'(x_k)$, называется антиградиентом, который направлен в сторону наискорейшего убывания функции f(x). Выбирая в качестве направления спуска p_k антиградиент $-f'(x_k)$ в точке x_k, мы приходим к итерационному процессу вида:

$$x_{k+1} = x_k - \alpha_k f'(x_k),\ \alpha_k>0,\ k = 0,1,2,\dots .$$

В координатной форме этот процесс записывается следующим образом:

$$x_{k+1}^{(i)} = x_k^{(i)} - \alpha_k \frac{\partial f}{\partial x_i}(x_k),\ i = 1,2,...,n.$$

Все итерационные процессы, в которых направление движения на каждом шаге совпадает с антиградиентом функции, называются градиентными методами. Они отличаются друг от друга только способом выбора шага α_k. Существует много различных способов выбора α_k, но наиболее распространены: метод с постоянным шагом, метод с дроблением шага и метод наискорейшего спуска.

Градиентный метод с постоянным шагом

Основная проблема в градиентных методах – это выбор шага α_k. Достаточно малый шаг α_k обеспечивает убывание функции, то есть выполнение неравенства:

$$f(x_k - \alpha_k f'(x_k)) < f(x_k),$$

но может привести к неприемлемо большому количеству итераций, необходимых для достижения точки минимума. С другой стороны, слишком большой шаг может вызвать неожиданный рост функции (невыполнение условия убывания) либо привести к колебаниям около точки минимума. Однако проверка условия убывания на каждой итерации является довольно трудоемкой, поэтому в методе градиентного спуска с постоянным шагом задают $\alpha=\alpha_k$ постоянным и достаточно малым, чтобы можно было использовать этот шаг на любой итерации. При этом приходится мириться с возможно большим количеством итераций. Утешением является лишь то, что трудоемкость каждой итерации, в этом случае, минимальна (нужно вычислять только градиент $f'(x_k)$).

Алгоритм градиентного метода с постоянным шагом Шаг 1.

Задаются начальное приближение x_0, постоянный шаг α, условия останова алгоритма ε_3. Вычисляется значение градиента $f'(x_k)$ – направление поиска. Присваивается k = 0.

Шаг 2.

Определяется точка очередного эксперимента: $x_{k+1} = x_k - \alpha f'(x_k)$,

или, в координатной форме:

$$x_{k+1}^{(i)} = x_k^{(i)} - \alpha \frac{\partial f}{\partial x_i}(x_{k(1)},...,x_{k(n)}),\ где\ i = 1,2,...,n.$$

Шаг 3.

Вычисляется значение градиента в точке $x_{к+1}$: $\nabla f(x_{k+1})$, или, в координатной форме:

$$\left\{ \frac{\partial f}{\partial x_1}(x_{(k+1)1},\ldots,x_{(k+1)n}),\ldots,\frac{\partial f}{\partial x_n}(x_{(k+1)1},\ldots,x_{(k+1)n}) \right\}$$

Шаг 4.

Если $||\nabla f(x_{k+1})|| \le \varepsilon_3$, то поиск заканчивается, при этом:

$\tilde{x} = x_{k+1}, \tilde{y} = f(x_{k+1})$.

Иначе k =k+1 и переходим к шагу 2.

Градиентный метод с дроблением шага

В методе градиентного спуска с дроблением шага величина шага $\alpha_к$ выбирается так, чтобы было выполнено неравенство:

$$f(x_k-\alpha_k \nabla f(x_k))-f(x_k) \le -\delta\alpha_k||\nabla f(x_k)||^2,$$

где $0<\delta<1$ – произвольно выбранная постоянная (одна и та же для всех итераций). Это требование на выбор шага $\alpha_к$ более жесткое, чем условие убывания, но имеет тот же смысл: функция должна убывать от итерации к итерации. Однако при выполнении неравенства функция будет уменьшаться на гарантированную величину, определяемую правой частью неравенства.

Процесс выбора шага протекает следующим образом. Выбираем число $\alpha>0$, одно и то же для всех итераций. На к-й итерации проверяем выполнение неравенства при $\alpha_к=\alpha$. Если оно выполнено, полагаем $\alpha_к=\alpha$ и переходим к следующей итерации. Если нет, то шаг $\alpha_к$ дробим, например уменьшаем каждый раз в два раза, до тех пор, пока оно не выполнится.

Алгоритм градиентного метода с дроблением шага Шаг 1.

Задаются x0, ε3, δ и начальное значение шага α. Вычисляется значение градиента $\nabla f(x_0)$ – направление поиска. Присваивается к=0.

Шаг 2.

Проверяется условие: $f(xk-\alpha \nabla f(x_k)) \le -\delta\alpha||\nabla f(x_k)||2$. Если выполняется, то переходим к шагу 3, иначе дробим значение α ($\alpha=\alpha/2$) и повторяем шаг 2. Шаг 3. Определяется точка очередного эксперимента: xк+1 = xк - $\alpha \nabla f(x_k)$. Шаг 4. Вычисляется значение градиента в точке xк+1: $\nabla f(x_{k+1})$. Шаг 5.

Если $||\nabla f(x_{k+1})|| \le \varepsilon$3, то поиск заканчивается, при этом:

$$\tilde{x} = x_{k+1}, \tilde{y} = f(x_{k+1}).$$

Иначе к=к+1 и переходим к шагу 2.

Метод наискорейшего спуска

В градиентном методе с постоянным шагом величина шага, обеспечивающая убывание функции f(x) от итерации к итерации, оказывается очень малой, что приводит к необходимости проводить большое количество итерации для достижения точки минимума. Поэтому методы спуска с переменным шагом являются более экономными.

Алгоритм, на каждой итерации которого шаг $\lambda_к$ выбирается из условия минимума функции f(x) в направлении движения, то есть:

$$f(x_k - \lambda_k f'(x_k)) = \min_{\lambda \geq 0} f(x_k - \lambda f'(x_k)).$$

называется методом наискорейшего спуска. Разумеется, этот способ выбора $\lambda_к$ сложнее ранее рассмотренных вариантов.

Реализация метода наискорейшего спуска предполагает решение на каждой итерации довольно трудоемкой вспомогательной задачи одномерной минимизации. Как правило, метод наискорейшего спуска, тем не менее, дает выигрыш в числе машинных операций, поскольку обеспечивает движение с самым выгодным шагом, ибо решение задачи одномерной минимизации связано с дополнительными вычислениями только самой функции f(x), тогда как основное машинное время тратится на вычисление ее градиента $f'(x_k)$.

Следует иметь в виду, что одномерную минимизацию можно производить любым методом одномерной оптимизации, что порождает различные варианты метода наискорейшего спуска.

Алгоритм Метода наискорейшего спуска Шаг 1.

Задаются x_0, ε_3. Вычисляется градиент $f'(x_0)$, направление поиска. Присваивается k = 0.

Шаг 2.

Определяется точка очередного эксперимента: $x_{к+1} = x_к - \lambda_к f'(x_k)$,

$$\lambda_k = \arg\min_{\lambda \geq 0} f(x_k - \lambda f'(x_k))$$

где $\lambda_к$ – минимум задачи одномерной минимизации:

Шаг 3.

Вычисляется значение градиента в точке $x_{к+1}$: $f'(x_{k+1})$.

Шаг 4.

Если $\|f'(x_{k+1})\| \le \varepsilon_3$, то поиск точки минимума заканчивается и полагается:

$$\tilde{x} = x_{k+1}, \tilde{y} = f(x_{k+1}).$$

Иначе k = k+1 и переход к шагу 2.

Пример выполнения

Минимизируемая функция

$$F(x,y) := 4 \cdot x^4 + 3 \cdot x \cdot y + 3 \cdot y^2 - 5 \cdot x + 4 \cdot y - 12$$

Рассчитываем градиент

$$F1(x,y) := \frac{d}{dx} F(x,y)$$

$$F1(x,y) \text{ simplify } \rightarrow 16 \cdot x^3 + 3 \cdot y - 5$$

$$F2(x,y) := \frac{d}{dy} F(x,y)$$

$$F2(x,y) \text{ simplify } \rightarrow 3 \cdot x + 6 \cdot y + 4$$

Переходим к векторному аргументу

$$f(x) := F(x_1, x_2)$$

$$\text{gradf}(x) := \left(F1(x_1, x_2) \quad F2(x_1, x_2) \right)^T$$

Данная начальная точка поиска

$$X_1 := (0 \quad 0)^T \qquad f(X_1) = -12$$

Заданная точность поиска

$\varepsilon := 0.1$

Заданный постоянный шаг(который дробится на иттерациях)

$\alpha := 0.1$

Заданная величина δ алгоритма

$\delta := 0.5$

Шаг1

Дробим шаг в такой степени чтобы обеспечить выполнение неравенства

$$f(x_k - \alpha_k f'(x_k)) - f(x_k) \le -\delta \alpha_k \| f'(x_k) \|^2,$$

$$f(X_1 - \alpha \cdot \mathrm{gradf}(X_1)) - f(X_1) + \delta \cdot \alpha \cdot \mathrm{gradf}(X_1)^T \cdot \mathrm{gradf}(X_1) = -1.92$$

Шаг α подходит

$X_2 := X_1 - \alpha \cdot \mathrm{gradf}(X_1)$

$X_2 = \begin{pmatrix} 0.5 \\ -0.4 \end{pmatrix}$ $\qquad f(X_2) = -15.97$

$f(X_2) - f(X_1) = -3.97$

Проверка сходимости $\quad \mathrm{gradf}(X_2) = \begin{pmatrix} -4.2 \\ 3.1 \end{pmatrix}$

В градиенте есть компоненты большие 0.1 сходимость не достигнута

Шаг2

Дробим шаг в такой степени чтобы обеспечить выполнение неравенства

$$f(x_k - \alpha_k f'(x_k)) - f(x_k) \leq -\delta\alpha_k \| f'(x_k) \|^2,$$

$$f(X_2 - \alpha \cdot \mathrm{gradf}(X_2)) - f(X_2) + \delta \cdot \alpha \cdot \mathrm{gradf}(X_2)^T \cdot \mathrm{gradf}(X_2) = 0.311$$

Шаг α не подходит

$$\alpha 1 := \frac{\alpha}{2}$$

$$f(X_2 - \alpha 1 \cdot \mathrm{gradf}(X_2)) - f(X_2) + \delta \cdot \alpha 1 \cdot \mathrm{gradf}(X_2)^T \cdot \mathrm{gradf}(X_2) = -0.36$$

Шаг $\alpha 1$ подходит

$$X_3 := X_2 - \alpha 1 \cdot \mathrm{gradf}(X_2)$$

$X_3 = \begin{pmatrix} 0.71 \\ -0.555 \end{pmatrix} \qquad f(X_3) = -17.012$

Проверка сходимости $\quad \mathrm{gradf}(X_3) = \begin{pmatrix} -0.938 \\ 2.8 \end{pmatrix}$

В градиенте есть компоненты большие 0.1 сходимость не достигнута

Шаг3

Дробим шаг в такой степени чтобы обеспечить выполнение неравенства

$$f(x_k - \alpha_k f'(x_k)) - f(x_k) \leq -\delta\alpha_k \| f'(x_k) \|^2,$$

$$f(X_3 - \alpha \cdot gradf(X_3)) - f(X_3) + \delta \cdot \alpha \cdot gradf(X_3)^T \cdot gradf(X_3) = -0.163$$

Шаг α подходит

$$X_4 := X_3 - \alpha \cdot gradf(X_3)$$

$$X_4 = \begin{pmatrix} 0.804 \\ -0.835 \end{pmatrix} \qquad f(X_4) = -17.611$$

Проверка сходимости $\quad gradf(X_4) = \begin{pmatrix} 0.806 \\ 1.402 \end{pmatrix}$

В градиенте есть компоненты большие 0.1 сходимость не достигнута

Шаг4

Дробим шаг в такой степени чтобы обеспечить выполнение неравенства

$$f(x_k-\alpha_k f'(x_k))-f(x_k)\le-\delta\alpha_k\|f'(x_k)\|^2,$$

$$f\left(X_4 - \alpha \cdot gradf\left(X_4\right)\right) - f\left(X_4\right) + \delta \cdot \alpha \cdot gradf\left(X_4\right)^T \cdot gradf\left(X_4\right) = 0.056$$

Шаг α не подходит

$$\alpha 1 := \frac{\alpha}{2}$$

$$f\left(X_4 - \alpha 1 \cdot gradf\left(X_4\right)\right) - f\left(X_4\right) + \delta \cdot \alpha 1 \cdot gradf\left(X_4\right)^T \cdot gradf\left(X_4\right) = -0.018$$

Шаг $\alpha 1$ подходит

$$X_5 := X_4 - \alpha 1 \cdot gradf\left(X_4\right)$$

$$X_5 = \begin{pmatrix} 0.764 \\ -0.905 \end{pmatrix} \qquad f\left(X_5\right) = -17.694$$

Проверка сходимости $\quad gradf\left(X_5\right) = \begin{pmatrix} -0.592 \\ 0.86 \end{pmatrix}$

В градиенте есть компоненты большие 0.1 сходимость не достигнута

Шаг5

Дробим шаг в такой степени чтобы обеспечить выполнение неравенства

$$f(x_k-\alpha_k f'(x_k))-f(x_k)\le-\delta\alpha_k\|f'(x_k)\|^2,$$

$$f(X_5-\alpha\cdot gradf(X_5))-f(X_5)+\delta\cdot\alpha\cdot gradf(X_5)^T\cdot gradf(X_5)=4.059\times10^{-3}$$

Шаг α не подходит

$$\alpha1:=\frac{\alpha}{2}$$

$$f(X_5-\alpha1\cdot gradf(X_5))-f(X_5)+\delta\cdot\alpha1\cdot gradf(X_5)^T\cdot gradf(X_5)=-0.013$$

Шаг $\alpha1$ подходит

$$X_6:=X_5-\alpha1\cdot gradf(X_5)$$

$$X_6=\begin{pmatrix}0.793\\-0.948\end{pmatrix}\qquad f(X_6)=-17.734$$

Проверка сходимости $\quad gradf(X_6)=\begin{pmatrix}0.14\\0.691\end{pmatrix}$

В градиенте есть компоненты большие 0.1 сходимость не достигнута

Шаг6

Дробим шаг в такой степени чтобы обеспечить выполнение неравенства

$$f(x_k-\alpha_k f'(x_k))-f(x_k)\le-\delta\alpha_k\| f'(x_k)\|^2,$$

$$f(X_6-\alpha\cdot gradf(X_6))-f(X_6)+\delta\cdot\alpha\cdot gradf(X_6)^T\cdot gradf(X_6)=-4.7\times10^{-3}$$

Шаг α подходит

$$X_7 := X_6-\alpha\cdot gradf(X_6)$$

$$X_7=\begin{pmatrix}0.779\\-1.017\end{pmatrix} \qquad f(X_7)=-17.764$$

Проверка сходимости $\quad gradf(X_7)=\begin{pmatrix}-0.483\\0.234\end{pmatrix}$

В градиенте есть компоненты большие 0.1 сходимость

Шаг8

Дробим шаг в такой степени чтобы обеспечить выполнение неравенства

$$f(x_k-\alpha_k f'(x_k))-f(x_k)\le-\delta\alpha_k\| f'(x_k)\|^2,$$

$$f(X_7-\alpha\cdot gradf(X_7))-f(X_7)+\delta\cdot\alpha\cdot gradf(X_7)^T\cdot gradf(X_7)=0.019$$

Шаг α не подходит

$$\alpha1 := \frac{\alpha}{2}$$

$$f(X_7-\alpha1\cdot gradf(X_7))-f(X_7)+\delta\cdot\alpha1\cdot gradf(X_7)^T\cdot gradf(X_7)=1.03\times10^{-3}$$

