African Institute for Mathematical Sciences

Title of the course : Continuous Numerical Optimization

Lecturer: Justin Buhendwa Nyenyezi from Namur University, Belgium

---

Assignment n° 1

Group n° 3

Group members :

Member 1: Bolaji Rolayo BAMIRO
Member 2: Idowu Esther IJAODORO

Skills block n° 3

Report written by: Bolaji Rolayo BAMIRO

October 27, 2019

Topic: Analysis of the behavior of Steepest Descent, Preconditioned Steepest and Newton Algorithm method to understand the role of a preconditioner and the parameter α from Line Search algorithms of the given functions.

Summary of Work outline: Initial condition x_0 was fixed and different alphas were used for every preconditioners shown below. α= 1, 0.5 and 0.25 was used for both functions. The graphs of the gradiet norm against the number of iteration was generated using Octave. The results were presented on a table showing the various methods for comparing. The solutions and observations gotten from the results were shown in this work.

Note that: *The Armijo condition was used to determine Alpha and the Line search for all the methods.Newton was programmed to be set at $\alpha = 1$ in this work regardless of the change in alpha. That is the Newton Algorithm Method will be omitted for lower valued alpha(s). Maximum Iteration was set at 50 and tolerance at* $1e^{-6}$. The Preconditioners are;

Preconditioners steepest(PS1) 1: $D_k(i,i) = max(1, \frac{\delta^2 f(x_k)}{\delta x_i^2})^{-1}$

Preconditioners steepest 2: $D_k(i,i) = max(1, \frac{\delta^2 f(x_0)}{\delta x_i^2})^{-1}$

Preconditioners steepest 3: $D_k(i,i) = \frac{1}{max(1,|(x_k)_i|)}$

Preconditioners steepest 4: $D_k(i,i) = \frac{1}{max(1,|(x_0)_i|)}$

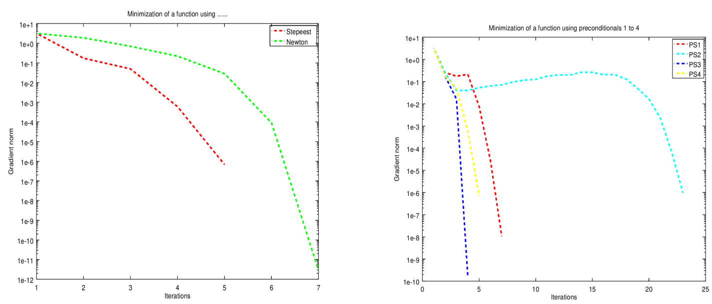
1. We optimize the function:

$$f(x) = \frac{1}{2}(x_1^2 + x_2^2)e^{(x_1^2 - x_2^2)}$$

$$g(x) = \begin{bmatrix} (x_1 + x_1^3 + x_1 x_2^2)e^{x_1^2 - x_2^2} \\ (x_2 - x_2 x_1^2 - x_2^3)e^{x_1^2 - x_2^2} \end{bmatrix}$$

$$H(x) = \begin{bmatrix} (1 + 5x_1^2 + 2x_1^4 + x_2^2 + 2x_1^2 x_2^2)e^{x_1^2 - x_2^2} & -2(x_1^3 x_2 - x_1 x_2^3)e^{x_1^2 - x_2^2} \\ -2(x_1^3 x_2 - x_1 x_2^3)e^{x_1^2 - x_2^2} & (1 - 5x_2^2 - x_1^2 + 2x_2^2 x_1^2 + 2x_2^4)e^{x_1^2 - x_2^2} \end{bmatrix}$$

   (a) Using Initial Condition $x_0 = [1, 1]$, We set $\alpha_{min} = 0$ and $\alpha_{max} = 2$. With one iteration, the α given by the armijo is 1.

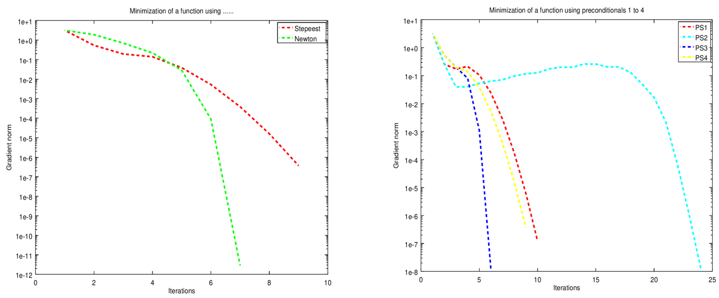


   Graph 1: showing the Gradient norm against number of iterations (**left:** shows the Newton(green) and steepest descent method(red), **right:** shows the Preconditionals Steepest (1:red, 2:cyan, 3:blue and 4:yellow) method)

   (b) Fixing the Initial Condition at $x_0 = [1, 1]$ constant, We set $\alpha_{min} = 0$ and $\alpha_{max} = 1$. With one iteration, the α given by the armijo condition is 0.5.

