Федеральное агентство по образованию

Московский авиационный институт
(государственный технический университет)

Кафедра 304
(вычислительные машины, системы и сети)

Лабораторная работа по курсу «Автоматизация проектирования»

Отчёт по работе *№1*.

Методы решения задач безусловной оптимизации
(наименование работы)

Вариант задания *№2*.

Лабораторную работу выполнил:

студент гр. 13-501, Резвяков Денис Михайлович ______________
(должность) (Ф. И. О.) (подпись)

Лабораторную работу принял:

доцент каф. 304, Силаева Татьяна Александровна ______________.
(должность) (Ф. И. О.) (подпись)

«___» ______________ 2010 г.
(дата приёма)

Цель работы: Изучить и практически овладеть основными методами решения задач безусловной оптимизации.

Задание

Безусловно оптимизировать функцию: $\underset{X}{\operatorname{extr}} f(X)$,

$$f(X) = x_1^3 + x_2^2 + x_3^2 - 3x_1 - 2x_2 - 4x_3 + 8.$$

Порядок выполнения работы

1. Найти решение задачи безусловной оптимизации функции *f(X)* с помощью следующих методов: *а)* классического; *б)* Ньютона; *в)* градиентного спуска при заданном h = 0,2; *г)* наискорейшего спуска.

В трёх итерационных методах выполнить вручную только по одной итерации, используя одинаковое значение $X^{(0)}$, компоненты которого должны отличаться от найденного по классическому методу экстремума на +1.

2. Провести сравнительный анализ полученных на первой итерации результатов решения заданной задачи безусловной оптимизации функции с помощью трёх итерационных методов, сформулировать выводы.

3. Выполнить на ЭВМ расчёт безусловного экстремума функции *f(X)* с точностью $\varepsilon = 10^{-4}$ согласно методам: Ньютона и трём градиентным, включая метод сопряжённых градиентов.

4. Выполнить сравнительный анализ результатов решения одной и той же задачи безусловной оптимизации функции с помощью пяти используемых в работе методов и сформулировать выводы.

1. Выполнение первых итераций для 4х методов вручную

а) Классический метод:

Необходимое условие экстремума функции $f(X)$:

$$\begin{cases} \frac{\partial f}{\partial x_1}(X) = 3x_1^2 - 3 = 0, \\ \frac{\partial f}{\partial x_2}(X) = 2x^2 - 2 = 0, \\ \frac{\partial f}{\partial x_3}(X) = 2x_3 - 4 = 0. \end{cases} \Rightarrow \begin{cases} x_1 = \sqrt{\frac{3}{3}} = \pm 1, \\ x_2 = \frac{2}{2} = 1, \\ x_3 = \frac{4}{2} = 2. \end{cases} \Rightarrow X_1 = (1;1;2) \text{ и } X_2 = (-1;1;2).$$

Достаточное условие экстремума в точках X_1 и X_2:

матрица Гессе $A(X) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(X)\right] \Rightarrow \begin{bmatrix} 6x_1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \Rightarrow$ $\Delta_3 = 6x_1 \cdot 2 \cdot 2 = 24x_1$, $\Delta_2 = 6x_1 \cdot 2 = 12x_1$, $\Delta_1 = 6x_1$.

- для точки X_1: $\Delta_3 = 24$, $\Delta_2 = 12$, $\Delta_1 = 6$ — точка минимума.
- для точки X_2: $\Delta_3 = -24$, $\Delta_2 = -12$, $\Delta_1 = -6$ — точка перегиба.

$$f(X_1) = 1^3 + 1^2 + 2^2 - 3\cdot 1 - 2\cdot 1 - 4\cdot 2 + 8 = 1; \quad f(X_2) = 5.$$

б) Метод Ньютона:

Найдём $\varphi_i(X) = \frac{\partial f}{\partial x_i}(X)$: $\varphi_1(X) = 3x_1^2 - 3$, $\varphi_2(X) = 2x_2 - 2$, $\varphi_3(X) = 2x_3 - 4$.

Пусть начальная точка будет: $X^{(0)} = (1+1; 1+1; 2+1) = (2;\ 2;\ 3)$ — по заданию.