Шаг α1 не подходит

$$\alpha2 := \frac{\alpha1}{2}$$

$$f(X_7-\alpha2\cdot gradf(X_7))-f(X_7)+\delta\cdot\alpha2\cdot gradf(X_7)^T\cdot gradf(X_7)=-1.565\times10^{-3}$$

Шаг α2 подходит

$X_8 := X_7 - \alpha 2 \cdot \mathrm{gradf}(X_7)$

$X_8 = \begin{pmatrix} 0.791 \\ -1.023 \end{pmatrix}$ $f(X_8) = -17.769$

Проверка сходимости $\mathrm{gradf}(X_8) = \begin{pmatrix} -0.143 \\ 0.235 \end{pmatrix}$

В градиенте есть компоненты большие 0.1 сходимость не достигнута

+

Шаг9

Дробим шаг в такой степени чтобы обеспечить выполнение неравенства

$f(x_k - \alpha_k f'(x_k)) - f(x_k) \le -\delta\alpha_k \| f'(x_k) \|^2,$

$f(X_8 - \alpha \cdot \mathrm{gradf}(X_8)) - f(X_8) + \delta \cdot \alpha \cdot \mathrm{gradf}(X_8)^T \cdot \mathrm{gradf}(X_8) = -2.779 \times 10^{-5}$

Шаг α подходит

$X_9 := X_8 - \alpha \cdot \mathrm{gradf}(X_8)$

$X_9 = \begin{pmatrix} 0.806 \\ -1.047 \end{pmatrix}$ $f(X_9) = -17.773$

Проверка сходимости $\mathrm{gradf}(X_9) = \begin{pmatrix} 0.224 \\ 0.137 \end{pmatrix}$

В градиенте есть компоненты большие 0.1 сходимость не достигнута

Шаг10

Дробим шаг в такой степени чтобы обеспечить выполнение неравенства

$f(x_k - \alpha_k f'(x_k)) - f(x_k) \le -\delta\alpha_k \| f'(x_k) \|^2,$

$f(X_9 - \alpha \cdot \mathrm{gradf}(X_9)) - f(X_9) + \delta \cdot \alpha \cdot \mathrm{gradf}(X_9)^T \cdot \mathrm{gradf}(X_9) = 5.72 \times 10^{-3}$

Шаг α не подходит

$$\alpha 1 := \frac{\alpha}{2}$$

$$f\left(X_9 - \alpha 1 \cdot gradf\left(X_9\right)\right) - f\left(X_9\right) + \delta \cdot \alpha 1 \cdot gradf\left(X_9\right)^T \cdot gradf\left(X_9\right) = 5.843 \times 10^{-4}$$

Шаг α1 не подходит

$$\alpha 2 := \frac{\alpha 2}{2}$$

$$f\left(X_9 - \alpha 2 \cdot gradf\left(X_9\right)\right) - f\left(X_9\right) + \delta \cdot \alpha 2 \cdot gradf\left(X_9\right)^T \cdot gradf\left(X_9\right) = -2.865 \times 10^{-4}$$

Шагα2 подходи

$$X_{10} := X_9 - \alpha 2 \cdot gradf\left(X_9\right)$$

$$X_{10} = \begin{pmatrix} 0.803 \\ -1.048 \end{pmatrix} \qquad f\left(X_{10}\right) = -17.774$$

Проверка сходимости $\quad gradf\left(X_{10}\right) = \begin{pmatrix} 0.132 \\ 0.118 \end{pmatrix}$

В градиенте есть компоненты большие 0.1 сходимость не достигнута

Шаг11

Дробим шаг в такой степени чтобы обеспечить выполнение неравенства

$$f(x_k - \alpha_k f'(x_k)) - f(x_k) \le -\delta \alpha_k \| f'(x_k) \|^2,$$

$$f\left(X_{10} - \alpha \cdot gradf\left(X_{10}\right)\right) - f\left(X_{10}\right) + \delta \cdot \alpha \cdot gradf\left(X_{10}\right)^T \cdot gradf\left(X_{10}\right) = 1.986 \times 10^{-3}$$

Шаг α не подходит

$$\alpha 1 := \frac{\alpha}{2}$$

$$f\left(X_{10} - \alpha 1 \cdot gradf\left(X_{10}\right)\right) - f\left(X_{10}\right) + \delta \cdot \alpha 1 \cdot gradf\left(X_{10}\right)^T \cdot gradf\left(X_{10}\right) = 1.066 \times 10^{-4}$$

Шаг α1 не подходит

$$\alpha 2 := \frac{\alpha 2}{2}$$

$$f\left(X_{10} - \alpha 2 \cdot gradf\left(X_{10}\right)\right) - f\left(X_{10}\right) + \delta \cdot \alpha 2 \cdot gradf\left(X_{10}\right)^T \cdot gradf\left(X_{10}\right) = -8.435 \times 10^{-5}$$

Шагα2 подходи

$$X_{11} := X_{10} - \alpha 2 \cdot \mathrm{gradf}(X_{10})$$

$$X_{11} = \begin{pmatrix} 0.802 \\ -1.049 \end{pmatrix} \qquad f(X_{11}) = -17.774$$

Проверка сходимости $\quad \mathrm{gradf}(X_{11}) = \begin{pmatrix} 0.104 \\ 0.112 \end{pmatrix}$

В градиенте есть компоненты большие 0.1 сходимость не достигнута

Шаг12

Дробим шаг в такой степени чтобы обеспечить выполнение неравенства

$$f(x_k - \alpha_k f'(x_k)) - f(x_k) \le -\delta \alpha_k \| f'(x_k) \|^2,$$

$$f(X_{11} - \alpha \cdot \mathrm{gradf}(X_{11})) - f(X_{11}) + \delta \cdot \alpha \cdot \mathrm{gradf}(X_{11})^T \cdot \mathrm{gradf}(X_{11}) = 1.222 \times 10^{-3}$$

Шаг α не подходит

$$\alpha 1 := \frac{\alpha}{2}$$

$$f(X_{11} - \alpha 1 \cdot \mathrm{gradf}(X_{11})) - f(X_{11}) + \delta \cdot \alpha 1 \cdot \mathrm{gradf}(X_{11})^T \cdot \mathrm{gradf}(X_{11}) = 1.583 \times 10^{-}$$

Шаг $\alpha 1$ не подходит

$$\alpha 2 := \frac{\alpha 2}{2}$$

$$f(X_{11} - \alpha 2 \cdot \mathrm{gradf}(X_{11})) - f(X_{11}) + \delta \cdot \alpha 2 \cdot \mathrm{gradf}(X_{11})^T \cdot \mathrm{gradf}(X_{11}) = -3.41 \times 10^{-5}$$

Шаг $\alpha 2$ подходит

$$X_{12} := X_{11} - \alpha 2 \cdot \mathrm{gradf}(X_{11})$$

$$X_{12} = \begin{pmatrix} 0.802 \\ -1.049 \end{pmatrix} \qquad f(X_{12}) = -17.774$$

Проверка сходимости $\quad \mathrm{gradf}(X_{12}) = \begin{pmatrix} 0.093 \\ 0.108 \end{pmatrix}$

В градиенте есть компоненты большие 0.1 сходимость не достигнута

Шаг13

Дробим шаг в такой степени чтобы обеспечить выполнение неравенства

$$f(x_k-\alpha_k f'(x_k))-f(x_k)\le-\delta\alpha_k\| f'(x_k)\|^2,$$

$$f(X_{12}-\alpha\cdot gradf(X_{12}))-f(X_{12})+\delta\cdot\alpha\cdot gradf(X_{12})^T\cdot gradf(X_{12}) = 9.642\times10^{-4}$$

Шаг α не подходит

$$\alpha1 := \frac{\alpha}{2}$$

$$f(X_{12}-\alpha1\cdot gradf(X_{12}))-f(X_{12})+\delta\cdot\alpha1\cdot gradf(X_{12})^T\cdot gradf(X_{12}) = -1.337\times10^{-5}$$

Шаг $\alpha1$ подходит

$$X_{13} := X_{12}-\alpha1\cdot gradf(X_{12})$$

$$X_{13} = \begin{pmatrix} 0.797 \\ -1.055 \end{pmatrix} \qquad f(X_{13}) = -17.774$$

Проверка сходимости $\quad gradf(X_{13}) = \begin{pmatrix} -0.066 \\ 0.062 \end{pmatrix}$

Все компоненты градиента по модулю меньше заданной точности 0.1 достигнута сходимость

Ответ

$$xmin := X_{13} = \begin{pmatrix} 0.797 \\ -1.055 \end{pmatrix}$$

$$fmin := f(xmin) = -17.774$$

Проверка по встроенной функции

$$x := (0 \;\; 0)^T$$

$$xmin1 := Minimize(f, x)$$

$$xmin1 = \begin{pmatrix} 0.8 \\ -1.067 \end{pmatrix}$$

$$fmin1 := f(xmin1) = -17.775$$

$$d := \sqrt{(xmin - xmin1)^T \cdot (xmin - xmin1)}$$

$$d = 0.012$$

Вывод: Точка минимума найдена с абсолютной погрешностью 0.012

Индивидуальные задания

№	$f\square x_1,x_2\square$	№	$f\square x_1,x_2\square$
1	$2\square x_1^2 \square 0.2\square x_1 \square x_2 \square 4\square x_2^2 \square 3\square x_1 \square 6\square x_2 \square 2$	2	$3\square x_1^2 \square 0.7\square x_1 \square x_2 \square 5\square x_2^2 \square 6\square x_1 \square 4\square x_2 \square 2$
3	$3\square x_1^2 \square 0.2\square x_1 \square x_2 \square 4\square x_2^2 \square 3\square x_1 \square 6\square x_2 \square 2$	4	$4\square x_1^2 \square 0.7\square x_1 \square x_2 \square 5\square x_2^2 \square 6\square x_1 \square 4\square x_2 \square 2$
5	$3\square x_1^2 \square 0.2\square x_1 \square x_2 \square 5\square x_2^2 \square 3\square x_1 \square 6\square x_2 \square 2$	6	$3\square x_1^2 \square 0.7\square x_1 \square x_2 \square 5\square x_2^2 \square 8\square x_1 \square 4\square x_2 \square 2$
7	$3\square x_1^2 \square 0.2\square x_1 \square x_2 \square 5\square x_2^2 \square 3\square x_1 \square 2\square x_2 \square 2$	8	$3\square x_1^2 \square 1.7\square x_1 \square x_2 \square 5\square x_2^2 \square 6\square x_1 \square 4\square x_2 \square 2$
9	$3\square x_1^2 \square 0.2\square x_1 \square x_2 \square 5\square x_2^2 \square 3\square x_1 \square 2\square x_2 \square 2$	10	$3\square x_1^2 \square 1.7\square x_1 \square x_2 \square 4\square x_2^2 \square 6\square x_1 \square 4\square x_2 \square 2$
11	$3\square x_1^2 \square 0.2\square x_1 \square x_2 \square 5\square x_2^2 \square 3\square x_1 \square 2\square x_2 \square 2$	12	$3\square x_1^2 \square 1.3\square x_1 \square x_2 \square 5\square x_2^2 \square 6\square x_1 \square 4\square x_2 \square 2$
13	$3\square x_1^2 \square 0.2\square x_1 \square x_2 \square 5\square x_2^2 \square 3\square x_1 \square 4\square x_2 \square 2$	14	$4\square x_1^2 \square 1.3\square x_1 \square x_2 \square 5\square x_2^2 \square 6\square x_1 \square 4\square x_2 \square 2$
15	$3\square x_1^2 \square 0.4\square x_1 \square x_2 \square 5\square x_2^2 \square 3\square x_1 \square 4\square x_2 \square 2$	16	$3\square x_1^2 \square 1.2\square x_1 \square x_2 \square 5\square x_2^2 \square 6\square x_1 \square 4\square x_2 \square 2$
17	$3\square x_1^2 \square 0.5\square x_1 \square x_2 \square 5\square x_2^2 \square 3\square x_1 \square 4\square x_2 \square 2$	18	$3\square x_1^2 \square 1.3\square x_1 \square x_2 \square 5\square x_2^2 \square 3\square x_1 \square 4\square x_2 \square 2$
19	$3\square x_1^2 \square 0.7\square x_1 \square x_2 \square 5\square x_2^2 \square 3\square x_1 \square 4\square x_2 \square 2$	20	$2\square x_1^2 \square 1.3\square x_1 \square x_2 \square 5\square x_2^2 \square 6\square x_1 \square 4\square x_2 \square 2$
21	$2\square x_1^2 \square 0.5\square x_1 \square x_2 \square 5\square x_2^2 \square 3\square x_1 \square 4\square x_2 \square 2$	22	$3\square x_1^2 \square 1.3\square x_1 \square x_2 \square 6\square x_2^2 \square 6\square x_1 \square 4\square x_2 \square 2$

Контрольные вопросы

1. Что называется градиентом функции нескольких переменных, антиградиентом.
2. Описать общую схему градиентных методов.
3. Что означает дробление шага в градиентном методе.

4. Каков критерий останова вычислений в градиентных методах.

5. В чем основная идея метода наискорейшего спуска.

Лабораторная работа 4. Методы второго порядка. Метод Ньютона

Цель работы: Изучить метод Ньютона в задачах минимизации без ограничений.

Задание. Применить метод Ньютона для минимизации функции двух переменных из заданной начальной точки (1,0) и до достижения заданной точности нахождения экстремума 0.001 по всем компонентам градиента в текущей точке.

Краткие теоретические сведения

Метод оптимизации Ньютона предназначен для решения многомерных задач локальной безусловной оптимизации. Метод относится, с одной стороны, к классу методов оптимизации *k*-го порядка, где $k = 2$ (т.е. к классу методов оптимизации второго порядка), а с другой стороны – к классу детерминированных методов оптимизации. Одна итерация метода требует вычисления в текущей точке компонент вектора градиента минимизируемой функции, а также компонент матрицы Гессе в этой точке.

Матрицей Гессе $H(x)$ функции $f(x)$ называется $(n*n)$-матрица вторых частных производных этой функции. Здесь n – размерность вектора x.

Положим, что функция $f(x)$ всюду дважды дифференцируема в n-мерном евклидовом пространстве R^n.

Рассмотрим следующую многомерную задачу локальной безусловной оптимизации: найти минимум критерия оптимальности $f(x)$, определенного в nмерном евклидовом пространстве R^n, $\min_{x \in R^n} f(x) = f(x^*) = f^*$.

Обоснование метода оптимизации Ньютона

Рассмотрим первые три члена разложения функции $f(x)$ в ряд Тейлора в окрестности точки x^r:

$$\tilde{f}^r(x) = f(x^r) + (\nabla f(x^r),(x - x^r)) + \frac{1}{2}(H(x^r)(x - x^r),(x - x^r)).$$

$$f(x) \approx \tilde{f}^r(x)$$

Здесь $H(x)$ – матрица Гессе функции $f(x)$. Из следует, что градиент функции $\tilde{f}^r(x)$ равен

$$\nabla \tilde{f}^r(x) = \nabla f(x^r) + H(x^r)(x - x^r).$$

Если матрица Гессе $H(x^r)$ положительно определена, то функция $\tilde{f}^{\,r}(x)$ достигает минимума в точке, в которой градиент этой функции равен нулевому вектору.

Таким образом, в точке x^{r+1} минимума функции $\tilde{f}^{\,r}(x)$ справедливо равенство

$$\nabla f(x^r) + H(x^r)(x^{r+1} - x^r) = 0,$$

где 0 – n-мерный вектор нулей. Отсюда получаем итерационную формулу

$$x^{r+1} = x^r - H^{-1}(x^r)\nabla f(x^r) \qquad (1)$$

для отыскания очередного приближения к точке минимума функции $f(x)$.

Отметим трудности, которые могут возникать при использовании итерационной формулы (1):

1. Если размерность пространства R^n велика, то обращение на каждой итерации матрицы Гессе $H(x^r)$ может потребовать значительных вычислительных ресурсов.

