Министерство образования и науки Украины
Национальный технический университет
«Харьковский политехнический институт»

Кафедра компьютерной математики и анализа данных

ЛАБОРАТОРНАЯ РАБОТА №7

МЕТОД СОПРЯЖЕННЫХ ГРАДИЕНТОВ

ст. гр. КН-118 Тепляков А. Д.

Харьков, 2020

<u>Цель:</u> Реализовать метод сопряженных градиентов поиска экстремумов функции многих переменных.

<u>Задачи</u>

1. Изучить метод сопряженных градиентов.
2. Реализовать метод сопряженных градиентов поиска экстремума функции.
3. Исследовать скорость сходимости метода для квадратичной формы

$f(x) = (Ax,\ x) + (\mathbf{b}, x)$

где $\mathbf{x}, \mathbf{b} \in R^2,\ A \in R^{2\times 2}$ – положительно определена:

- с разной ориентацией осей:

  $\gamma = k \cdot \pi/4,\ (k = \overline{0,8}),$

  γ – угол между осью абсцисс и большою полуосью квадратичной формы;

- с разной эллиптичностью линий уровня:

  $\varepsilon = \dfrac{b}{a};\ \varepsilon = \{1,\ ...,\ 100\},$

  где b и a большая и малая полуоси эллипса соответственно;

4. Сравнить скорости сходимости метода наискорейшего спуска, метода Ньютона, метода ДФП и метода сопряженных градиентов для квадратичной формы большой размерности: $\dim x = \{2, ..., 1000\}$.

1. Теоретическая часть

Пусть имеется функция $f(x) = (Ax, x) + (b, x)$ (1), матрица A – положительно определена и симметрична. Требуется найти минимум данной функции.

В таком случае найти минимум можно при помощи метода сопряженных градиентов (далее – МСГ) по следующим формулам:

$$\begin{cases} x_{k+1} = x_k + \alpha_k d_k \\ k = 0 \ldots n-1, n = \dim(A) \\ d_0 = -\nabla f(x_0) \\ d_k = -\nabla f(x_k) + \beta_{k-1} d_{k-1} \\ \alpha_k = argminf(x_k) = -\frac{(Ax_k + b, d_k)}{2(Ad_k, d_k)} \\ \beta_{k-1} = \frac{(Ad_{k-1}, \nabla f_k)}{(Ad_{k-1}, d_{k-1})} \end{cases} \quad (2)$$

Где:

x_i – вектор в пространстве R^n

d_i – вектор из A – сопряженной системы векторов

$f(x)$ – функция вида (1)

$\alpha_i, \beta_i \in R^1$

Существуют теоремы, доказывающие, что МСГ является методом сопряженных направлений, и что для любой выбранной точки x_0 точка x_n, полученная по схеме (2), будет координатой $minf(x)$ и будет найдена не более чем за n шагов.

2. Результаты работы программы

Рассмотрим результаты работы программы для функции вида (1):

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, b = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

С различными точками x_0:

$$x_0 = (0;\ 0)$$

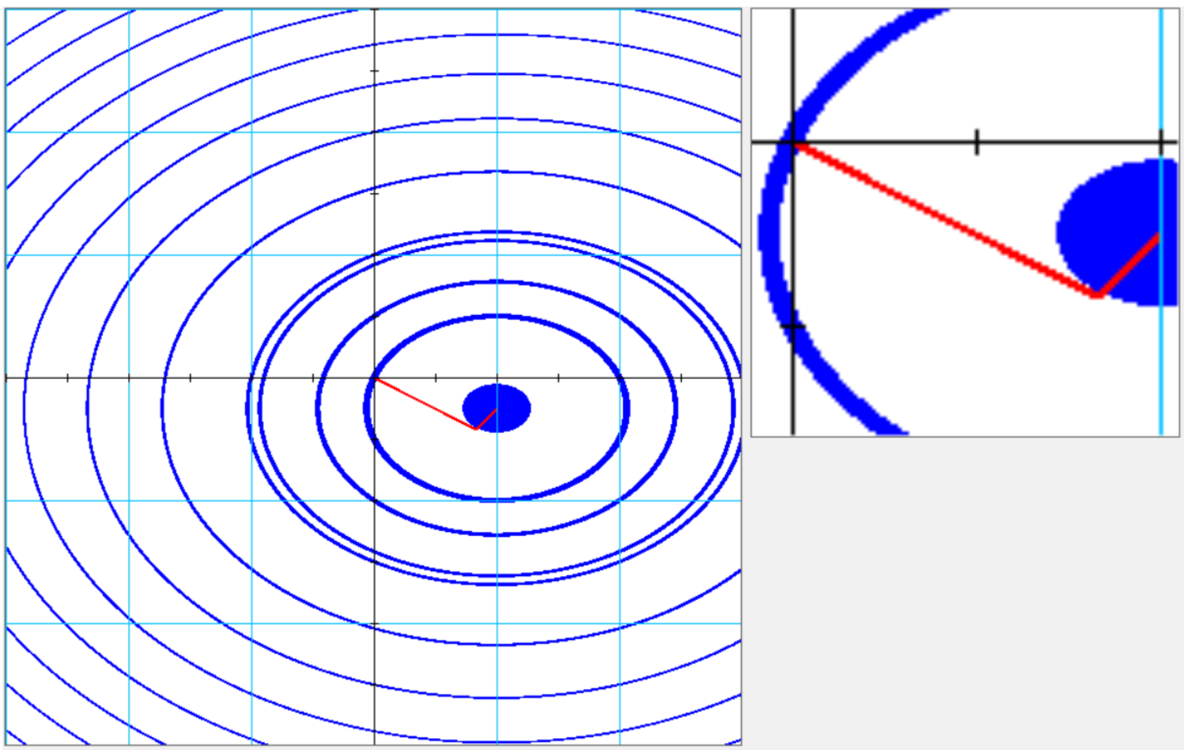

Рис.1 – результат работы МСГ для квадратичной формы

Iteration	Lambda	X1	X2	F_x	norm_X
	Forma				
0		0	0		
1	0,416661	0,833323	-0,41666	-1,04167	0,931683
2	0,3	0,999987	-0,25	-1,125	0,235699
3	0,413058	0,999998	-0,25	-1,125	1,20E-05
	Extremum	0,999998	-0,25	-1,125	