Формируем систему уравнений: $\varepsilon = 10^{-4}$; $f(X^{(0)}) = 7$.

$$\begin{cases} 6x_1^{(k)} t_1^{(k)} = -3\left[x_1^{(k)}\right]^2 + 3, \\ 2t_2^{(k)} = -2x_2^{(k)} + 2, \\ 2t_3^{(k)} = -2x_3^{(k)} + 4. \end{cases} \Rightarrow \begin{cases} 6\cdot 2\cdot t_1^{(0)} = -3\cdot 2^2 + 3 = -9, \\ 2\cdot t_2^{(0)} = -2\cdot 2 + 2 = -2, \\ 2\cdot t_3^{(0)} = -2\cdot 3 + 4 = -2. \end{cases} \Rightarrow \begin{cases} t_1^{(0)} = -\frac{9}{12}, \\ t_2^{(0)} = -\frac{2}{2}, \\ t_3^{(0)} = -\frac{2}{2}. \end{cases} \Rightarrow T^{(0)} = \left(-\frac{3}{4}; -1; -1\right).$$

Найдём $X^{(1)}$: $X^{(1)} = X^{(0)} + T^{(0)} = (1{,}25;\ 1;\ 2)$. $f(X^{(1)}) = 3{,}1094$.

в) Метод градиентного спуска:

$X^{(0)}=(2;\ 2;\ 3)$, $\varepsilon=10^{-4}$, $h^{(k)}=h=0{,}2$ (по заданию).

Градиент: $\nabla f(X)=\begin{pmatrix}3x_1^2-3\\2x_2-2\\2x_3-4\end{pmatrix}$, $\nabla f(X^{(0)})=\begin{pmatrix}3\cdot2^2-3\\2\cdot2-2\\2\cdot3-4\end{pmatrix}=\begin{pmatrix}9\\2\\2\end{pmatrix}$,

$\|\nabla f(X^{(0)})\|=9{,}434$; $f(X^{(0)})=7$.

Найдём $X^{(1)}$ для следующей итерации: $X^{(k+1)}=X^{(k)}-h\nabla f(X^{(k)})$.

$$X^{(1)}=\begin{pmatrix}2\\2\\3\end{pmatrix}-0{,}2\begin{pmatrix}9\\2\\2\end{pmatrix}=\begin{pmatrix}0{,}2\\1{,}6\\2{,}6\end{pmatrix},\quad \nabla f(X^{(1)})=\begin{pmatrix}-2{,}88\\1{,}2\\1{,}2\end{pmatrix};\ \|\nabla f(X^{(1)})\|=3{,}343;$$
$$f(X^{(1)})=3{,}128.$$

г) Метод наискорейшего спуска:

Начальные вычисления совпадают с предыдущим методом до определения $h^{(0)}$. Найдём $h^{(0)}$: $z(h^{(0)})=f(X^{(1)})=f(X^{(0)}-h^{(0)}\nabla f(X^{(0)}))=$

$$=f\begin{pmatrix}2-h^{(0)}\cdot9\\2-h^{(0)}\cdot2\\3-h^{(0)}\cdot2\end{pmatrix}=(-9h^{(0)}+2)^3+(-2h^{(0)}+2)^2+(-2h^{(0)}+3)^2-3(-9h^{(0)}+2)-2(-2h^{(0)}+2)-4(-2h^{(0)}+3)+8.$$

$$z(h^{(0)})'=0\Rightarrow 3(-9h^{(0)}+2)^2\cdot(-9)+2(-2h^{(0)}+2)\cdot(-2)+2(-2h^{(0)}+3)\cdot(-2)+27+4+8=$$
$$=-27(81(h^{(0)})^2-36h^{(0)}+4)-4(-2h^{(0)}+2)-4(-2h^{(0)}+3)+39=0$$
$$-2187[h^{(0)}]^2+988h^{(0)}-89=0$$

$D=988^2-4\cdot(-2187)\cdot(-89)=\sqrt{444{,}491}$

$h^{(0)}_{1,2}=\dfrac{-988\pm444{,}491}{-2\cdot2187}$

$\underline{h^{(0)}_1=0{,}1243}$; $h^{(0)}_2=0{,}3275$.