2. Значение минимизируемой функции $f(x)$ в точке $x^{r+1} = x^r + p^r$ может превышать значение функции в предыдущей точке x^r вследствие того, что направление p^r ведет к уменьшению $f(x)$, но величина шага слишком велика.

3. Направление спуска, определяемое вектором $p^r = -H^{-1}(x^r)\nabla f(x^r)$, ведет к убыванию целевой функции только при положительной определенности матрицы Гессе $H(x^r)$. Это приводит к тому, что на каждой итерации необходимы вычислительные затраты на проверку обусловленности этой матрицы. Указанная матрица может быть плохо обусловленной. Более того, указанная матрица может быть вырожденной и поэтому не иметь обратной матрицы.

Вследствие этих трудностей итерационная формула (1) в «чистом» виде не используется в вычислительной практике.

Для того чтобы избежать обращения матрицы Гессе, на практике вектор p^r находят обычно из следующей системы линейных алгебраических уравнений (СЛАУ), вытекающей из равенства (1):

$$H(x^r)p^r = -\nabla f(x^r). \qquad (2)$$

СЛАУ (2) может быть решена различными численными методами (например, прямыми методами, итерационными методами).

Величина шага в направлении p^r, которая приводит к убыванию функции $f(x)$, может быть обеспечена путем добавления в итерационную формулу (1) коэффициента $\alpha^r \in (0;1]$, т.е. путем использования вместо формулы (1) итерационной формулы

$$x_{r+1} = x_r - \alpha_r H^{-1}(x_r)\nabla f(x_r) = x_r + \alpha_r p_r, \qquad (3)$$

где коэффициент α^r выбирают тем или иным способом так, чтобы обеспечить условие $f(x^{r+1}) < f(x^r)$.

Для того чтобы направление спуска независимо от определенности матрицы Гессе $H(x^r)$ вело к убыванию функции $f(x)$, в качестве вектора p^r можно использовать вектор

$$p^r = -(\mu^r I + H(x^r))^{-1} \nabla f(x^r), \tag{4}$$

где I – $(n*n)$ – единичная матрица, а μ^r – параметр, выбираемый так, чтобы матрица $\mu^r I + H(x^r)$ являлась положительно определенной.

Алгоритм метода оптимизации Ньютона

Рассмотрим алгоритм одной из модификаций метода оптимизации Ньютона, в которой используется итерационная формула (3) и вектор p^r находят путем решения на каждой итерации СЛАУ (2).

Шаг 1. Задаем начальную точку x^0, начальную величину шага $\alpha^0 = 1$ и коэффициент дробления шага $\beta \in (0;1]$. Полагаем счетчик числа итераций $r = 0$.

Шаг 2. Вычисляем в точке x^r вектор градиента $\nabla f(x^r)$ и матрицу Гессе $H(x^r)$.

Шаг 3. Решаем СЛАУ (2) и находим вектор p^r.

Шаг 4. По формуле (3) вычисляем компоненты вектора x^{r+1}.

Шаг 5. Вычисляем величину $f(x^{r+1})$ – значение функции $f(x)$ в точке x^{r+1}.

Шаг 6. Проверяем условие окончания поиска (5). Если условие окончания поиска выполнено, то полагаем $x^* = x^{r+1}$ и завершаем итерации. Иначе – переходим к следующему шагу.

Шаг 7. Если $f(x^{r+1}) < f(x^r)$, то полагаем $r = r + 1$ и переходим к шагу 2. Иначе – фиксированное число раз полагаем $\alpha^r = \beta\alpha^r$ и переходим к шагу 4.

В качестве условия окончания поиска можно использовать одно из стандартных условий окончания итераций:

$$\|x^{r+1} - x^r\| \le \varepsilon_x, \quad |f(x^{r+1}) - f(x^r)| \le \varepsilon_f$$

или условие $\|\nabla f^r\| \le \varepsilon_\nabla$, (5)

где ε_∇ – константа, определяющая требуемую точность решения по градиенту функции $f(x)$.

Пример выполнения

Минимизируемая функция

$$F(x,y) := 4\cdot x^4 + 3\cdot x\cdot y + 3\cdot y^2 - 5\cdot x + 4\cdot y - 12$$

Рассчитываем градиент

$$F1(x,y) := \frac{d}{dx}F(x,y)$$

$$F1(x,y) \text{ simplify } \rightarrow 16\cdot x^3 + 3\cdot y - 5$$

$$F2(x,y) := \frac{d}{dy}F(x,y)$$

$$H(x,y) := \begin{bmatrix} \frac{d}{dx}\left(\frac{d}{dx}F(x,y)\right) & \frac{d}{dx}\left(\frac{d}{dy}F(x,y)\right) \\ \frac{d}{dx}\left(\frac{d}{dy}F(x,y)\right) & \frac{d}{dy}\left(\frac{d}{dy}F(x,y)\right) \end{bmatrix}$$

$$H(x,y) \text{ simplify } \rightarrow \begin{pmatrix} 48\cdot x^2 & 3 \\ 3 & 6 \end{pmatrix}$$

Переходим к векторному аргументу

$$f(x) := F(x_1, x_2)$$

$$gradf(x) := \left(F1(x_1, x_2) \quad F2(x_1, x_2) \right)^T$$

$$hf(x) := H(x_1, x_2)$$

Данная начальная точка поиска

$$X_1 := (1 \quad 0)^T$$

Заданная точность поиска

$$\varepsilon := 0.001$$

Заданный начальный шаг λ

$$\lambda := 1$$

$$I := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Шаг1

$$q := hf(X_1)^{-1} \cdot gradf(X_1) \cdot \lambda$$

Проверка допустимости шага

$$f(X_1 - q) - f(X_1) = -4.753$$

Шаг подходит

$$X_2 := X_1 - q$$

$$X_2 = \begin{pmatrix} 0.839 \\ -1.086 \end{pmatrix} \qquad f(X_2) = -17.753$$

Проверка окончания

$$gradf(X_2) = \begin{pmatrix} 1.182 \\ 6.538 \times 10^{-12} \end{pmatrix}$$

Точность не достигнута

Шаг2

$$q := hf(X_2)^{-1} \cdot gradf(X_2) \cdot \lambda$$

Проверка допустимости шага

$$f(X_2 - q) - f(X_2) = -0.022$$

Шаг подходит

$$X_3 := X_2 - q$$

$$X_3 = \begin{pmatrix} 0.802 \\ -1.068 \end{pmatrix} \qquad f(X_3) = -17.775$$

Проверка окончания

$$gradf(X_3) = \begin{pmatrix} 0.053 \\ -2.298 \times 10^{-14} \end{pmatrix}$$

Точность не достигнута

Шаг3

$$q := hf(X_3)^{-1} \cdot gradf(X_3) \cdot \lambda$$

Проверка допустимости шага

$$f(X_3 - q) - f(X_3) = -4.825 \times 10^{-5}$$

Шаг подходит

$$X_4 := X_3 - q$$

$$X_4 = \begin{pmatrix} 0.8 \\ -1.067 \end{pmatrix} \qquad f(X_4) = -17.775$$

Проверка окончания

$$gradf(X_4) = \begin{pmatrix} 1.262 \times 10^{-4} \\ 3.244 \times 10^{-14} \end{pmatrix}$$

Точность достигнута

Ответ

$$xmin := X_4 = \begin{pmatrix} 0.8 \\ -1.067 \end{pmatrix}$$

$$fmin := f(xmin) = -17.775$$

Индивидуальные задания

№	f □x_1,x_2□	№	f □x_1,x_2□
1	2□ x_1^2 □ 0.2□ x_1 □ x_2 □ 4□ x_2^2 □3□ x_1 □ 6□ x_2 □ 2	2	3□ x_1^2 □ 0.7□ x_1 □ x_2 □5□ x_2^2 □ 6□ x_1 □ 4□ x_2 □ 2
3	3□ x_1^2 □ 0.2□ x_1 □ x_2 □ 4□ x_2^2 □3□ x_1 □ 6□ x_2 □ 2	4	4□ x_1^2 □ 0.7□ x_1 □ x_2 □5□ x_2^2 □ 6□ x_1 □ 4□ x_2 □ 2
5	3□ x_1^2 □ 0.2□ x_1 □ x_2 □5□ x_2^2 □3□ x_1 □ 6□ x_2 □ 2	6	3□ x_1^2 □ 0.7□ x_1 □ x_2 □5□ x_2^2 □8□ x_1 □ 4□ x_2 □ 2
7	3□ x_1^2 □ 0.2□ x_1 □ x_2 □5□ x_2^2 □3□ x_1 □ 2□ x_2 □ 2	8	3□ x_1^2 □1.7□ x_1 □ x_2 □5□ x_2^2 □ 6□ x_1 □ 4□ x_2 □ 2
9	3□ x_1^2 □ 0.2□ x_1 □ x_2 □5□ x_2^2 □3□ x_1 □ 2□ x_2 □ 2	10	3□ x_1^2 □1.7□ x_1 □ x_2 □ 4□ x_2^2 □ 6□ x_1 □ 4□ x_2 □ 2
11	3□ x_1^2 □ 0.2□ x_1 □ x_2 □5□ x_2^2 □3□ x_1 □ 2□ x_2 □ 2	12	3□x_1^2 □1.3□x_1 □x_2 □5□x_2^2 □6□ x_1 □4□x_2 □ 2
13	3□ x_1^2 □ 0.2□ x_1 □ x_2 □5□ x_2^2 □3□ x_1 □ 4□ x_2 □ 2	14	4□ x_1^2 □1.3□ x_1 □ x_2 □5□ x_2^2 □ 6□ x_1 □ 4□ x_2 □ 2

15	$3 \square x_1^2 \square 0.4 \square x_1 \square x_2 \square 5 \square x_2^2 \square 3 \square x_1 \square 4 \square x_2 \square 2$	16	$3 \square x_1^2 \square 1.2 \square x_1 \square x_2 \square 5 \square x_2^2 \square 6 \square x_1 \square 4 \square x_2 \square 2$
17	$3 \square x_1^2 \square 0.5 \square x_1 \square x_2 \square 5 \square x_2^2 \square 3 \square x_1 \square 4 \square x_2 \square 2$	18	$3 \square x_1^2 \square 1.3 \square x_1 \square x_2 \square 5 \square x_2^2 \square 3 \square x_1 \square 4 \square x_2 \square 2$
19	$3 \square x_1^2 \square 0.7 \square x_1 \square x_2 \square 5 \square x_2^2 \square 3 \square x_1 \square 4 \square x_2 \square 2$	20	$2 \square x_1^2 \square 1.3 \square x_1 \square x_2 \square 5 \square x_2^2 \square 6 \square x_1 \square 4 \square x_2 \square 2$
21	$2 \square x_1^2 \square 0.5 \square x_1 \square x_2 \square 5 \square x_2^2 \square 3 \square x_1 \square 4 \square x_2 \square 2$	22	$3 \square x_1^2 \square 1.3 \square x_1 \square x_2 \square 6 \square x_2^2 \square 6 \square x_1 \square 4 \square x_2 \square 2$

Контрольные вопросы

1. Что называется матрицей Гессе дифференцируемой функции двух переменных.
2. Записать разложение Тейлора для функции двух переменных до членов второго порядка малости.
3. Выписать формулу Ньютона для оптимизации второго порядка.
4. Как дробится шаг в формуле Ньютона, если он оказался слишком велик.

Лабораторная работа 5. Условный экстремум при ограничениях типа **равенств. Метод множителей Лагранжа**

Цель работы: Изучить метод Лагранжа решения экстремальных задач с ограничениями.

Задание. Найти точку абсолютного минимума в данной задаче с ограничениями.

Краткие теоретические положения

Задачей гладкой конечномерной оптимизации с ограничениями типа равенств и неравенств называется задача вида:

$f_0(x) \square \min;\ f_i(x) \square 0, i \square 1,\ldots, m';$ (P) $f_i(x) \square 0, i \square m' \square 1,\ldots, m;$

Теорема 1. (Метод Лагранжа). Пусть точка $x^* \square (x_1^* x_2^* \square\square x_n^*)^T$ является точкой локального минимума в задаче (1). Рассмотрим функцию Лагранжа:

$L(x,\square) \square \sum_{i \square 0}^{m} \square_i \square f_i(x)$. Найдется такой вектор $\square \square (\square_0,\ldots,\square_m)^T$, что в точке x^* выполняются условия:

1) Стационарности функции Лагранжа:

$$\frac{\partial L(x_*,\lambda)}{\partial x_j} = 0,\; j = 1,\dots,m. \qquad (2)$$

2) Дополняющей нежесткости:

$$\lambda_i \cdot f_i(x^*) = 0,\; i = 1,\dots,m'. \qquad (3)$$

3) Неотрицательности $\lambda_i \ge 0, i = 1,\dots,m'$. (4)

Условия (2)–(4) позволяют выделить конечное множество точек $A_1(x_1, y_1),\dots, A_k(x_k, y_k)$ подозрительных на экстремум, из которых, опираясь на соответствующие значения целевой функции $f(A_1(x_1, y_1)),\dots, f(A_k(x_k, y_k))$ выделить точку абсолютного минимума.

При этом используется теорема Вейерштрасса:

Теорема 2. Пусть непрерывная функция $f(x) = f(x_1,\dots, x_n)$ обладает свойством $f(x) \to \infty \; (\|x\| \to \infty)$, тогда она имеет абсолютный минимум на любом ограниченном замкнутом подмножестве в R^n.

Пример выполнения $f_0(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 \to \min$;

$2x_1 - x_2 + x_3 \le 5$; $x_1 + x_2 + x_3 = 3$; Решение.

1. Функция Лагранжа

$$L(x_1,x_2,x_3,\lambda_0,\lambda_1,\lambda_2) = \lambda_0(x_1^2 + x_2^2 + x_3^2) + \lambda_1(2x_1 - x_2 + x_3 - 5) + \lambda_2(x_1 + x_2 + x_3 - 3)$$

2. Условия стационарности:

$$\frac{\partial L(x,\lambda)}{\partial x_1} = 2\lambda_0 x_1 + 2\lambda_1 + \lambda_2 = 0;$$

$$\frac{\partial L(x,\lambda)}{\partial x_2} = 2\lambda_0 x_2 - \lambda_1 + \lambda_2 = 0;$$

$$\frac{\partial L(x,\lambda)}{\partial x_3} = 2\lambda_0 x_3 + \lambda_1 + \lambda_2 = 0;$$

3. Условия дополняющей нежесткости:

$$\lambda_1(2x_1 - x_2 + x_3 - 5) = 0;$$

4. Неотрицательносьти $\lambda_1 \ge 0$.
5. Связи:

$$x_1 + x_2 + x_3 = 3.$$

Рассмотрим сначала случай $\lambda_0 = 0$. Получим систему:

$$\begin{cases} \lambda_1 + 2\lambda_2 = 0; \\ \lambda_1 + \lambda_2 = 0; \quad \lambda_1 + \lambda_2 = 0; \end{cases}$$

Из этой системы определяем, что $\lambda_0 = \lambda_1 = \lambda_2 = 0$. Но это невозможно, так как вектор множителей Лагранжа должен быть отличен от 0. Поэтому $\lambda_0 \neq 0$, можно взять $\lambda_0 = 1/2$. Получаем систему:

$$\begin{cases} x_1 + \lambda_1 + 2\lambda_2 = 0; \\ x_2 + \lambda_1 + \lambda_2 = 0; \\ x_3 + \lambda_1 + \lambda_2 = 0; \end{cases} \quad (5)$$

Рассмотрим 2 подслучая: А) $\lambda_1 \neq 0$.