Рис. 2 – результат работы МСГ для квадратичной формы

$x_0 = (2;\ 2)$

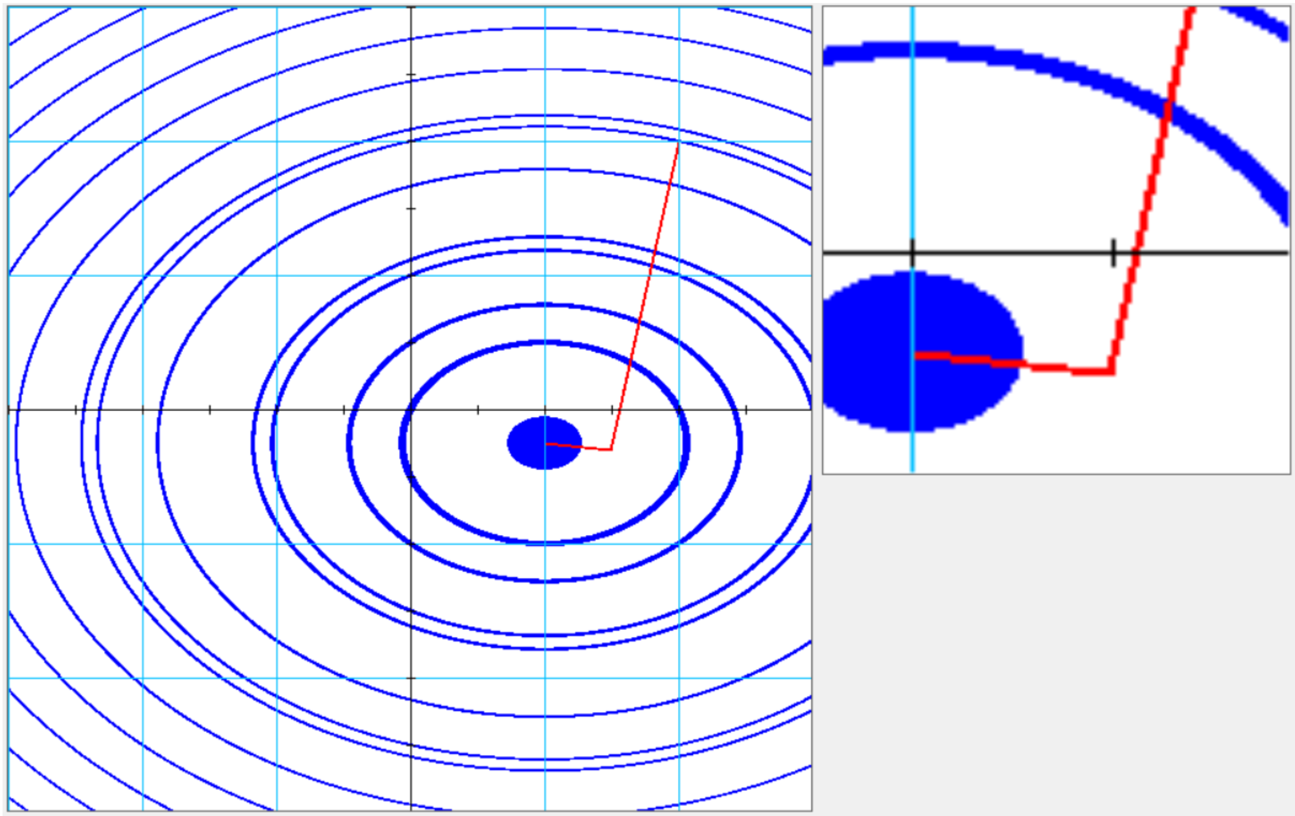

Рис.3 – МСГ для другой начальной точки

Iteration	Lambda	X1	X2	F_x	norm_X
	Forma				
0		2	2		
1	0,256008	1,487984	-0,30407	-0,88102	2,360276
2	0,488247	1,000051	-0,24986	-1,125	0,490936
3	0,254861	1,000019	-0,25	-1,125	0,00015
	Extremum	1,000019	-0,25	-1,125	

Рис.4 – МСГ для другой начальной точки

3. Исследование скорости сходимости МСГ для квадратичной формы

Исследуем сходимость МСГ для квадратичной формы в зависимости от угла наклона большей полуоси фигуры к оси абсцисс с помощью поворота фигуры матрицей поворота, где

$$\gamma = k\frac{\pi}{4}, (k = \overline{0,8}),$$

γ – угол наклона фигуры к оси абсцисс,

k – коэфициент угла наклона.

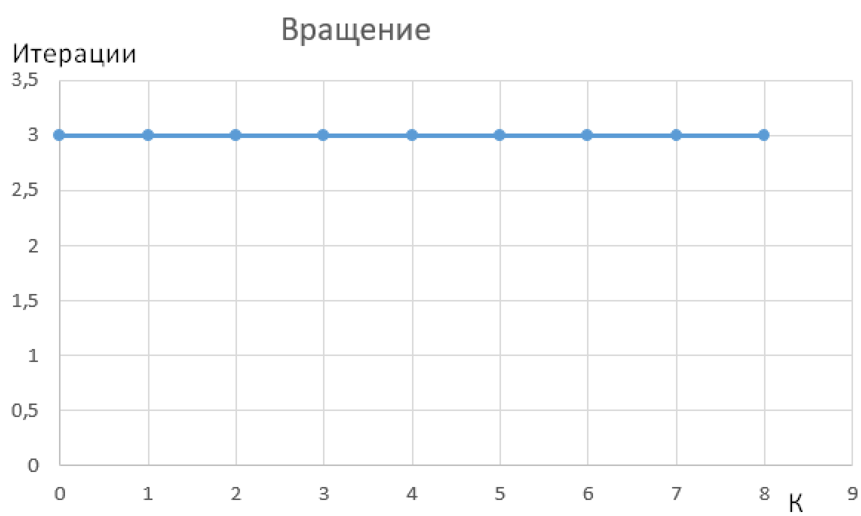

Рис.5 – вращение квадратичной формы

Исследуем сходимость МСГ для различной эллиптичности линий уровня квадратичной формы по формуле

$$\varepsilon = \frac{b}{a},$$

где b и a – большая и меньшая полуоси соответственно.