Descent method	Local Extrema(x_1, x_2)	Gradient Norm	No of iterations
Steepest	(-0.021536,4.304)	$6.798e^{-7}$	5
Newton	(0,0)	$2.875e^{-12}$	7
Preconditional 1	(-0.0336,4.8)	$1.05e^{-8}$	7
Preconditional 2	(1.99,4.7675)	$9.5127e^{-7}$	23
Preconditional 3	(-0.045,5.235)	$1.7409e^{-10}$	4
Preconditional 4	(-0.021536,4.304)	$6.798e^{-7}$	5

Table 1: Table to show results of the iteration of the various methods

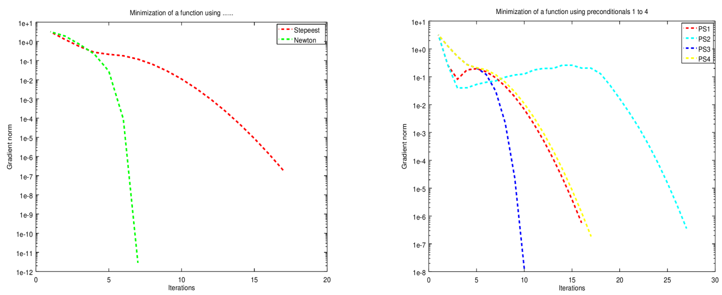


Graph 2: showing the Gradient norm against number of iterations

Descent method	Local Extrema(x_1, x_2)	Gradient Norm	No of iterations
Steepest	(-0.001829,4.38)	$3.7e^{-7}$	9
Preconditional 1	(-0.0083143,4.515)	$1.22e^{-7}$	10
Preconditional 2	(-0.22,4.78)	$1.311e^{-8}$	24
Preconditional 3	(0,4.78)	$1.1644e^{-8}$	6
Preconditional 4	(-0.0018,4.38)	$3.7e^{-7}$	9

Table 2: Table to show results of the iteration of the various methods

(c) Fixing the Initial Condition at $x_0 = [1, 1]$ constant, We set $\alpha_{min} = 0$ and $\alpha_{max} = 0.5$. With one iteration, the α given by the armijo is 0.25.



Graph 3: showing the Gradient norm against number of iterations

Descent method	Local Extrema(x_1, x_2)	Gradient Norm	No of iterations
Steepest	(0.0167,4.465)	$1.8522e^{-7}$	17
Preconditional 1	(0.0029,4.3253)	$5.744e^{-7}$	16
Preconditional 2	(0,4.384)	$3.58e^{-7}$	27
Preconditional 3	(0.055,4.778)	$1.2706e^{-8}$	10
Preconditional 4	(0.0167,4.45)	$1.8522e^{-7}$	17

Table 3: Table to show results of the iteration of the various methods

Obsevation and Conclusion: From **Table1 and Graph 1**, it was observed that the Newton Algorithm method gave a distinctive solution using seven iterations with quadratic convergence. The steepest method is quite similar to the **preconditional steepest 4** due to the initial condition x0 being (1,1) making it an identity matrix. **Preconditional Steepest 3** used the least number of iteration to acheive a solution which makes it a good preconditional for this function. However, **preconditional steepest 2** used the highest number of iteration to achieve its solution and is a bad precondition for this function. The local minimizers inferred from **Graph1 is (0,0) and (0, [4.3 - 5.24]); Graph 2 gives (0,[4.38 - 4.787]); Graph 3 gives (0,[4.32 - 4.77]).** PS: Since the Newton method works best with quadratic equations and this function is not, the Newwton algorithm method is unreliable for this function.