Проверим условие минимума:

$z(h^{(0)})''=-4374h^{(0)}+988$.

$\underline{z(h^{(0)}_1)''=444{,}312}$; $z(h^{(0)}_2)''=-444{,}485$.

$$X^{(1)}=X^{(0)}-h^{(0)}\nabla f(X^{(0)})=(2;\ 2;\ 3)-0{,}1243\cdot(9;\ 2;\ 2)=(0{,}8813;\ 1{,}7514;\ 2{,}7514).$$
$\nabla f(X^{(1)})=(-0{,}6699;\ 1{,}5028;\ 1{,}5028)$; $\|\nabla f(X^{(1)})\|=2{,}2284$; $f(X^{(1)})=2{,}1698$.

2. Сравнительный анализ полученных результатов и вывод

Классический метод решения задачи наиболее точный и если есть возможность его использовать, то следует воспользоваться им. Остальные методы приближённого нахождения решения предназначены для поиска результата в сложных задачах.

Метод Ньютона на первой же итерации показал более близкий к истинному значению минимума результат. Однако сам по себе метод достаточно сложный и требует вычисления вторых производных исходной функции по всем её переменным, что не только увеличивает сложность алгоритма, но и значительно увеличивает время обработки каждой итерации.

Метод градиентного спуска на первой итерации показал самый дальний от истинного значения минимума результат. Также метод не предусматривает расчёта величины шага h, который влияет на скорость и возможность получения результата. В то же время данный метод вообще не требует нахождения каких-либо вторых производных, что значительно упрощает его алгоритм и ускоряет обработку каждой итерации.

Метод наискорейшего спуска на первой итерации показал более близкий результат, чем в предыдущем методе, но в то же время значительно более дальний, чем метод Ньютона. Данный метод лишён недостатка предыдущего метода, связанного с выбором шага итерации h, но из-за этого он требует дополнительных непростых расчётов и нахождения второй производной. Хотя, в общем, обработка каждой итерации всё равно остаётся проще, чем в методе Ньютона.

3. Результаты выполнения расчётов на ЭВМ

а) Метод Ньютона:

Входные данные:

```
f(x)=x1^3+x2^2+x3^2-3x1-2x2-4x3+8
h=0.2;eps=1.0E-4
X(0)=(2;2;3)
```

Результаты вычислений:

```
X(0) = ( 2.0000, 2.0000, 3.0000), f(X(0)) = 7.0000
k = 0           T(0) = (-0.7500,-1.0000,-1.0000)
X(1) = ( 1.2500, 1.0000, 2.0000), f(X(1)) = 1.2031
k = 1           T(1) = (-0.2250, 0.0000, 0.0000)
X(2) = ( 1.0250, 1.0000, 2.0000), f(X(2)) = 1.0019
k = 2           T(2) = (-0.0247, 0.0000, 0.0000)
X(3) = ( 1.0003, 1.0000, 2.0000), f(X(3)) = 1.0000
k = 3           T(3) = (-0.0003, 0.0000, 0.0000)
X(4) = ( 1.0000, 1.0000, 2.0000), f(X(4)) = 1.0000
k = 4           T(4) = ( 0.0000, 0.0000, 0.0000)
X(5) = ( 1.0000, 1.0000, 2.0000), f(X(5)) = 1.0000
```

Для нахождения точки минимума с заданной точностью ε методом Ньютона потребовалось 5 итераций.