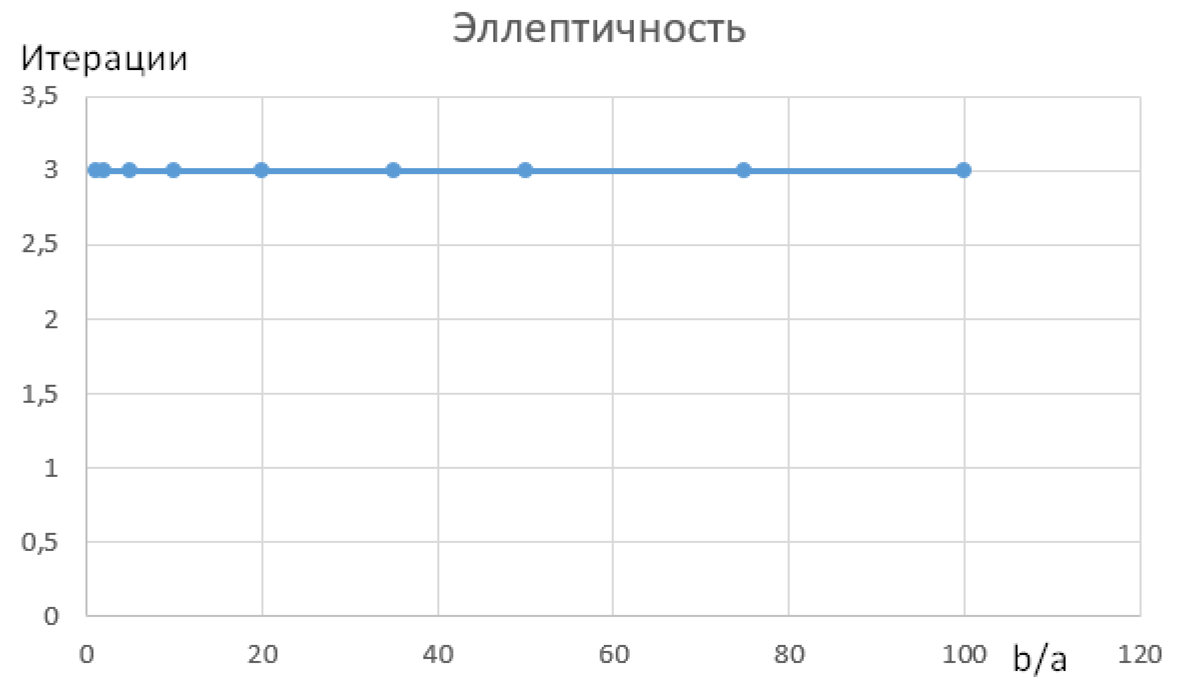

Рис.6 – эллиптичность квадратичной формы

Как можно видеть из приведенных выше графиков, количество итераций МСГ не изменяется с поворотом формы или с изменением эллиптичности линий уровня.