Тогда из условия 2 и 4 получаем $2x_1 + x_2 + x_3 - 5 = 0$; $x_1 + x_2 + x_3 = 3$;

Используя (5) данную систему можно записать в виде:

$$\begin{cases} -2(\lambda_1 + 2\lambda_2) - (\lambda_1 + \lambda_2) - (\lambda_1 + \lambda_2) - 5 = 0; \\ -(\lambda_1 + 2\lambda_2) - (\lambda_1 + \lambda_2) - (\lambda_1 + \lambda_2) = 3; \end{cases}$$

Эту систему можно записать в виде:

$$\begin{cases} -6\lambda_1 - 2\lambda_2 = 5; \\ -2\lambda_1 - 3\lambda_2 = 3. \end{cases}$$

Решая эту систему получаем $\lambda_1 = -\dfrac{9}{14}, \lambda_2 = -\dfrac{4}{7}$. При этом не выполняется условие неотрицательности 3. Следовательно, данный случай невозможен.

Подслучай В) $\lambda_1 = 0$.

Получаем: $x_1 + \lambda_2 = 0$;

$x_2 + \lambda_2 = 0$; $x_3 + \lambda_2 = 0$;

Следовательно, $x_1 = x_2 = x_3 = -\lambda_2$. Подставляя данное соотношение в 4, получаем $x_1 = 1$, $x_2 = 1$, $x_3 = 1$, $f(1,1,1) = 3$.

Используя теорему Вейершрасса будем иметь $x^* = (1,1,1) \in abs\min(P)$.

Индивидуальные задания

1. $x_1x_2x_3 \to \mathrm{extr}; \quad x_1^2 + x_2^2 + x_3^2 \leqslant 1.$

2. $\max x_j^2 \to \mathrm{extr}; \quad \sum\limits_{j=1}^{n} x_j^4 \leqslant 1.$

3. $\sum\limits_{j=1}^{n} x_j^4 \to \mathrm{extr}; \quad \sum\limits_{j=1}^{n} x_j^2 \leqslant 1.$

4. $e^{x_1 - x_2} - x_1 - x_2 \to \mathrm{extr}; \quad x_1 + x_2 \leqslant 1, \quad x_1 \geqslant 0, \quad x_2 \geqslant 0.$

5. $x_1^2 + x_2 \to \mathrm{extr}; \quad x_1^2 + x_2^2 \leqslant 1, \quad x_1 \geqslant 0, \quad x_2 \geqslant 0.$

6. $x_1^2 + x_2^2 + x_3^2 \to \mathrm{extr}; \quad x_1 + x_2 + x_3 \leqslant 12, \quad x_1 \geqslant 0, \quad x_2 \geqslant 0, \quad x_3 \geqslant 0.$

7. $x_1^2 + 4x_2^2 + x_3^2 \to \mathrm{extr}; \quad x_1 + x_2 + x_3 \leqslant 12, \quad x_1 \geqslant 0, \quad x_2 \geqslant 0, \quad x_3 \geqslant 0.$

8. $x_1^2 + x_2^2 + x_3^2 \to \mathrm{extr}; \quad x_1 + x_2 + x_3 \leqslant 1, \quad x_1 \geqslant 0, \quad x_1 + x_2 - x_3 = \dfrac{1}{2}.$

9. $2x_1^2 + 2x_1 + 4x_2 - 3x_3 \to \mathrm{extr}; \quad 8x_1 - 3x_2 + 3x_3 \leqslant 40,$
$-2x_1 + x_2 - x_3 = -3, \quad x_2 \geqslant 0.$

Контрольные вопросы

1. Сформулировать теорему Лагранжа.
2. Сформулировать теорему Вейерштрасса.
3. Привести определение конечномерной гладкой задачи с ограничениями.
4. Дать определение локальной точки экстремума.
5. Дать определение глобальной точки экстремума.

Лабораторная работа 6. Симплекс-метод решения задачи линейного программирования

Цель работы: получение навыков применения симплекс-метод для решения задач линейного программирования.

Задание. Решить данную стандартную задачу линейного программирования симплекс-методом.

Краткие теоретические сведения

Канонической формой задачи линейного программирования называется следующая задача оптимизации:

$$z(x) = c^T \cdot x \to \max; \qquad (1)$$

$$A \cdot x = b; \qquad (2)$$

$$x \geq 0; \qquad (3)$$

где

$x = (x_1\ x_2\ \ldots\ x_n)^T;$

$c = (c_1\ c_2\ \ldots\ c_n)^T;$

$b^T = (b_1\, b_2 \ldots b_m) \geq 0;$

$A = A[m,n], rg(A) = m; m < n.$

Разобьем матрицу A на две подматрицы B и N, т.е. запишем $A = [B\,|\,N]$, где B – невырожденная подматрица $m \times m$. Тогда система (2) приобретет вид

$$B \cdot x_b + N \cdot x_N = b\,, \qquad (4)$$

где вектор x разбит на базисную и небазисную часть следующим образом:

$$x = \begin{pmatrix} x_b \\ x_N \end{pmatrix},$$

Из (4) следует, что

$$B \cdot x_b = -N \cdot x_N + b \qquad (5)$$

Так как по предположению B - невырожденная матрица, то

$$x_b = -B^{-1}N \cdot x_N + B^{-1}b \qquad (6).$$

Формула (6) означает, что m базисных переменных x_b выражены через $n - m$ небазисных переменных x_N. Вектор x называется базисным решением системы (2), если в нем все небазисные компоненты равны 0, т.е. если $x_N = 0$ и. следовательно, $x_b = B^{-1}b$.

Допустимое базисное решение- это такое базисное решение, в котором все элементы базисного вектора неотрицательны, т.е. выполняется условие $x_b = B^{-1}b \geq 0$. Такое преобразование, выражение базисных неизвестных через небазисные выполняется также и в целевой функции. Разобьем вектор цен на базисную и небазисную части:

$c = \begin{pmatrix} c_b \\ c_N \end{pmatrix}$. Тогда целевая функция задачи может быть преобразована следующим образом:

$$z(x) = c^T x = \begin{pmatrix} c_b \\ c_N \end{pmatrix}^T \begin{pmatrix} x_b \\ x_N \end{pmatrix} = \begin{pmatrix} c_b^T & c_N^T \end{pmatrix} \begin{pmatrix} x_b \\ x_N \end{pmatrix} = c_b^T \cdot x_b + c_N^T \cdot x_N \qquad (7)$$

В данном выражении выразим базисные переменные через небазисные по формуле (6), получим $z(x) = c_b^T \cdot \left(-B^{-1}N \cdot x_N + B^{-1}b\right) + c_N^T \cdot x_N = c_b^T \cdot B^{-1}b - \left(c_b^T \cdot B^{-1}N - c_N^T\right) \cdot x_N$. (8)

Таким образом, исходная задача линейного программирования может быть представлена в виде:

$$z(x) = c_b^T B^{-1} b - \left(c_b^T B^{-1} N - c_N^T\right) x_N \to \max;$$
$$x_b + B^{-1} N x_N = B^{-1} b \ge 0; \quad (9) \quad x_N \ge 0$$

Обозначим $\beta = B^{-1}b$, $\pi = \left(c_b^T B^{-1}\right)^T$ задачу (9) можно записать в виде:

$$z(x) = c_b^T \beta - \left(\pi^T N - c_N^T\right) x_N \to \max;$$
$$x_b + B^{-1} N x_N = \beta \ge 0; \quad (10) \quad x_N \ge 0$$

Вектор-строка $d = \pi^T N - c_N^T$ называется строкой относительных оценок небазисных переменных. Необходимым и достаточным условием оптимальности текущего базисного решения $x = \begin{pmatrix} x^b = \beta \\ 0 \end{pmatrix}$ является условие $d = \pi^T N - c_N^T \ge 0$, так как если относительная оценка d_j некоторой небазисной переменной x_j отрицательна, то эту небазисную переменную можно увеличить от текущего нулевого значения и в соответствии с уравнением $z(x) = c_b^T \beta - \left(\pi^T N - c_N^T\right) x_N$ получить приращение целевой функции.

На основе анализа относительных оценок небазисных переменных и строится алгоритм симплекс-метода. Предположим, у нас имеется текущее допустимое базисное решение с базисной подматрицей B и неотрицательной правой частью $\beta = B^{-1} b$.

Шаг 1. Определение переменной вводимой в базис. Если все относительные оценки неотрицательны, задача решена, текущее базисное решение является оптимальным. Иначе находим наименьшую отрицательную оценку $d_q < 0$, небазисная переменная $x_q = 0$ будет увеличиваться и становиться положительной, т.е. будет вводиться в базис.

Шаг 2. Определение переменной выводимой из базиса. При увеличении x_q какие то базисные переменные н7ачнут уменьшаться, но они не могут быть отрицательными, первая базисная переменная, которая обратиться в 0 или любая из них, если таких переменных будет несколько удаляется из базиса.

Т.е. переменную x_p, выводимую из базиса, мы определяем из решения следующей задачи:

$$\frac{\beta_p}{\alpha_{pq}} = \min_{\alpha_{iq} > 0} \frac{\beta_i}{\alpha_{iq}} \quad (11).$$

Если такая переменная x_p не существует, т.е. не определяется из уравнения (9), то целевая функция не ограничена и может быть увеличена до ∞, работа алгоритма закончена.

Шаг 3. Преобразование симплекс таблицы методом Гаусса с заданным ведущим элементом. Столбец p, равный удаляется из базисной подматрицы B, а на его место ставится столбец $□_q$, что делается преобразованием Гаусса с ведущим элементом $□_{p,q}$, при это столбец $□_q$ должен в позиции $□_{p,q}$ получить значение, равное 1, а в остальных позициях получить значение равное 0.

Преобразование Гаусса распространяется и на столбец □ правых частей и на строку d относительных оценок небазисных переменных, в результате они приобретают обновленные значения, соответствующие новому базису. После этого возвращаемся к шагу 1.

Все эти преобразования выполняются внутри специальной таблицы, называемой таблицей симплекс-метода.

Таблица симплекс метода имеет следующую структуру:

Базис	z	x_1	□	x_n	Решение	$b_i/A_{i,q}$
z	1	d_1	□	d_n	$z□x□$	-
x_b	0	$B^{□1} □ A□ □$			$□□ B^{□1} □b$	$\frac{□_i}{□_{i,q}}$

Пример выполнения.

Решить задачу линейного программирования:

$z = x_1 + 2 x_2 \to \max;$

$2 x_1 + x_2 \le 10;$

$x_1 + 2 x_2 \le 10;\ x_{1,2} \ge 0.$

1. Приводим задачу к каноническому виду (1)–(3): $z \to \max;\ z - x_1 - 2 x_2 - 0 s_1 - 0 s_2 = 0;\ 2 x_1 + x_2 + 1 s_1 + 0 s_2 = 10;\ x_1 + 2 x_2 + 0 s_1 + 1 s_2 = 10;\ x_{1,2}, s_1, s_2 \ge 0.$

При этом мы ввели в рассмотрение так называемые остаточные переменные s_1, s_2 .

2. Заполняем начальную симплекс-таблицу и производим ее анализ в соответствии с алгоритмом симплекс-метода.

Базис	z	x_1	x_2	$s_1 □ x_3$	$s_2 □ x_4$	Реш.	Bi/Aiq
z	1	-1	-2	0	0	0	-
s_1	0	2	1	1	0	10	
s_2	0	1	2	0	1	10	

В начальной таблице базисными переменными являются остаточные переменные s_1, s_2 , которые являются 3 и 4 переменной задачи, таким образом, имеем $□ □ □ s_1, s_2 □$. Базисным переменным соответствует единичная подматрица

□1 0□

0 1 матрицы симплекс-таблицы. Переменные x_1, x_2 являются небазисными, их

относительные оценки равны соответственно $d_1$1,$d_2$2.

Наименьшей отрицательной относительной оценкой обладает переменная x_2, она направляется в базис, что можно записать в виде: x_2. Это означает, что столбец

x_2
–2
1
2

Симплекс таблицы является ведущим и должен быть сделан единичным методом Гаусса. Мы должны определить переменную выводимую из базиса. Для этого рассчитаем отношения правых частей к положительным элементам ведущего столбца и выберем наименьшее:

Bi/Aiq
-
10/1=10
10/2=5

Рассчитанном столбце наименьшее найденное отношение, равное 5, располагается во второй строке симплекс таблицы (отметим, что строки симплекс таблицы нумеруются с 0, причем нулевая строка- это z-строка). Поэтому вторая строка симплекс таблицы будет ведущей, ей соответствует переменная s_2, которая выводится из базиса, элемент $_{2,2}$ 2на пересечении ведущей строки и ведущего столбца будет ведущим и он делается равным 1 по методу Гаусса:

А) Разделим ведущую строчку на ее ведущий элемент, получим переход от строки

s_2	0	1	2	0	1	10	5

К строке

s_2	0	1/2	1	0	1/2	5	

В) Вычтем из z-строки ведущую строку, умноженную на (–2) и из первой строки ведущую строку, умноженную на 1, т.е. сделаем следующие преобразования симплекс-таблицы:

Базис	z	x_1	x_2	s_1 x_3	$s_2$$x_4$	Реш.	Bi/Aiq
z	1-0·(-2)	-1-1/2·(-2)	-2-1·(-2)	0-0·(-2)	0-1/2·(-2)	0-5·(-2)	-
s_1	0-0·(1)	2-1/2·(1)	1-1·(1)	1-0·(1)	0-1/2·(1)	10-5·(1)	
s_2	0	1/2	1	0	1/2	5	

В результате получим следующую симплекс-таблицу:

Базис	z	x_1	x_2	$s_1 \square x_3$	$s_2 \square x_4$	Реш.	Bi/Aiq
z	1	0	0	0	1	10	-
s_1	0	3/2	0	1	-1/2	5	
x_2	0	1/2	1	0	1/2	5	

Обрабатываем вторую симплекс-таблицу. В ней относительные оценки всех переменных, т.е. элементы z-строки неотрицательны, получена оптимальная симплекс-таблица. Выписываем ответ задачи:

z=10 – максимальная прибыль;

$x_1=0$ (небазисная переменная)- производство первого продукта;

$x_2=5$ – производство второго продукта; $s_1=5$ – остаток первого ресурса; $s_2=0$(небазисная переменная) – остаток второго ресурса. **Индивидуальные задания**

№	Задача линейного программирования	№	Задача линейного программирования
1	$z \square 3\square x_1 \square 2\square x_2 \square \max;$ $2\square x_1 \square x_2 \square 10;$ $x_1 \square 2\square x_2 \square 10;$ $x_{1,2} \square 0.$	2	$z \square x_1 \square 2\square x_2 \square \max;$ $2\square x_1 \square x_2 \square 15;$ $x_1 \square 2\square x_2 \square 10;$ $x_{1,2} \square 0.$
3	$z \square 3\square x_1 \square 2\square x_2 \square \max;$ $2\square x_1 \square x_2 \square 10;$ $x_1 \square 2\square x_2 \square 20;$ $x_{1,2} \square 0.$	4	$z \square x_1 \square x_2 \square \max;$ $2\square x_1 \square x_2 \square 10;$ $x_1 \square 4\square x_2 \square 10;$ $x_{1,2} \square 0.$
5	$z \square x_1 \square x_2 \square \max;$ $2\square x_1 \square x_2 \square 12;$ $x_1 \square 3\square x_2 \square 10;$ $x_{1,2} \square 0.$	6	$z \square x_1 \square 2\square x_2 \square \max;$ $5\square x_1 \square x_2 \square 10;$ $x_1 \square 4\square x_2 \square 10;$ $x_{1,2} \square 0.$
7	$z \square 3\square x_1 \square 2\square x_2 \square \max;$ $2\square x_1 \square x_2 \square 12;$ $x_1 \square 2\square x_2 \square 10;$ $x_{1,2} \square 0.$	8	$z \square x_1 \square 3\square x_2 \square \max;$ $2\square x_1 \square x_2 \square 10;$ $x_1 \square 2\square x_2 \square 10;$ $x_{1,2} \square 0.$
9	$z \square x_1 \square 4\square x_2 \square \max;$ $2\square x_1 \square x_2 \square 10;$ $x_1 \square 2\square x_2 \square 10;$ $x_{1,2} \square 0.$	10	$z \square x_1 \square 2\square x_2 \square \max;$ $2\square x_1 \square x_2 \square 10;$ $3\square x_1 \square 2\square x_2 \square 20;$ $x_{1,2} \square 0.$
11	$z \square 4\square x_1 \square 2\square x_2 \square \max;$ $2\square x_1 \square x_2 \square 10;$	12	$z \square x_1 \square 2\square x_2 \square \max;$ $3\square x_1 \square x_2 \square 12;$