б) Метод градиентного спуска:

Входные данные:

```
x1(3)+x2(2)+x3(2)-3x1-2x2-4x3+8
 2.0000000000E-01
 1.0000000000E-04
3
 2.0000000000E+00
 2.0000000000E+00
 3.0000000000E+00
```

Результаты вычислений:

```
k = 0,        X = ( 2.0000, 2.0000, 3.0000),  f(X) =  7.0000
Gr_f(X) = ( 9.0000, 2.0000, 2.0000), Norma_gr_f(X) = 9.43398
k = 1,        X = ( 0.2000, 1.6000, 2.6000),  f(X) =  3.1280
Gr_f(X) = (-2.8800, 1.2000, 1.2000), Norma_gr_f(X) = 3.34281
k = 2,        X = ( 0.7760, 1.3600, 2.3600),  f(X) =  1.3985
Gr_f(X) = (-1.1935, 0.7200, 0.7200), Norma_gr_f(X) = 1.56881
k = 3,        X = ( 1.0147, 1.2160, 2.2160),  f(X) =  1.0940
Gr_f(X) = ( 0.0888, 0.4320, 0.4320), Norma_gr_f(X) = 0.61736
k = 4,        X = ( 0.9969, 1.1296, 2.1296),  f(X) =  1.0336
Gr_f(X) = (-0.0184, 0.2592, 0.2592), Norma_gr_f(X) = 0.36702
k = 5,        X = ( 1.0006, 1.0778, 2.0778),  f(X) =  1.0121
Gr_f(X) = ( 0.0036, 0.1555, 0.1555), Norma_gr_f(X) = 0.21997
k = 6,        X = ( 0.9999, 1.0467, 2.0467),  f(X) =  1.0044
Gr_f(X) = (-0.0007, 0.0933, 0.0933), Norma_gr_f(X) = 0.13197
k = 7,        X = ( 1.0000, 1.0280, 2.0280),  f(X) =  1.0016
Gr_f(X) = ( 0.0001, 0.0560, 0.0560), Norma_gr_f(X) = 0.07918
k = 8,        X = ( 1.0000, 1.0168, 2.0168),  f(X) =  1.0006
Gr_f(X) = (-0.0000, 0.0336, 0.0336), Norma_gr_f(X) = 0.04751
k = 9,        X = ( 1.0000, 1.0101, 2.0101),  f(X) =  1.0002
Gr_f(X) = ( 0.0000, 0.0202, 0.0202), Norma_gr_f(X) = 0.02850
k = 10,        X = ( 1.0000, 1.0060, 2.0060),  f(X) =  1.0001
Gr_f(X) = (-0.0000, 0.0121, 0.0121), Norma_gr_f(X) = 0.01710
k = 11,        X = ( 1.0000, 1.0036, 2.0036),  f(X) =  1.0000
Gr_f(X) = ( 0.0000, 0.0073, 0.0073), Norma_gr_f(X) = 0.01026
```

```
k = 12,        X = ( 1.0000, 1.0022, 2.0022),  f(X) =  1.0000
Gr_f(X) = (-0.0000, 0.0044, 0.0044), Norma_gr_f(X) = 0.00616
k = 13,        X = ( 1.0000, 1.0013, 2.0013),  f(X) =  1.0000
Gr_f(X) = ( 0.0000, 0.0026, 0.0026), Norma_gr_f(X) = 0.00369
k = 14,        X = ( 1.0000, 1.0008, 2.0008),  f(X) =  1.0000
Gr_f(X) = (-0.0000, 0.0016, 0.0016), Norma_gr_f(X) = 0.00222
k = 15,        X = ( 1.0000, 1.0005, 2.0005),  f(X) =  1.0000
Gr_f(X) = ( 0.0000, 0.0009, 0.0009), Norma_gr_f(X) = 0.00133
k = 16,        X = ( 1.0000, 1.0003, 2.0003),  f(X) =  1.0000
Gr_f(X) = (-0.0000, 0.0006, 0.0006), Norma_gr_f(X) = 0.00080
k = 17,        X = ( 1.0000, 1.0002, 2.0002),  f(X) =  1.0000
Gr_f(X) = ( 0.0000, 0.0003, 0.0003), Norma_gr_f(X) = 0.00048
k = 18,        X = ( 1.0000, 1.0001, 2.0001),  f(X) =  1.0000
Gr_f(X) = (-0.0000, 0.0002, 0.0002), Norma_gr_f(X) = 0.00029
k = 19,        X = ( 1.0000, 1.0001, 2.0001),  f(X) =  1.0000
Gr_f(X) = ( 0.0000, 0.0001, 0.0001), Norma_gr_f(X) = 0.00017
k = 20,        X = ( 1.0000, 1.0000, 2.0000),  f(X) =  1.0000
Gr_f(X) = (-0.0000, 0.0001, 0.0001), Norma_gr_f(X) = 0.00010
k = 21,        X = ( 1.0000, 1.0000, 2.0000),  f(X) =  1.0000
Gr_f(X) = ( 0.0000, 0.0000, 0.0000), Norma_gr_f(X) = 0.00006
```

Для нахождения точки минимума с заданной точностью ε методом градиентного спуска потребовалось 21 итерация.