4. Сравнение МСГ с другими методами

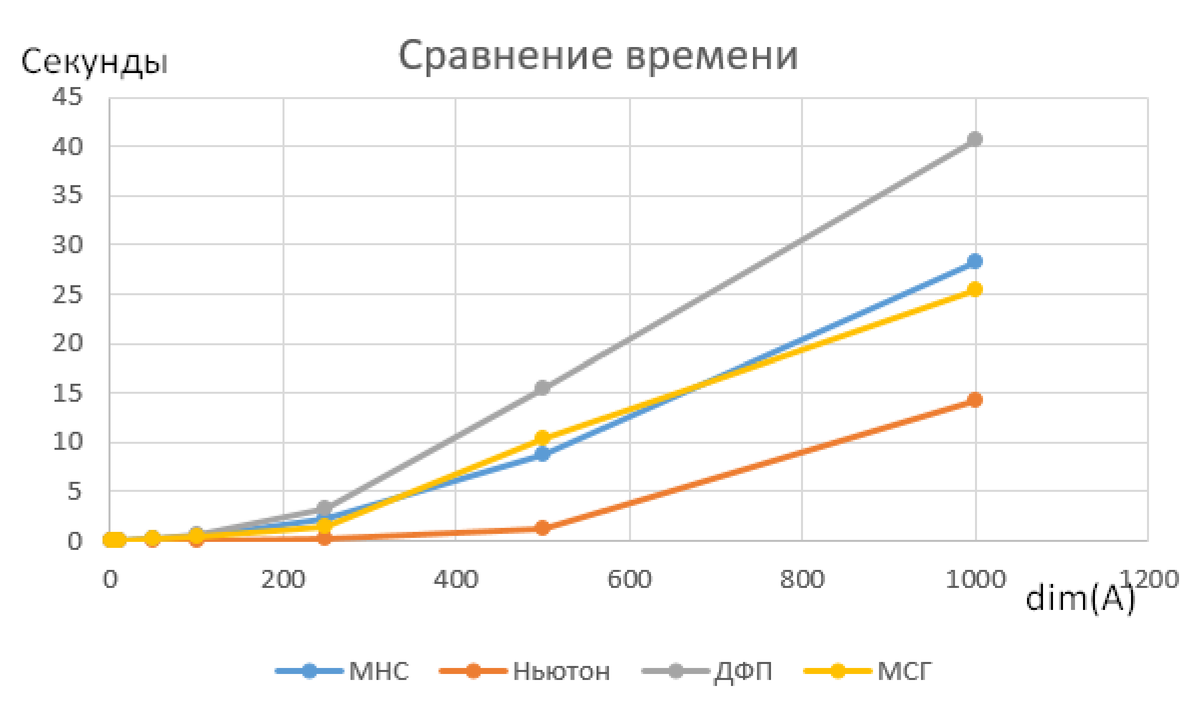

Рис.7 – сравнение времени работы методов

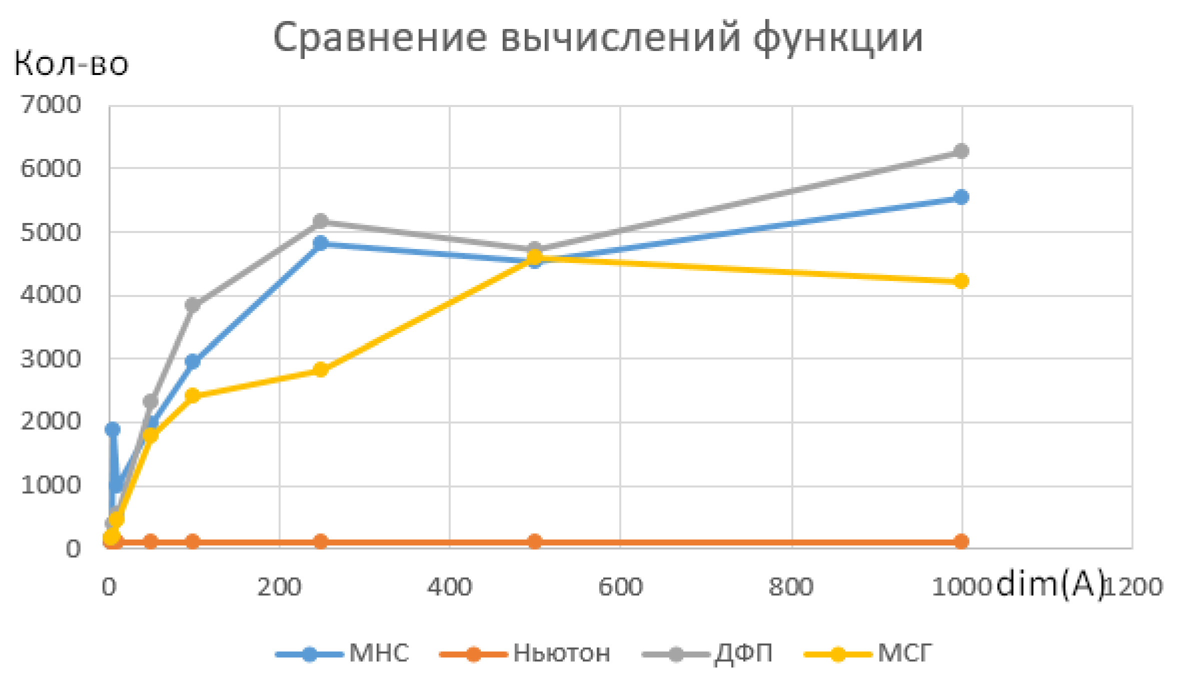

Рис.8 – сравнение количества вычислений функции

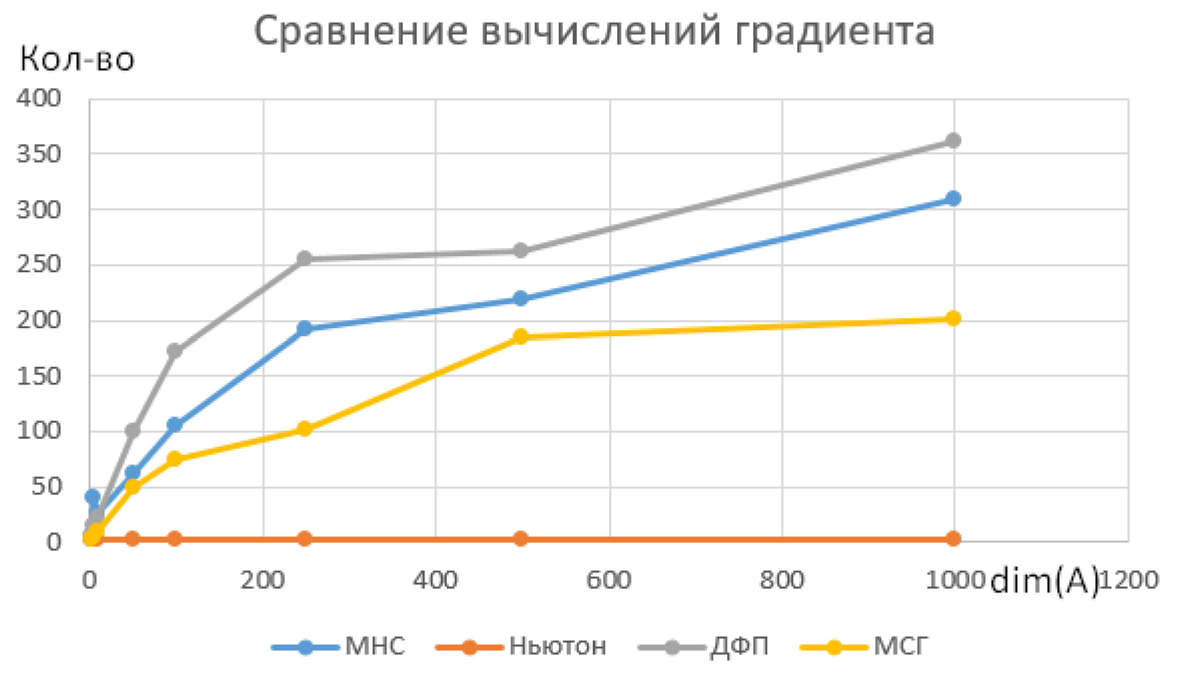

Рис.9 – сравнение количества вычислений градиента

Из приведенных выше графиков видно, что метод Ньютона наилучший – это объясняется низким количеством дополнительных вычислений, помимо обращения матрицы. МСГ оказывается на втором месте.

Так же следует учитывать, что полученные результаты могут несколько отличаться в зависимости от программной реализации методов.

5. Выводы

Основная идея методов сопряженных направлений – использование системы А – сопряженных векторов вместо направлений спуска. И, так как система таких векторов может быть выбрана различными способами, существует несколько методов сопряженной направлений.

В данной работе был рассмотрен и реализован один из методов сопряженных направлений – метод сопряженных градиентов. От ранее рассмотренных он отличается прежде всего тем, что не является методом спуска. Выбор системы А – сопряженных векторов делается при помощи градиента функции.

Так же было проведено сравнение работы МСГ для различных параметров квадратичной формы и выявлено, что те не влияют на его скорость сходимости.

Было проведено сравнение МСГ с методами спуска, в котором лучше него показал себя метод Ньютона, так как он требует меньше сопутствующих вычислений.