	$x_1 \square 2\square x_2 \square 15;$ $x_{1,2} \square 0.$		$x_1 \square 2\square x_2 \square 10;$ $x_{1,2} \square 0.$
№	Задача линейного программирования	№	Задача линейного программирования
13	$z \square x_1 \square 2\square x_2 \square \max;$ $2\square x_1 \square x_2 \square 10;$ $x_1 \square 3\square x_2 \square 12;\ x_{1,2} \square 0.$	14	$z \square x_1 \square 2\square x_2 \square \max;$ $2\square x_1 \square x_2 \square 10;$ $x_1 \square 2\square x_2 \square 8;$ $x_{1,2} \square 0.$
15	$z \square 5\square x_1 \square 2\square x_2 \square \max;$ $2\square x_1 \square x_2 \square 10;$ $x_1 \square 2\square x_2 \square 10;$ $x_{1,2} \square 0.$	16	$z \square 3\square x_1 \square 2\square x_2 \square \max;$ $2\square x_1 \square x_2 \square 12;$ $x_1 \square 2\square x_2 \square 10;$ $x_{1,2} \square 0.$
17	$z \square x_1 \square 4\square x_2 \square \max;$ $2\square x_1 \square x_2 \square 10;$ $x_1 \square 2\square x_2 \square 10;$ $x_{1,2} \square 0.$	18	$z \square x_1 \square 5\square x_2 \square \max;$ $2\square x_1 \square x_2 \square 10;$ $x_1 \square 2\square x_2 \square 10;$ $x_{1,2} \square 0.$
19	$z \square x_1 \square 2\square x_2 \square \max;$ $4\square x_1 \square x_2 \square 20;$ $x_1 \square 2\square x_2 \square 10;$ $x_{1,2} \square 0.$	20	$z \square x_1 \square 3\square x_2 \square \max;$ $3\square x_1 \square x_2 \square 10;$ $x_1 \square 2\square x_2 \square 10;$ $x_{1,2} \square 0.$
21	$z \square x_1 \square 4\square x_2 \square \max;$ $2\square x_1 \square x_2 \square 10;$ $x_1 \square 2\square x_2 \square 30;$ $x_{1,2} \square 0.$	22	$z \square x_1 \square 2\square x_2 \square \max;$ $2\square x_1 \square 4\square x_2 \square 20;$ $x_1 \square 2\square x_2 \square 10;$ $x_{1,2} \square 0.$
23	$z \square 4\square x_1 \square 2\square x_2 \square \max;$ $2\square x_1 \square x_2 \square 20;$ $x_1 \square 2\square x_2 \square 10;$ $x_{1,2} \square 0.$	24	$z \square x_1 \square 5\square x_2 \square \max;$ $2\square x_1 \square x_2 \square 10;$ $x_1 \square 4\square x_2 \square 10;$ $x_{1,2} \square 0.$
25	$z \square x_1 \square 2\square x_2 \square \max;$ $2\square x_1 \square x_2 \square 20;$ $x_1 \square 2\square x_2 \square 15;$ $x_{1,2} \square 0.$	26	$z \square x_1 \square 2\square x_2 \square \max;$ $2\square x_1 \square x_2 \square 10;$ $x_1 \square 3\square x_2 \square 18;\ x_{1,2} \square 0.$

Контрольные вопросы

1. Какой вид имеет каноническая задача линейного программирования.
2. Что такое базисное решение, допустимое базисное решение.
3. Какой вид имеет преобразованная задача линейного программирования после исключения базисных переменных.
4. Какой критерий оптимальности в преобразованной задаче линейного программирования.
5. Назвать шаги алгоритма симплекс-метода.

Лабораторная работа 7. Методы решения транспортной задачи

Цель работы: Изучение метода потенциалов решения транспортной задачи.

Задание

Решить индивидуальное задание по образцу приведенного решения.

Краткие теоретические положения

Транспортная задача – это задача о поиске оптимального распределения поставок однородного товара от поставщиков к потребителям при известных затратах на перевозку (тарифах) между пунктами отправления и назначения

Транспортная задача может быть записана в виде прямоугольной таблицы Пример подобной таблицы приведен ниже

	Потребитель B_1, потребность 20 кг	**Потребитель B_2, потребность 30 кг**	**Потребитель B_3, потребность 30 кг**	**Потребитель B_4, потребность 10 кг**
Поставщик A_1, запас 30 кг	C_{11}=2 руб./кг	C_{12}=3 руб./кг	C_{13}=2 руб./кг	C_{14}=4 руб./кг
Поставщик A_2, запас 40 кг	C_{21}=3 руб./кг	C_{22}=2 руб./кг	C_{23}=5 руб./кг	C_{24}=1 руб./кг
Поставщик A_3, запас 20 кг	C_{31}=4 руб./кг	C_{32}=3 руб./кг	C_{33}=2 руб./кг	C_{34}=6 руб./кг

Цена перевозки (например, в рублях за 1 килограмм груза) C_{ij} записывается в ячейки таблицы на пересечении соответствующего потребителя и поставщика (цена может быть и отрицательной – в этом случае она представляет собой прибыль) Неизвестной (искомой) величиной в задаче являются такие объемы перевозки x_{ij} от поставщиков к потребителям, чтобы минимизировать общие затраты на транспортировку

В табличной записи цены отделяют от объемов перевозки косой чертой или квадратным уголком, в этой статье из соображений лучшей доходчивости они подписаны. При решении транспортной задачи единственными необходимыми арифметическими действиями являются сложение и вычитание.

Для поиска начального решения применяют метод северо-западного угла, метод минимальных тарифов или метод Фогеля, а для окончательной оптимизации – метод потенциалов. В то же время, транспортная задача является подмножеством задач линейного программирования и может решаться симплексметодом.

Если сумма запасов равна сумме потребностей, то транспортная задача называется *закрытой*. Если равенство не соблюдается, то задача называется *открытой*. Для решения транспортной задачи необходимо, чтобы она была приведена к закрытому виду.

В показанном выше примере, сумма запасов = 30 + 40 + 20 = 90 кг, а сумма потребностей = 20 + 30 + 30 + 10 кг = 90 кг (запасы и потребности равны между собой, задача закрытая).

Если это равенство не соблюдено, необходимо ввести фиктивного поставщика или фиктивного потребителя на недостающий или избыточный объем товара, которому нужно приписать нулевую цену доставки. Этот объем будет соответствовать недопоставке или, напротив, избытку товара на складе.

Решение транспортной задачи начинается с поиска допустимого начального решения (плана перевозок), чтобы все запасы поставщиков были распределены по потребителям. Допустимое начальное решение не обязательно оказывается оптимальным, а метод его нахождения может быть как простейшим (метод северо-западного угла или аналоги) или более сложным и приближенным к оптимальному решению (метод минимальных тарифов), или же вообще произвольным.

Допустимое (но не всегда оптимальное с точки зрения стоимости доставки) начальное решение транспортной задачи можно построить, последовательно перебирая строки таблицы (то есть поставщиков) сверху вниз.

В пределах каждой строки нужно перебрать слева направо не охваченных или не полностью охваченных поставками потребителей, записывая в соответствующие ячейки объем поставляемого груза от поставщика в данной строке, и так до исчерпания возможностей поставщика.

Таким образом, весь груз от поставщиков будет распределен по потребителям. Этот метод был предложен Данцигом в 1951 г. и назван Чарнесом и Купером «правилом северо-западного угла».

	B_1, 20 кг	**B_2, 30 кг**	**B_3, 30 кг**	**B_4, 10 кг**

A_1, 30 кг	X_{11}=20 кг	X_{12}=10 кг		
A_2, 40 кг		X_{22}=20 кг	X_{23}=20 кг	
A_3, 20 кг			X_{33}=10 кг	X_{34}=10 кг

В таблице здесь и далее зеленым цветом отмечены ячейки с ненулевыми объемами перевозки груза (базисные ячейки).

Другой метод получения начального решения — записывать отгрузки в первую очередь в те ячейки, где тариф минимален. Этот метод позволяет получить более приближенное к оптимальному решению, которое, однако, может потребовать дальнейшей оптимизации. Метод минимальных тарифов с его модификациями (минимальный тариф по строке или минимальный тариф по столбцу) был описан Данцигом в работе 1951 г.

Для поиска начального решения транспортной задачи можно применить также метод Фогеля, который обычно дает еще более приближенное к оптимальному решение.

Метод потенциалов позволяет за несколько шагов (итераций) найти полностью оптимальное решение транспортной задачи. Перед решением задачи этим методом нужно найти допустимое начальное решение одним из методов, описанных в разделе выше. Поскольку у нас нет ограничений на черно-белую полиграфию, для большей ясности ячейки транспортной таблицы в этой статье отмечены разными цветами.

Эта проверка не входит в алгоритм метода потенциалов, но может потребоваться для исключения арифметических ошибок (при ручном расчете на бумаге) или самопроверки алгоритма при компьютерных вычислениях.

Особенностью распределения груза по транспортной таблице является совпадение суммы объемов по строкам с запасами соответствующего поставщика, а суммы объемов по столбцам – с потребностями соответствующих потребителей. В показанном выше примере,

- Для 1-й строки: 30 кг = 20 + 10 кг
- Для 2-й строки: 40 кг = 20 + 20 кг
- Для 3-й строки: 20 кг = 10 + 10 кг
- Для 1-го столбца: 20 кг = 20 кг
- Для 2-го столбца: 30 кг = 10 + 20 кг
- Для 3-го столбца: 30 кг = 20 + 10 кг
- Для 4-го столбца: 10 кг = 10 кг

Этот шаг также не входит в сам алгоритм метода потенциалов, но он полезен для распечатки результатов и показа, что алгоритм движется в правильном направлении, уменьшая на каждом (или не на каждом) шаге общую себестоимость перевозки. Для всех ячеек цена умножается на объем перевозки и полученный результат суммируется.

	B_1, 20 кг	**B_2, 30 кг**	**B_3, 30 кг**	**B_4, 10 кг**
A_1, 30 кг	C_{11}=2 руб./кг, X_{11}=20 кг	C_{12}=3 руб./кг, X_{12}=10 кг	C_{13}=2 руб./кг	C_{14}=4 руб./кг
A_2, 40 кг	C_{21}=3 руб./кг	C_{22}=2 руб./кг, X_{22}=20 кг	C_{23}=5 руб./кг, X_{23}=20 кг	C_{24}=1 руб./кг
A_3, 20 кг	C_{31}=4 руб./кг	C_{32}=3 руб./кг	C_{33}=2 руб./кг, X_{33}=10 кг	C_{34}=6 руб./кг, X_{34}=10 кг

В нашем примере, сумма затрат на перевозку груза составляет

$2×20 + 3×10 + 2×20 + 5×20 + 2×10 + 6×10 = 290$ руб.

Ячейки (клетки) транспортной таблицы с ненулевыми перевозками называются *базисными*, а клетки с нулевыми объемами перевозки – *свободными*

Базисных (см. выше) ячеек таблицы должно быть равно m+n-1, где *m* и *n* – соответственно, число поставщиков и потребителей, иначе решение считается *вырожденным* и требует введения в базис одной из ячеек с нулевой перевозкой (чтобы алгоритм не впал в бесконечный цикл, эта ячейка должна быть случайной).

Для исключения ситуаций с вырожденностью к объемам потребления добавляют небольшие *возмущения* – числа, заведомо ничтожные при перевозках (такие как 0.00001), при этом к объему поставки одного из поставщиков добавляют их сумму (или наоборот)

Каждому поставщику A_i соответствует потенциал U_i, а каждому потребителю B_j соответствует потенциал V_j. Данциг называет потенциалы U_i и V_j симплекс-множителями или неявными ценами. Чтобы определить эти потенциалы, полагают, что U_1=0, а остальные потенциалы вычисляют из соотношения $U_i + V_j = C_{ij}$ для всех занятых (базисных) ячеек таблицы (отмечены зеленым).

	V_1	**V_2**	**V_3**	**V_4**

U_1=0	C_{11}=2 руб./кг	C_{12}=3 руб./кг		
U_2		C_{22}=2 руб./кг	C_{23}=5 руб./кг	
U_3			C_{33}=2 руб./кг	C_{34}=6 руб./кг

$U_1+V_1=2$. Поскольку $U_1=0$, $0+V_1=2$, следовательно, $V_1=2$ руб./кг

$U_1+V_2=3$. Поскольку $U_1=0$, $0+V_2=3$, следовательно, $V_2=3$ руб./кг

$U_2+V_2=2$. Поскольку $V_2=3$, $U_2+3=2$, следовательно, $U_2=–1$ руб./кг

$U_2+V_3=5$. Поскольку $U_2=–1$, $–1+V_3=5$, следовательно, $V_3=6$ руб./кг

$U_3+V_3=2$. Поскольку $V_3=6$, $U_3+6=2$, следовательно, $U_3=–4$ руб./кг

$U_3+V_4=6$. Поскольку $U_3=–4$, $–4+V_4=6$, следовательно, $V_4=10$ руб./кг

При компьютерной реализации удобно использовать рекурсию: взаимный вызов двух функций, которые отрабатывают алгоритм, соответственно, по строкам и по столбцам..

Для всех незанятых ячеек (с нулевым объемом перевозки) вычисляют оценки клеток распределительной таблицы *Δij* по формуле $\Delta_{ij} = C_{ij} – U_i – V_j$, где U_i и V_j берутся из вычислений, выполненных в разделе выше (здесь они вписаны в заголовки таблицы).

Для всех занятых ячеек (с ненулевыми объемами перевозки, отмечены зеленым цветом) полагают $\Delta_{ij}=0$, поскольку на следующем шаге нам потребуется значение с минимальной оценкой в незанятых ячейках.

	V_1=2	**V_2=3**	**V_3=6**	**V_4=10**
U_1=0	Δ_{11}=0	Δ_{12}=0	Δ_{13} = 2–0–6 = –4	Δ_{14}=4–0–10 = –6
U_2=–1	Δ_{21}=3–(–1)–2 = 2	Δ_{22}=0	Δ_{23}=0	Δ_{24}=1–(–1)–10 = –8
U_3=–4	Δ_{31}=4–(–4)–2 = 6	Δ_{32}=3–(–4)–3 = 4	Δ_{33}=0	Δ_{34}=0

Если в получившейся таблице нет отрицательных значений Δ_{ij}, то план перевозок оптимален и задача решена (переход к шагу 10).

В нашем примере есть отрицательные значения. Наличие отрицательных значений Δ_{ij} означает, что решение не оптимально.

	B₁	**B₂**	**B₃**	**B₄**
A₁	$\Delta_{11}=0$	$\Delta_{12}=0$	$\Delta_{13}=–4$	$\Delta_{14}=–6$
A₂	$\Delta_{21}=2$	$\Delta_{22}=0$	$\Delta_{23}=0$	$\Delta_{24}=–8$
A₃	$\Delta_{31}=6$	$\Delta_{32}=4$	$\Delta_{33}=0$	$\Delta_{34}=0$

Наименьшее отрицательное значение Δ_{24} = –8 (начальная вершина для цикла перераспределения поставок, о котором (см. ниже) отмечено красным цветом. Если одинаковых отрицательных значений несколько, то берется любое.