в) Методом наискорейшего спуска:

Входные данные:

```
x1(3)+x2(2)+x3(2)-3x1-2x2-4x3+8
 2.0000000000E-01
 1.0000000000E-04
3
 2.0000000000E+00
 2.0000000000E+00
 3.0000000000E+00
```

Результаты вычислений:

```
k = 0
X= ( 2.0000, 2.0000, 3.0000), Gr_f(X) = ( 9.0000, 2.0000, 2.0000)
Norma_Gr_f(X) = 9.43398, f(X) =  7.0000, h = 0.12426
k = 1
X= ( 0.8817, 1.7515, 2.7515), Gr_f(X) = (-0.6680, 1.5030, 1.5030)
Norma_Gr_f(X) = 2.22800, f(X) =  2.1698, h = 0.42164
k = 2
X= ( 1.1633, 1.1178, 2.1178), Gr_f(X) = ( 1.0599, 0.2355, 0.2355)
Norma_Gr_f(X) = 1.11104, f(X) =  1.1121, h = 0.16517
k = 3
X= ( 0.9882, 1.0789, 2.0789), Gr_f(X) = (-0.0701, 0.1577, 0.1577)
Norma_Gr_f(X) = 0.23382, f(X) =  1.0129, h = 0.42351
k = 4
X= ( 1.0179, 1.0121, 2.0121), Gr_f(X) = ( 0.1086, 0.0241, 0.0241)
Norma_Gr_f(X) = 0.11382, f(X) =  1.0013, h = 0.17586
k = 5
X= ( 0.9988, 1.0078, 2.0078), Gr_f(X) = (-0.0070, 0.0156, 0.0156)
Norma_Gr_f(X) = 0.02319, f(X) =  1.0001, h = 0.42378
k = 6
X= ( 1.0018, 1.0012, 2.0012), Gr_f(X) = ( 0.0107, 0.0024, 0.0024)
Norma_Gr_f(X) = 0.01125, f(X) =  1.0000, h = 0.17715
k = 7
X= ( 0.9999, 1.0008, 2.0008), Gr_f(X) = (-0.0007, 0.0015, 0.0015)
Norma_Gr_f(X) = 0.00228, f(X) =  1.0000, h = 0.42381
```

```
k = 8
X= ( 1.0002, 1.0001, 2.0001), Gr_f(X) = ( 0.0011, 0.0002, 0.0002)
Norma_Gr_f(X) = 0.00111, f(X) =  1.0000, h = 0.17728
k = 9
X= ( 1.0000, 1.0001, 2.0001), Gr_f(X) = (-0.0001, 0.0002, 0.0002)
Norma_Gr_f(X) = 0.00022, f(X) =  1.0000, h = 0.42381
k = 10
X= ( 1.0000, 1.0000, 2.0000), Gr_f(X) = ( 0.0001, 0.0000, 0.0000)
Norma_Gr_f(X) = 0.00011, f(X) =  1.0000, h = 0.17729
k = 11
X= ( 1.0000, 1.0000, 2.0000), Gr_f(X) = (-0.0000, 0.0000, 0.0000)
Norma_Gr_f(X) = 0.00002, f(X) =  1.0000, h = 0.42381
```

Для нахождения точки минимума с заданной точностью ε методом наискорейшего спуска потребовалось 11 итераций.

г) Методом сопряжённых градиентов:

Входные данные:

```
x1(3)+x2(2)+x3(2)-3x1-2x2-4x3+8
 2.0000000000E-01
 1.0000000000E-04
3
 2.0000000000E+00
 2.0000000000E+00
 3.0000000000E+00
```

Результаты вычислений:

```
k = 0
X= ( 2.0000, 2.0000, 3.0000), Gr_f(X) = ( 9.0000, 2.0000, 2.0000)
Norma_Gr_f(X) = 9.43398, f(X) =  7.0000, h = 0.12426
k = 1
X= ( 0.8817, 1.7515, 2.7515), Gr_f(X) = (-0.6680, 1.5030, 1.5030)
Norma_Gr_f(X) = 2.22800, f(X) =  2.1698, h = 0.46924
k = 2
X= ( 0.9596, 0.9939, 1.9939), Gr_f(X) = (-0.2377,-0.0122,-0.0122)
Norma_Gr_f(X) = 0.23833, f(X) =  1.0049, h = 0.16824
k = 3
X= ( 0.9999, 0.9928, 1.9928), Gr_f(X) = (-0.0007,-0.0143,-0.0143)
Norma_Gr_f(X) = 0.02027, f(X) =  1.0001, h = 0.48202
k = 4
X= ( 1.0011, 0.9997, 1.9997), Gr_f(X) = ( 0.0064,-0.0006,-0.0006)
Norma_Gr_f(X) = 0.00648, f(X) =  1.0000, h = 0.17109
k = 5
X= ( 1.0000, 1.0001, 2.0001), Gr_f(X) = ( 0.0001, 0.0001, 0.0001)
Norma_Gr_f(X) = 0.00021, f(X) =  1.0000, h = 0.35581
k = 6
X= ( 1.0000, 1.0000, 2.0000), Gr_f(X) = (-0.0001, 0.0000, 0.0000)
Norma_Gr_f(X) = 0.00012, f(X) =  1.0000, h = 0.23378
k = 7
X= ( 1.0000, 1.0000, 2.0000), Gr_f(X) = (-0.0000,-0.0000,-0.0000)
Norma_Gr_f(X) = 0.00000, f(X) =  1.0000, h = 0.20084
```

Для нахождения точки минимума с заданной точностью ε методом сопряжённых градиентов потребовалось 7 итераций.

4. Сравнительный анализ методов и выводы

В данной лабораторной работе невозможно зафиксировать, сколько времени тратит машина на одну итерацию в каждом конкретном методе, поэтому воспользуемся теоретическими рассуждениями из ранее сделанных выводов (см. п. 2, стр 5).

Метод Ньютона дал результат нужной точности за 5 итераций — это самое малое из всех методов число шагов. Но по предыдущим выводам теоретически алгоритм данного метода является самым сложным и самым длительным из всех.

Метод градиентного спуска достиг требуемой по условию точности за 21 итерацию — самое большое из всех методов число шагов. В то же время метод самый простой из всех и теоретически каждая итерация отрабатывается достаточно быстро.

Метод наискорейшего спуска получил результат за 11 итераций. Однако данный метод требует перерасчёта размера шага h на каждой итерации, что в свою очередь усложняет алгоритм и замедляет обработку одной итерации, но в то же время исключает расходимость и обеспечивает достаточную сходимость, в отличие от предыдущего метода.

Метод сопряжённых градиентов закончил вычисление минимума за 7 итераций. Данный метод подбирает шаг поиска в цикле и скорость нахождения шага h на каждой итерации может сильно варьироваться в зависимости от исходной функции.

Метод Ньютона слишком сложный и не факт, что самый быстрый, а метод градиентного спуска не гарантирует сходимость, поэтому самыми оптимальными методами будут метод наискорейшего спуска и метод сопряжённых градиентов. Какой из методов будет работать быстрее, сказать заранее невозможно, т.к. скорость алгоритма метода наискорейшего спуска приблизительно пропорциональна сложности функции, а метода сопряжённых градиентов — зависит ещё и от вида самой функции.

Предполагаю, что для большинства задач лучшим методом поиска будет метод сопряжённых градиентов.

Если под рукой нет ни одного готового метода, то для единичных расчётов на ЭВМ можно воспользоваться и методом градиентного спуска для более быстрой разработки. Однако нужно предусмотреть возможность ручного изменения шага, в случае расходимости или слишком длительного вычисления.