Цикл перераспределения поставок представляет собой замкнутую ломаную линию, которая соединяет начальную вершину (отмечена красным цветом) и занятые (отмеченные в нашем примере зеленым цветом) ячейки транспортной таблицы по определенным правилам.

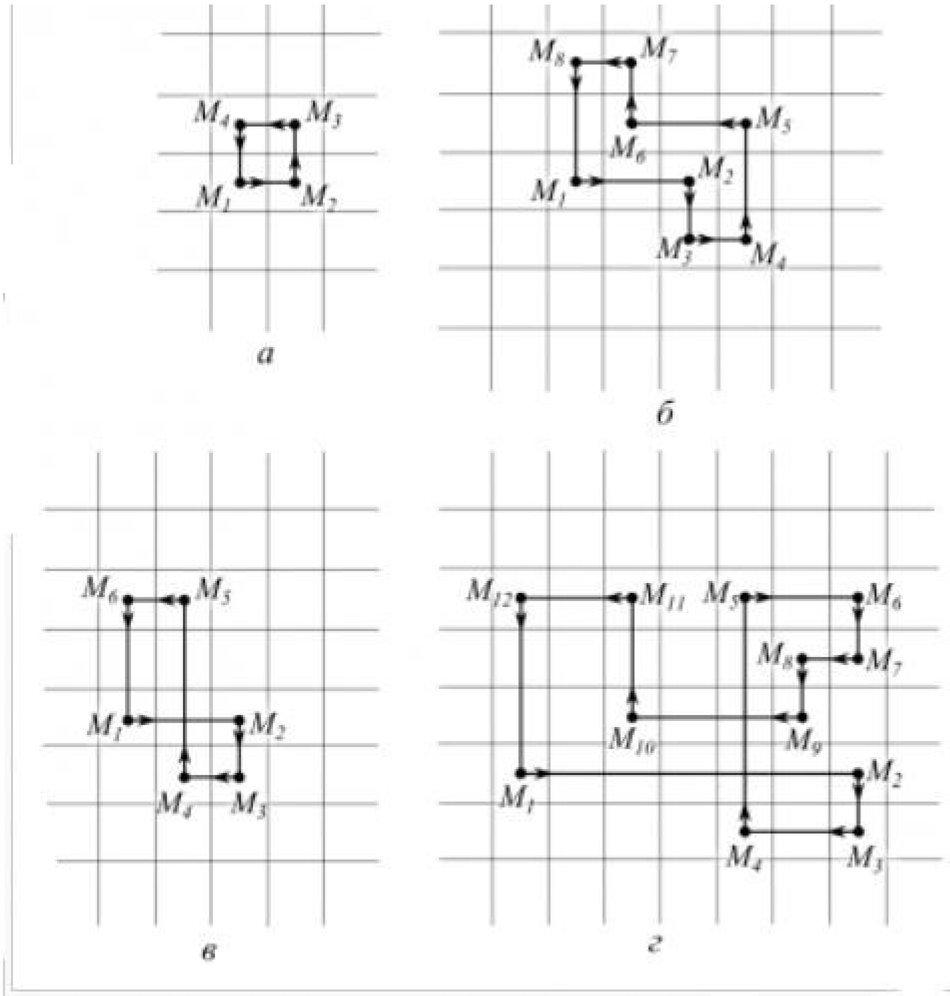
1. Все вершины, кроме начальной, находятся в занятых ячейках таблицы (ячейки с ненулевыми перевозками или «введенные в базис» на шаге 4 (в разделе «Проверка плана на вырожденность» ячейки с нулевой перевозкой – здесь они отмечены в примерах зеленым цветом), при этом охвачены циклом могут быть не все, а лишь некоторые занятые ячейки.

2. В каждой вершине цикла встречаются ровно два звена ломаной линии, причем одна из них находится по строке, а другая – по столбцу. Иначе говоря, они пересекаются под прямым углом.

3. Линия может пересекать занятые ячейки, не включая их в цикл (включение их в цикл не допускается). Другими словами, никакие три последовательные вершины не могут находиться в одной и той же строке или одном и том же столбце

4. Линия может пересекать саму себя, при этом точка пересечения не включается в цикл (исходя из п.2).

5.



Допустимые циклы для транспортной задачи.

В данном примере имеем:

	B_1, 20 кг	B_2, 30 кг	B_3, 30 кг	B_4, 10 кг
A_1, 30 кг	X_{11}=20 кг	X_{12}=10 кг		
A_2, 40 кг		X_{22}=20 кг	(*) X_{23}=20 кг	(*)
A_3, 20 кг			(*) X_{33}=10 кг	(*) X_{34}=10 кг

Вершины цикла в этом примере помечены звездочкой (*). По вершинам цикла нужно перераспределить объемы, чтобы получить следующее приближение к оптимальному решению задачи, как это показано далее.

«Красной» ячейке цикла присваиваем знак (+), следующей по циклу (начать двигаться можно в любом направлении) – знак (–), следующей ячейке цикла – опять (+) и так далее. Находим минимальную поставку по отмеченным знаком (–) вершинам цикла и обозначаем ее θ. Эта вершина цикла X_{34}=10 кг помечена желтым цветом. Значение θ вычитаем из вершин цикла, которые помечены знаком (–) и прибавляем его к вершинам цикла, которые помечены знаком (+).

	B_1, 20 кг	**B_2, 30 кг**	**B_3, 30 кг**	**B_4, 10 кг**
A_1, 30 кг	X_{11}=20 кг	X_{12}=10 кг		
A_2, 40 кг		X_{22}=20 кг	(–) X_{23}=20 кг	(+)
A_3, 20 кг			(+) X_{33}=10 кг	(–) X_{34}=10 кг

Получаем новое решение, которое чуть-чуть оптимальнее

	B_1, 20 кг	**B_2, 30 кг**	**B_3, 30 кг**	**B_4, 10 кг**
A_1, 30 кг	X_{11}=20 кг	X_{12}=10 кг		
A_2, 40 кг		X_{22}=20 кг	X_{23}=10 кг	X_{24}=10 кг
A_3, 20 кг			X_{33}=20 кг	

Поскольку алгоритм является циклическим (итерационным), переходим к пункту 1.

Примечание: есть опасность, что алгоритм впадет в бесконечный цикл из-за вырожденности. Впрочем, по мнению Данцига, те меры, которые можно предпринять для исключения вырожденности (см. Вырожденность в транспортной задаче) приводят к успеху в 100 % случаев. Для подстраховки можно применить метод Фогеля, который не склонен «впадать» в бесконечные циклы, и выдает более или менее приближенное к оптимальному решение за ограниченное число шагов.

Пример решения типового варианта Задание.

Решить транспортную задачу методом потенциалов.

	B1	B2	B3	Запас
A1	4	1	2	15
A2	3	3	2	17
A3	1	2	5	18
Спрос	20	20	10	50 50

Решение.

Задача является сбалансированной, так как суммарный запас поставщиков 50 и суммарный спрос потребителей 50 совпадают.

1. Находим начальный опорный план (базис) транспортных поставок методом минимальных тарифов. В ходе этого требуемый объем поставок 50 и спроса 50 должен быть распределен по m+n-1=3+3-1=5 базисным клеткам в первую очередь в клетки с минимальными тарифами.

Напомним, что базис-это связное ациклическое множество m+n–1 клеток таблицы транспортной задачи. Получаем план вида:

	B1		B2		B3		Запас
A1	4		1	15	2		15
A2	3	2	3	5	2	10	17
A3	1	18	2		5		18
Спрос	20		20		10		50 50

Имеем начальный базис B0 ={(1,2),(2,1),(2,2),(2,3),(3,1)}. Его цена $z_0 = z(B_0) = 1\cdot 15 + 3\cdot 2 + 3\cdot 5 + 2\cdot 10 + 1\cdot 18 = 74$.

2. Оптимизация начального плана метод потенциалов.

2.1. Первая итерация.

Находим потенциалы строк $u_i, i=1,2,3$ и потенциалы столбцов v_j, $j=1,2,3$ исходя из системы уравнений $u_i + v_j = c_{i,j}, (i,j) \in B_0$, $u_1 = 0$ и относительные оценки небазисных клеток $\Delta_{i,j} = c_{i,j} - (u_i + v_j) = c_{i,j}, (i,j) \notin B_0$.

Процесс расчета описан выше в теоретических сведениях.

Получаем таблицу:

	B1 v1=1		B2 v2=1		B3 v3=0		Запас
A1 u1=0	4	d11=3	1	15	2	d13=2	15
A2 u2=2	3	2	3	5	2	10	17
A3 u3=0	1	18	2	d32=1	5	d33=5	18
Спрос	20		20		10		50 50

Текущий план является оптимальным, так как относительные оценки всех небазисных клеток неотрицательны. Получено оптимальное решение т.е. оптимальное решение получено сразу, без итераций метода потенциалов).

Описание решения:

1. поставщик поставляет 15 ед. продукции второму потребителю;

2. поставщик поставляет 2 е. первому потребителю, 5 ед. второму и 10 ед. третьему;

3. поставщик поставляет 18 ед. первому потребителю.

Минимальные транспортные издержки составляют 74 денежные единицы.

Индивидуальные задания

№	

		B1	B2	B3	Запас
1		B1	B2	B3	Запас
	A1	4	1	2	15
	A2	3	3	2	17
	A3	1	3	5	18
	Спрос	20	20	10	50

					50
2		B1	B2	B3	Запас
	A1	4	1	2	15
	A2	3	3	2	17
	A3	1	2	1	18
	Спрос	20	20	10	50 50
3		B1	B2	B3	Запас
	A1	2	1	2	15
	A2	3	3	2	17
	A3	1	2	5	18
	Спрос	20	20	10	50 50
4		B1	B2	B3	Запас
	A1	4	1	2	15
	A2	3	3	2	17
	A3	1	2	2	18
	Спрос	20	20	10	50 50
5		B1	B2	B3	Запас
	A1	4	1	2	15
	A2	1	3	2	17
	A3	1	2	5	18
	Спрос	20	20	10	50 50
6		B1	B2	B3	Запас
	A1	4	1	2	15
	A2	3	3	2	17
	A3	1	2	3	18
	Спрос	20	20	10	50 50
7		B1	B2	B3	Запас

	A1	3	1	2	15
	A2	3	3	2	17
	A3	1	2	5	18
	Спрос	20	20	10	50 50
8		B1	B2	B3	Запас
	A1	4	4	2	15
	A2	3	3	2	17
	A3	1	2	5	18
	Спрос	20	20	10	50 50
9		B1	B2	B3	Запас
	A1	4	5	2	15
	A2	3	3	2	17
	A3	1	2	5	18
	Спрос	20	20	10	50 50
10		B1	B2	B3	Запас
	A1	4	1	2	15
	A2	3	3	2	17
	A3	5	2	5	18
	Спрос	20	20	10	50 50
11		B1	B2	B3	Запас
	A1	4	1	4	15
	A2	3	3	2	17
	A3	1	2	5	18
	Спрос	20	20	10	50 50
12		B1	B2	B3	Запас
	A1	4	1	2	15
	A2	3	3	2	17
	A3	4	2	5	18
	Спрос	20	20	10	50 50
13		B1	B2	B3	Запас
	A1	4	1	2	15
	A2	3	3	2	17
	A3	5	2	5	18

		Спрос	20	20	10	50 50

Контрольные вопросы

1. Записать математическую модель транспортной задачи.
2. Что такое опорный план транспортной задачи.
3. Описать метод северозападного угла.
4. Описать метод потенциалов.

Лабораторная работа 8. Задача выпуклого программирования. Метод проекции градиента

Цель работы: Изучить основные понятия выпуклой оптимизации с ограничениями, разобрать метод условного градиента, метод возможных направлений.

Задание

Данную нелинейную задачу с нелинейными ограничениями типа неравенств решить методом возможных направлений.

Краткие теоретические положения

Случай линейных ограничений. $f(x) \to \min$,

$a_i x \le b_i, i \in I_1$, $a_i x = b_i, i \in I_2$

Пусть x– некоторая допустимая точка, рассмотрим множество $I^0(x) = \{i \mid i \in I_1, a_i x = b_i\} \cup I_2$.

Множество ограничений активных в данной точке.

Найдем допустимое направление d перехода в следующую точку $\bar{x}$, при котором значение функции $f(x)$ убывает в наибольшей степени.

Пусть $A^0 = (a_i, i \in I^0)$-подматрица строк активных ограничений.

Пусть $rg(A^0) = q$, рассмотрим подпространство $S^0 = \{d \in R^n, A^0 d = 0\}$. Ясно, что $\dim(S^0) = n - q$.

Рассмотрим оптимизационную задачу:

$\nabla f^T(x) d \to \min$,

$A^0 d = 0$, (1) $\|d\| \le 1$.

Теорема. Пусть A^0– матрица полного ранга q.

Доказательство. Запишем ортогональное разложение антиградиента:

$-\nabla f(x) = y + z$, $y \in S^0$, $z \in S^{0T}$

Тогда $-\nabla f(x)^T d = y^T d + z^T d = y^T d$,

Ясно что $z = A^{0T}u, u \in R^q$ Тогда $-\nabla f(x) = y + A^{0T}u$, Так как $y \in S^0$, то $-A^0\nabla f(x) = A^0A^{0T}u + A^0 y = 0$.

Откуда $u = -\left(A^0A^{0T}\right)^{-1}A^0\nabla f(x)$

И, следовательно, $y = -\nabla f(x) + A^{0T}\left(A^0A^{0T}\right)^{-1}A^0\nabla f(x) = -P^0\nabla f(x)$

, где $P_0 = I - A_{0T}\left(A_0A_{0T}\right)^{-1}A_0$

Метод возможных направлений Начальный этап.

Выбрать начальную точку x_1 для которой выполняются неравенства , $g_i(x_1) \le 0, i = 1,...,m$ т.е. найти любую допустимую точку.

Положить $k = 1$ и перейти к основному этапу.

Основной этап.

Шаг 1.

Положить $I = \{i: g_i(x_k) = 0\}$ - множество активных ограничений в текущей точке и решить следующую задачу:

$$z \to \min;$$

$$\nabla f(x_k)^T d - z \le 0;$$

$$\nabla g_i(x_k)^T d - z \le 0, i \in I;$$

$$-1 \le d_j \le 1, j = 1,...,n \qquad (2)$$

Задача (2) является задачей линейного программирования и может эффективно быть решена. Например, симплекс-методом.

Пусть (z_k, d_k) - оптимальное решение. Если $z_k = 0$. То остановиться , x_k является точкой Ф.Джона, решение получено. Если $z_k < 0$, то перейти к шагу 2. Шаг 2. Взять в качестве λ_k оптимальное решение следующей задачи:

$$f(x_k + \lambda d_k) \to \min;$$

$$0 \le \lambda \le \lambda_{max},$$

Где $\lambda_{max} = \sup\{\lambda: g_i(x_k + \lambda d_k) \le 0, i = 1,...,m\}$. Положить $x_{k+1} = x_k + \lambda_k d_k$, заменить k на $k+1$ и перейти к шагу 1.

Пример выполнения

Выполнить три итерации метода возможных направлений. Первую итерацию сделать вручную, вторую и третью в пакете MathCad.

Рассмотрим задачу

$2 \cdot x_1^2 + 2 \cdot x_2^2 - 2 \cdot x_1 \cdot x_2 - 4 \cdot x_1$

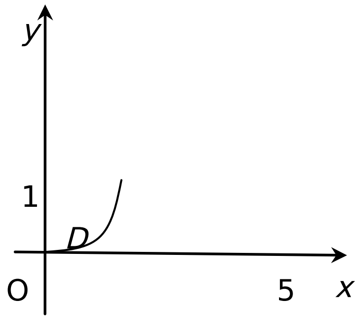
$-6 \cdot x_2$; $x_1 + 5 \cdot x_2 \le 5$;

$2 \cdot x_1^2 - x_2 \le 0$;

$-x_1 \le 0$;

$-x_2 \le 0$.

На геометрическом языке, речь идет о минимизации квадратичной функции двух переменных в области, расположенной в первом квадранте и ограниченной наклонной прямой и участком параболы.



<u>Начальный этап</u>.

Начнем процесс решения из начальной точки $x_1 = (0.00,\ 0.75)$ из чертежа видно, что данная начальная точка является допустимой.

Итерация 1.

А) Поиск направления.

В точке $x_1 = (0.00,\ 0.75)$ имеем

$\nabla f(x_1) = (4 x_1 - 2 x_2 - 4, 4 x_2 - 2 x_1 - 6) = (-5.5, -3)$;

А множество индексов активных ограничений $I = \{3\}$, так как точка x_1 расположена на граничной кривой $g_3(x_1, x_2) = -x_1 = 0$.

При этом $\nabla g_3(x_1) = (-1, 0)$;

Задача определения направления приобретает вид: $z \to \min$;

$-5.5 d_1 - 3 d_2 - z \le 0$;

$-d_1 - z \le 0$

$-1 \le d_j \le 1,\ j = 1,2$

Сведем эту задачу к стандартной задаче линейного программирования и решим симплекс-методом. Пусть $q_1 = d_1 + 1$, $q_2 = d_2 + 1$, $z_1 = -z + 1$. Получаем задачу: $z_1 \to maz$;

$-5.5\cdot(q_1 - 1) - 3\cdot(q_2 - 1) + z_1 \le 0;$

$-(q_1 - 1) + z_1 \le 0$

$0 \le q_j \le 2, j = 1,2$

Которую можно переписать в виде:

$z_1 \to \max;$

$-5.5q_1 - 3q_2 + z_1 \le 7.5;$

$-q_1 + z_1 \le 0$

$0 \le q_j \le 2, j = 1,2$

Приводим эту задачу к каноническому виду:

$w \to \max;$

$w - z_1 = 0;$

$-5.5q_1 - 3q_2 + z_1 + s_1 = 7.5;$

$-q_1 + z_1 + s_2 = 0$

$q_1 + s_3 = 2;$ $q_2 + s_4 = 2;$

$z_1, s_1, s_2, s_3, s_4 \ge 0.$

Применяем симплекс-метод в табличной форме:

Б	w	z_1	q_1	q_2	s_1	s_2	s_3	s_4	Реш	bi/aij
w	1	-1	0	0	0	0	0	0	0	-
s_1	0	1	-5.5	-3	1	0	0	0	7.5	7.5
s_2	0	1	-1	0	0	1	0	0	0	0
s_3	0	0	1	0	0	0	1	0	2	-
s_4	0	0	0	1	0	0	0	1	2	-

Б	w	z_1	q_1	q_2	s_1	s_2	s_3	s_4	Реш	bi/aij
w	1	0	-1	0	0	1	0	0	0	-
s_1	0	0	-4.5	-3	1	-1	0	0	7.5	-
z_1	0	1	-1	0	0	1	0	0	0	-
s_3	0	0	1	0	0	0	1	0	2	2
s_4	0	0	0	1	0	0	0	1	2	-

Б	w	z_1	q_1	q_2	s_1	s_2	s_3	s_4	Реш	bi/aij
w	1	0	0	0	0	1	1	0	2	-
s_1	0	0	0	-3	1	-1	4.5	0	16.5	-
s_2	0	1	0	0	0	1	1	0	2	-
q_1	0	0	1	0	0	0	1	0	2	2
s_4	0	0	0	1	0	0	0	1	2	-

Получено оптимальное решение:

z_{min}□ □(w_{max}□1) □ □1;

q_1□ 2□ d_1□1; q_2□ 0□ d_2

□ □1.

Итак, направление поиска минимума на данной итерации d □ □1,□1□.

Координаты произвольной точки вдоль этого направления имеют вид: x□□□□□ □□, 0.75□□□, при этом значение функции f_1□x□□□□□□□, 0.75□□□=

□ 2□□^{2} □ 2□□0.75□□□^{2} □ 2□□□□0.75□□□□ 4□□□6□□0.75□□□□□

□ 2□□^{2} □1.125□3□□□ 2□□^{2} □1.5□□□ 2□□^{2} □ 4□□□ 4.5□ 6□□□

□ 6□□^{2} □ 2.5□□□3.475

Определи величину □$_{max}$из неравенств: x_1

□5□x_2 □□□5□□0.75□□□□ 3.75□4□□□ 5;

2□x_1^2 □ x_2 □ 2□□^{2} □□0.75□□□□ 0;

□□□ 0;

□□0.75□□□□   0.

Или:

□□ □1.25 / 4 ;

2□□^{2} □□□0.75 □ 0 □□□□□0.914,0.414□□ 0□□□ 0.414.

□□ 0; □□ 0.75.

Получаем: □$_{max}$ □0.414.

Решаем одномерную задачу минимизации:

6□□^{2} □2.5□□□3.475□min;

□□□0,0.414□

Находим координаты вершины параболы:

$$\lambda_0 = \frac{2.5}{12} = 0.208 < 0.414 \Rightarrow \lambda_{opt} = 0.208$$

Получаем точку второй итерации: $x_2 = \left(\lambda, 0.75 - \lambda\right) = \left(0.208, 0.542\right)$.

Значение целевой функции при первой итерации уменьшилось с $f_1 = -3.375$ до $f_1 = -3.635$

Итерация 2.

Выполним данную итерацию в пакете MathCad.

$$f(x1,x2) := 2\cdot x1^2 + 2\cdot x2^2 - 2\cdot x1\cdot x2 - 4\cdot x1 - 6\cdot x2$$

$$F(P) := f\left(P_{1,1}, P_{1,2}\right)$$

$$gradf(x1,x2) := \left(\frac{d}{dx1}f(x1,x2) \quad \frac{d}{dx2}f(x1,x2)\right)$$

$$gradf(x1,x2) \text{ simplify } \rightarrow (4\cdot x1 - 2\cdot x2 - 4 \quad 4\cdot x2 - 2\cdot x1 - 6)$$

$$P2 := (0.208 \quad 0.542)$$

$$Gr(P) := gradf\left(P_{1,1}, P_{1,2}\right)$$

$$Gr(P2) = (-4.252 \quad -4.248)$$

Вектор-функция ограничений

$$g(P) := \begin{bmatrix} P_{1,1} + 5\cdot P_{1,2} - 5 \\ 2\cdot\left(P_{1,1}\right)^2 - P_{1,2} \\ -P_{1,1} \\ -P_{1,2} \end{bmatrix} \qquad gradg(P) := \begin{pmatrix} 1 & 5 \\ 4\cdot P_{1,1} & -1 \\ -1 & 0 \\ 0 & -1 \end{pmatrix}$$

Определение активных ограничений

```
Act(P,g,eps) := | k ← 0
                | m ← rows(g(P))
                | for i ∈ 1..m
                |   if g(P)_i > -eps
                |      | k ← k + 1
                |      | org_k ← i
                | (k  org)^T
```

Число активных ограничений в данной точке

$$\begin{pmatrix} Nog \\ Ogr \end{pmatrix} := Act(P2, g, 0.0000000001)$$

$Nog = 0$ активныхограниченийвданнойточке нет

Задача линейного программирования по нахождению направления поиска

минимизировать z

при условиях $-4.25d_1 - 4.25d_2 - z \leqslant 0,$

$-1 \leqslant d_j \leqslant 1, \quad j = 1, 2.$

$w(z, d1, d2) := z$

$d1 := 0$ $d2 := 0$ $z := 0$

Given

$d1 \le 1$ $d1 \ge -1$ $d2 \le 1$ $d2 \ge -1$

$$Gr(P2) \cdot \begin{pmatrix} d1 \\ d2 \end{pmatrix} - z \le 0$$

$$\begin{pmatrix} z \\ d1 \\ d2 \end{pmatrix} := Minimize(w, z, d1, d2)$$

$z = -8.5$ $d1 = 1$ $d2 = 1$

Линейный поиск

$x(\lambda) := P2 + \lambda \cdot (d1 \quad d2)$

$FL(t) := f\left(P2_{1,1} + d1 \cdot t, P2_{1,2} + d2 \cdot t\right)$

$t := 0$

Given

$t \ge 0$

$g(x(t))_{1,1} \le 0$ $\quad g(x(t))_{2,1} \le 0$

$g(x(t))_{3,1} \le 0$ $\quad g(x(t))_{4,1} \le 0$

$tm := \mathrm{Minimize}(FL, t)$

$P3 := x(tm)$

$$g(P3) = \begin{pmatrix} 0 \\ -0.273 \\ -0.555 \\ -0.889 \end{pmatrix}$$ Проверка допустимости

$P3 = (0.555 \quad 0.889)$

$F(P3) - F(P2) = -2.709$ Эффективность шага

Ттетья итерация

Число активных ограничений в данной точке

$$\begin{pmatrix} Nog \\ Ogr \end{pmatrix} := \mathrm{Act}(P3, g, 0.0000000001)$$

$Nog = 1$ одно активное ограничения

$Ogr = (1)$ это первое ограничения

Задача линейного программирования по нахождению направления поиска

$w(z, d1, d2) := z$

$d1 := 0$ $\quad d2 := 0$ $\quad z := 0$

Given

$d1 \le 1$ $\quad d1 \ge -1$ $\quad d2 \le 1$ $\quad d2 \ge -1$

$$Gr(P3)\cdot\begin{pmatrix} d1 \\ d2 \end{pmatrix} - z \le 0$$

$$\left[gradg(P3)\cdot\begin{pmatrix} d1 \\ d2 \end{pmatrix}\right]_1 - z \le 0$$

$$\begin{pmatrix} z \\ d1 \\ d2 \end{pmatrix} := \mathrm{Minimize}(w, z, d1, d2)$$

$z = -1.664$ $\quad$ $d1 = 1$ $\quad$ $d2 = -0.533$

Линейный поиск

$x(\lambda) := P3 + \lambda\cdot(\,d1 \quad d2\,)$

$FL(t) := f\left(P3_{1,1} + d1\cdot t, P3_{1,2} + d2\cdot t\right)$

$t := 0$

Given

$t \ge 0$

$g(x(t))_{1,1} \le 0$ $\quad$ $g(x(t))_{2,1} \le 0$

$g(x(t))_{3,1} \le 0$ $\quad$ $g(x(t))_{4,1} \le 0$

$tm := \mathrm{Minimize}(FL, t)$

$P4 := x(tm)$

Проверка допустимости-допустимо с точностью 0.0000001

$$g(P4) = \begin{pmatrix} -0.155 \\ 8.022\times 10^{-8} \\ -0.648 \\ -0.84 \end{pmatrix}$$

$P4 = (0.648 \;\; 0.84)$

$F(P4) - F(P3) = -0.123$ Эффективность шага

$F(P4) = -6.467$

Точное решение в пакете MathCad

$P := (0 \;\; 0)$

Given

$g(P)_{1,1} \le 0$ $\qquad$ $g(P)_{2,1} \le 0$

$g(P)_{3,1} \le 0$ $\qquad$ $g(P)_{4,1} \le 0$

$Pm := \mathrm{Minimize}(F, P)$

$Pm = (0.659 \;\; 0.868)$

$Fm := F(Pm)$

$Fm = -6.613$

Индивидуальные задания

№	Условие	№	Условие
1	$2\square x_1^2 \square 3\square x_2^2 \square 2\square x_1 \square x_2 \square 4\square x_1 \square 6\square x_2$; $x_1 \square 5\square x_2 \square 8$; $2\square x_1^2 \square x_2 \square 0$; $\square x_1 \square 0$; $\square x_2 \square 0$.	2	$2\square x_1^2 \square 2\square x_2^2 \square 2\square x_1 \square x_2 \square 3\square x_1 \square 8\square x_2$; $x_1 \square 5\square x_2 \square 5$; $2\square x_1^2 \square x_2 \square 0$; $\square x_1 \square 0$; $\square x_2 \square 0$.
3	$2\square x_1^2 \square 2\square x_2^2 \square 2\square x_1 \square x_2 \square 4\square x_1 \square 6\square x_2$; $2\square x_1 \square 5\square x_2 \square 5$; $2\square x_1^2 \square x_2 \square 0$; $\square x_1 \square 0$; $\square x_2 \square 0$.	4	$2\square x_1^2 \square 2\square x_2^2 \square 5\square x_1 \square x_2 \square 5\square x_1 \square 8\square x_2$; $x_1 \square 5\square x_2 \square 5$; $2\square x_1^2 \square x_2 \square 0$; $\square x_1 \square 0$; $\square x_2 \square 0$.
5	$3\square x_1^2 \square 2\square x_2^2 \square 2\square x_1 \square x_2 \square 4\square x_1 \square 6\square x_2$; $x_1 \square 5\square x_2 \square 12$; $2\square x_1^2 \square x_2 \square 0$;	6	$2\square x_1^2 \square 4\square x_2^2 \square 2\square x_1 \square x_2 \square 5\square x_1 \square 8\square x_2$; $x_1 \square 5\square x_2 \square 5$; $2\square x_1^2 \square x_2 \square 0$;

	$\square x_1 \square 0;$ $\square x_2 \square 0.$		$\square x_1 \square 0;$ $\square x_2 \square 0.$

7	$2\square x_1^2 \square 2\square x_2^2 \square 2\square x_1 \square x_2 \square 4\square x_1 \square 6\square x_2; x_1 \square 5\square x_2 \square 5;$ $2\square x_1^2 \square x_2 \square 0;$ $\square x_1 \square 0;$ $\square x_2 \square 0.$	8	$2\square x_1^2 \square 2\square x_2^2 \square 2\square x_1 \square x_2 \square 5\square x_1 \square 8\square x_2; x_1 \square 5\square x_2 \square 8;$ $2\square x_1^2 \square x_2 \square 0;$ $\square x_1 \square 0;$ $\square x_2 \square 0.$
9	$2\square x_1^2 \square 2\square x_2^2 \square 2\square x_1 \square x_2 \square 4\square x_1 \square 6\square x_2; x_1 \square 5\square x_2 \square 5;$ $2\square x_1^2 \square x_2 \square 0;$ $\square x_1 \square 0;$ $\square x_2 \square 0.$	10	$2\square x_1^2 \square 2\square x_2^2 \square 2\square x_1 \square x_2 \square 5\square x_1 \square 8\square x_2; x_1 \square 5\square x_2 \square 5;$ $2\square x_1^2 \square x_2 \square 0;$ $\square x_1 \square 0;$ $\square x_2 \square 0.$
11	$2\square x_1^2 \square 2\square x_2^2 \square 2\square x_1 \square x_2 \square 4\square x_1 \square 6\square x_2; x_1 \square 4\square x_2 \square 5;$ $2\square x_1^2 \square x_2 \square 0;$ $\square x_1 \square 0;$ $\square x_2 \square 0.$	12	$2\square x_1^2 \square 2\square x_2^2 \square 3\square x_1 \square x_2 \square 4\square x_1 \square 8\square x_2; x_1 \square 5\square x_2 \square 5;$ $2\square x_1^2 \square x_2 \square 0;$ $\square x_1 \square 0;$ $\square x_2 \square 0.$
13	$2\square x_1^2 \square 2\square x_2^2 \square 2\square x_1 \square x_2 \square 4\square x_1 \square 6\square x_2; x_1 \square 5\square x_2 \square 5;$ $3\square x_1^2 \square x_2 \square 0;$ $\square x_1 \square 0;$ $\square x_2 \square 0.$	14	$2\square x_1^2 \square 2\square x_2^2 \square 2\square x_1 \square x_2 \square 4\square x_1 \square 8\square x_2; x_1 \square 6\square x_2 \square 5;$ $2\square x_1^2 \square x_2 \square 0;$ $\square x_1 \square 0;$ $\square x_2 \square 0.$
15	$2\square x_1^2 \square 2\square x_2^2 \square 6\square x_1 \square x_2 \square 4\square x_1 \square 6\square x_2; x_1 \square 5\square x_2 \square 5;$ $2\square x_1^2 \square x_2 \square 0;$ $\square x_1 \square 0;$ $\square x_2 \square 0.$	16	$2\square x_1^2 \square 6\square x_2^2 \square 2\square x_1 \square x_2 \square 4\square x_1 \square 8\square x_2; x_1 \square 5\square x_2 \square 5;$ $2\square x_1^2 \square x_2 \square 0;$ $\square x_1 \square 0;$ $\square x_2 \square 0.$
17	$2\square x_1^2 \square 2\square x_2^2 \square 2\square x_1 \square x_2 \square 4\square x_1 \square 8\square x_2; x_1 \square 5\square x_2 \square 5;$ $2\square x_1^2 \square x_2 \square 0;$	18	$5\square x_1^2 \square 2\square x_2^2 \square 2\square x_1 \square x_2 \square 4\square x_1 \square 8\square x_2; x_1 \square 5\square x_2 \square 5;$ $2\square x_1^2 \square x_2 \square 0;$

	$\square x_1 \square\ 0;$ $\square x_2 \square\ 0.$		$\square x_1 \square\ 0;$ $\square x_2 \square\ 0.$
19	$2\square\ x_1^2 \square\ 2\square\ x_2^2 \square\ 2\square\ x_1 \square\ x_2 \square\ 4\square\ x_1 \square$ $6\square\ x_2;\ x_1 \square 5\square\ x_2 \square\ 4;$ $2\square\ x_1^2 \square\ x_2 \square\ 0;$ $\square x_1 \square\ 0;$ $\square x_2 \square\ 0.$	20	$7\square\ x_1^2 \square\ 2\square\ x_2^2 \square\ 2\square\ x_1 \square\ x_2 \square\ 4\square\ x_1$ $\square 8\square\ x_2;\ x_1 \square 5\square\ x_2 \square\ 5;$ $2\square\ x_1^2 \square\ x_2 \square\ 0;$ $\square x_1 \square\ 0;$ $\square x_2 \square\ 0.$
21	$2\square\ x_1^2 \square\ 2\square\ x_2^2 \square\ 2\square\ x_1 \square\ x_2 \square\ 4\square\ x_1 \square$ $6\square\ x_2;$ $3\square\ x_1 \square 5\square\ x_2 \square\ 5;$ $2\square\ x_1^2 \square\ x_2 \square\ 0;$ $\square x_1 \square\ 0;$ $\square x_2 \square\ 0.$	22	$2\square\ x_1^2 \square\ 2\square\ x_2^2 \square\ 2\square\ x_1 \square\ x_2 \square 5\square\ x_1$ $\square 8\square\ x_2;\ x_1 \square 5\square\ x_2 \square\ 5;$ $4\square\ x_1^2 \square\ x_2 \square\ 0;$ $\square x_1 \square\ 0;$ $\square x_2 \square\ 0.$
23	$2\square\ x_1^2 \square 3\square\ x_2^2 \square\ 2\square\ x_1 \square\ x_2 \square\ 4\square\ x_1 \square$ $6\square\ x_2;\ x_1 \square 5\square\ x_2 \square\ 5;$ $2\square\ x_1^2 \square\ x_2 \square\ 0;$ $\square x_1 \square\ 0;$ $\square x_2 \square\ 0.$	24	$4\square\ x_1^2 \square\ 2\square\ x_2^2 \square\ 2\square\ x_1 \square\ x_2 \square\ 4\square\ x_1 \square$ $6\square\ x_2;\ x_1 \square 5\square\ x_2 \square\ 5;$ $2\square\ x_1^2 \square\ x_2 \square\ 0;$ $\square x_1 \square\ 0;$ $\square x_2 \square\ 0.$
25	$4\square\ x_1^2 \square\ 2\square\ x_2^2 \square\ 2\square\ x_1 \square\ x_2 \square\ 4\square\ x_1 \square$ $6\square\ x_2;\ x_1 \square 5\square\ x_2 \square\ 5;$ $2\square\ x_1^2 \square\ x_2 \square\ 0;$ $\square x_1 \square\ 0;$ $\square x_2 \square\ 0.$	26	$2\square\ x_1^2 \square\ 2\square\ x_2^2 \square\ 2\square\ x_1 \square\ x_2 \square 5\square\ x_1$ $\square 8\square\ x_2;\ x_1 \square 5\square\ x_2 \square\ 7;$ $2\square\ x_1^2 \square\ x_2 \square\ 0;$ $\square x_1 \square\ 0;$ $\square x_2 \square\ 0.$

Контрольные вопросы

1. Как записывается задача нелинейного программирования с ограничениями.
2. Какое направление называется возможным.
3. Как записывается задача линейного программирования о поиске наилучшего возможного направления.
4. Привести формулу проекции антиградиента на линейное многообразие.

Лабораторная работа 9. Задача о распределении ресурсов. Метод динамического программирования.

Цель работы: Изучение основных положений метода динамического программирования и освоение практических способов решения конкретных задач.

Задание

Имеется данная сумма S средств для осуществления капиталовложений в начале года в данную группу из n предприятий. Величина S измеряется в некоторых целых условных единицах и возможный уровень x капиталовложений в некоторое предприятие является целым числом в пределах от 0 до S включительно. Эффективность капиталовложения x в каждое из предприятий характеризуется набором функций $f_i(x), i=1,...,n$.

Таким образом мы должны найти оптимальные уровни капиталовложений $x^*_i, i=1,...,n$, при которых достигается максимальная суммарная эффективность капиталовложений, таким образом, мы должны решить оптимизационную задачу:

$$Z(x_1,...,x_n) = f_1(x_1)+...+f_n(x_n) \to \max; \quad x_1+...+x_n \le S; \quad 0 \le x_i, i=1,...,n.$$

Краткие теоретические положения

Пусть у некоторой оптимизируемой системы имеется параметр состояния $y \in Y$, где $Y=\{s_1,...,s_m\}$- конечное множество всех возможных состояний, система подвергается воздействию вектора управления $\bar{x}=(x_1, x_2, \ldots, x_n)$, проходя, таким образом, этапов, на каждом этапе $k=1,...,n$ получается аддитивный вклад $z_k = f(x_k)$ в суммарный функционал качества

$z(x_1,...,x_n)=\sum_{k=1}^{n} f_k(x_k)$, который подлежит максимизации путем выбора оптимального вектора управления $\bar{x}^*=(x_1^*, x_2^*, \ldots, x_m^*)$. При этом система подчиняется некоторому уравнению состояния при каждом переходе: $y_k = g_k(y_{k-1}, x_k), k=1,...,n$, где y_{k-1}- состояние системы перед шагом k, x_k – решение, принимаемое на шаге $k=1,...,n$, где $x_k \in X_k = X(k, y_{k-1})$ – конечное множество возможных решений на шаге k, зависящее от состояния системы перед эти шагом и номера шага., при этом $y_0 \in Y$ – заданное исходное состояние системы перед первым шагом.

Метод динамического программирования состоит из двух этапов: обратного и прямого хода. Более сложным и ресурсоемким по памяти и процессорному времени является обратный ход, при котором последовательно с конца строятся векторы $Z_k(s_i)$, $X_k(s_i), i = n,...,1$, где $Z_k(s_i)$ – суммарная прибыль на последних nk+1 этапах k,k+1,…,n и $X_k(s_i) = x_k^*(s_i)$ – оптимальное решение, принимаемое на kом этапе в зависимости от состояния s_i системы перед этим этапом. При этом используются следующие уравнения Беллмана:

$Z_n(s_i) = \max\limits_{x \in X(n,s_i)} f_n(x)$, $X_n(s_i) = \arg\max\limits_{x \in X(n,s_i)} f_n(x)$, – для последнего этапа и

$$Z_k(s_i) = \max_{x \in X(k,s_i)} f_k(x) + Z_{k+1}(g_k(s_i, x)), k = n-1,...,1 \quad (1)$$

$$X_k(s_i) = \arg\max_{x \in X(k,s_i)} f_k(x) + Z_{k+1}(g_k(s_i, x)), k = n-1,...,1 \quad (2)$$

Для предыдущих этапов, т.е. после инициализации векторов $Z_k(s_i)$, $X_k(s_i), i = n,...,1$ для последнего этапа, они далее рассчитываются по реккурентным формулам (1),(2) на этапах k = n–1,..,1.

Рассмотрим далее конкретный пример задачи об оптимальном распределении ресурсов между тремя предприятиями.

Пусть $Z_k^*(s), k = 1,...,n, 0 \le s \le S$- максимальный доход от выделения суммы $0 \le s \le S$ предприятиям $k, k = 1,...,n$. Имеем $Z_n^*(s) = f_n(s)$ - известная функция. Для $k = n-1, n-2,...,1$ используем уравнение Беллмана:

$$Z_k^*(s) = \max_{x \in [0,s]} \{f_k(x) + Z_{k+1}^*(s-x)\} \quad (3)$$

Пусть также

$$x_k(s) = \operatorname{argmax}_{x \in [0,s]} \{f_k(x) + Z_{k+1}^*(s-x)\} \quad (4)$$

– уровень капиталовложений x, при котором в правой части формулы (3) достигается максимум.

Функции $Z_k^*(s), x_k(s)$ на практике рассчитываются табличным способом с конца, начиная с известной функции $Z_n(s)$ при обратном ходе алгоритма.

С помощью уравнений (3-4) оптимальный вектор $\bar{x}^* = (x_1^* \; ... \; x_n^*)$ с максимальным значением эффективности $Z^* = f_1(x_1^*) + ... + f_n(x_n^*)$ строится в соответствии с прямым ходом метода динамического программирования:

$x_1^* = x_1(S)$;

... (5)

$$x_k^* = x_k\left(S - \sum_{i=1}^{k-1} x_i^*\right), k = 2,...,n$$

Пример выполнения

Начальная сумма S = 3 капитала распределяется между тремя предприятиями, для которых эффективность капиталовложений задается следующей таблицей.

x	$f_1(x)$	$f_2(x)$	$f_3(x)$
0	0	0	0
1	0.5	0.1	0.6
2	1.0	0.5	1.1
3	1.4	1.8	1.5

Найти оптимальный план капиталовложений в целом.

Решение Обратный ход

Оптимизация на третьем шаге

s	$Z_3^*(s)$	$X_3(s)$
0	0	0
1	0.6	1
2	1.1	2
3	1.5	3

Оптимизация на втором шаге

s	x_2	$Z_2(s,x_2) = f_2(x_2) + Z_3^*(s - x_2)$	$Z_2^*(s)$	$X_2(s)$
0	0	0	0	0
1	0	$f_2(0)+Z_3^*(1)$=0+0.6=0.6	Max(0.6,0.1)=0.6	0
	1	$f_2(1)+Z_3^*(0)$=0.1+0=0.1		
2	0	$f_2(0)+Z_3^*(2)$=0+1.1=1.1	Max(1.1,0.7,0.5)=1.1	0
	1	$f_2(1)+Z_3^*(1)$=0.1+0.6=0/.7		
	2	$f_2(2)+Z_3^*(0)$=0.5+0=0.54		
3	0	$f_2(0)+Z_3^*(3)$=0+1.5=1.5	Max(1.5,1.2,1.1,1.8)=1.8	3
	1	$f_2(1)+Z_3^*(2)$=0.1+1.1=1.2		
	2	$f_2(2)+Z_3^*(1)$=0.5+0.6=1.1		
	3	$f_2(3)+Z_3^*(0)$=1.8+0=1.8		

Оптимизация на первом шаге

s	x_1	$Z_1(s,x_1) = f_1(x_1) + Z_2^*(S -$	$Z_1^*(s)$	$X_1(s)$

		x_1		
0	0	0	0	0
1	0	$f_1(0)+Z_2^*(1)=0+0.6=0.6$	Max(0.6,0.5)=0.6	0
	1	$f_1(1)+Z_2^*(0)=0.5+0=0.5$		
2	0	$f_1(0)+Z_2^*(2)=0+1.1=1.1$	Max(1.1,1.1,1.0)=1.1	0
	1	$f_1(1)+Z_2^*(1)=0.5+0.6=1.1$		
	2	$f_1(2)+Z_2^*(0)=1.0+0=1.0$		
3	0	$f_1(0)+Z_2^*(3)=0+1.8=1.8$	Max(1.8,1.6,1.6,1.4)=1.8	0
	1	$f_1(1)+Z_2^*(2)=0.5+1.1=1.6$		
	2	$f_1(2)+Z_2^*(1)=1.0+0.6=1.6$		
	3	$f_1(3)+Z_2^*(0)=1.4+0=1.4$		

<u>Прямой ход</u> Исходный каптал S=3.

Оптимальное выделение средств первому предприятию: $x_1^* = X_1(3) = 0$

Капитал после выделения средств первому предприятию:

$S_1 = S - x_1^* = 3 - 0 = 3$;

Оптимальное выделение средств второму предприятию:

$x_2^* = X_2(S_1) = x_2(3) = 3$; Капитал после выделения средств второму предприятию: $S_2 = S_1 - x_2^* = 3 - 3 = 0$;

Оптимальное выделение средств третьему предприятию:

$x_3^* = X_3(S_2) = x_3^*(0) = 0$;

Ответ:

Оптимальная программа выделения средств: $x^* = (x_1^* \;\; x_2^* \;\; x_3^*) = (0 \;\; 3 \;\; 0)$;

Максимальный экономический эффект:

$Z^* = Z(x^*) = f_1(x_1^*) + f_2(x_2^*) + f_3(x_3^*) = f_1(0) + f_2(3) + f_3(0) = 0 + 1.8 + 0 = 1.8$

Использование ЭВМ

Программа MathCad

```
Rasp_res(S, f) :=  n ← cols(f)
                   for k ∈ n, n − 1 .. 1
                     for s ∈ 0 .. S            if k = n
                       Z[s, k−1] ← f[s, k−1]
                       X[s, k−1] ← s
                     for s ∈ 0 .. S            otherwise
                       for x ∈ 0 .. S
                         H[x] ← 0
                       for x ∈ 0 .. s
                         H[x] ← f[x, k−1] + Z[s−x, k]
                       Z[s, k−1] ← max(H)
                       for x ∈ 0 .. s
                         if Z[s, k−1] = H[x]
                           X[s, k−1] ← x
                           break
                   for i ∈ 1 .. n
                     x[i−1] ← X[S, i−1]
                     S ← S − x[i−1]
                   (Z
                    X
                    x)
```

Исходные данные

x	$f_1(x)$	$f_2(x)$	$f_3(x)$
0	0	0	0
1	0.5	0.1	0.6
2	1.0	0.5	1.1
3	1.4	1.8	1.5

$$f := \begin{pmatrix} 0 & 0 & 0 \\ 0.5 & 0.1 & 0.6 \\ 1.0 & 0.5 & 1.1 \\ 1.4 & 1.8 & 1.5 \end{pmatrix}$$

Выполнение программы

$$R := Rasp\_res(3, f)$$

$$R_0 = \begin{pmatrix} 0 & 0 & 0 \\ 0.6 & 0.6 & 0.6 \\ 1.1 & 1.1 & 1.1 \\ 1.8 & 1.8 & 1.5 \end{pmatrix}$$

$$R_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 2 \\ 0 & 3 & 3 \end{pmatrix}$$

$$R_2 = \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix}$$

Контрольные вопросы

1. Записать оптимизационную задачу об оптимальном распределении капиталовложений.
2. Выписать уравнения Беллмана.
3. Выписать уравнения прямого хода для задачи динамического программирования.