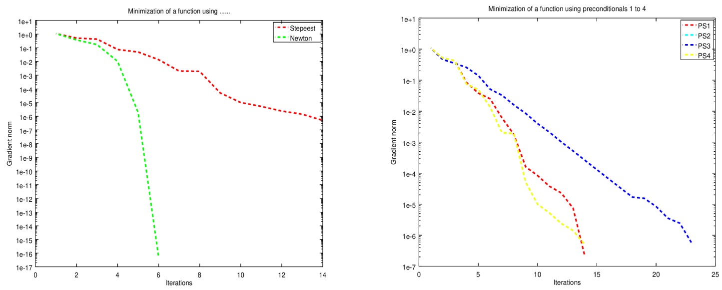
2. We optimize the function:

$$f(x) = \frac{1}{2}x_1^2 + x_1^2 cos(x_2)$$

$$g(x) = \begin{bmatrix} x_1 + cosx_2 \\ -xcosx_2 \end{bmatrix}$$

$$H(x) = \begin{bmatrix} 1 & -sinx_2 \\ -sinx_2 & -xcosx_2 \end{bmatrix}$$

(a) Using Initial Condition $x_0 = [1, 2]$, We set $\alpha_{min} = 0$ and $\alpha_{max} = 2$. With one iteration, the α given by the armijo is 1.

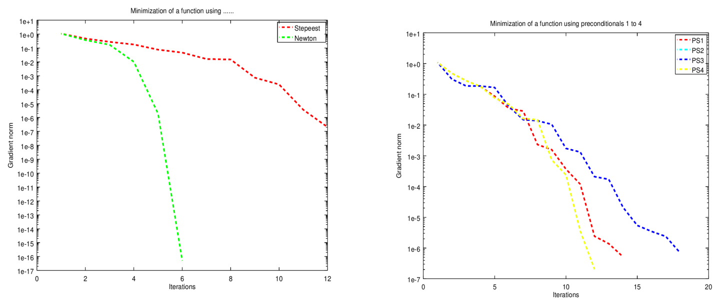


Graph 4: showing the Gradient norm against number of iterations

Descent method	Local Extrema(x_1, x_2)	Gradient Norm	No of iterations
Steepest	(1,3.14159)	$4,96e^{-7}$	14
Newton	(0,1.571)	$5.149e^{-17}$	6
Preconditional 1	(1,3.14159)	$2.02e^{-7}$	14
Preconditional 2	(1,3.14159)	$4.96e^{-7}$	14
Preconditional 3	(1,3.14159)	$5.62e^{-7}$	23
Preconditional 4	(1,3.14159)	$4.96e^{-7}$	14

Table 4: Table to show results of the iteration of the various methods

(b) Fixing the Initial Condition at $x_0 = [1, 2]$ constant, We set $\alpha_{min} = 0$ and $\alpha_{max} = 1$. With one iteration, the α given by the armijo is 0.5.

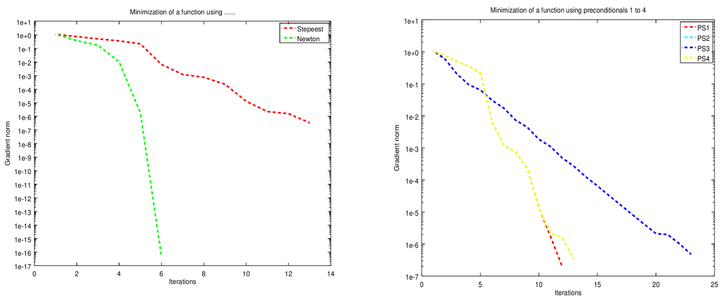


Graph 5: showing the Gradient norm against number of iterations

Descent method	Local Extrema(x_1, x_2)	Gradient Norm	No of iterations
Steepest	(1,3.1416)	$4.96e^{-7}$	12
Preconditional 1	(1,3.1416)	$5.195e^{-7}$	14
Preconditional 2	(1,3.1416)	$2.065e^{-7}$	12
Preconditional 3	(1,3.1416)	$7.058e^{-7}$	18
Preconditional 4	(1,3.1416)	$2.065e^{-7}$	12

Table 5: Table to show results of the iteration of the various methods

(c) Fixing the Initial Condition at $x_0 = [1, 1]$ constant, We set $\alpha_{min} = 0$ and $\alpha_{max} = 0.5$. With one iteration, the α given by the armijo is 0.25.



Graph 6: showing the Gradient norm against number of iterations

Descent method	Local Extrema(x_1, x_2)	Gradient Norm	No of iterations
Steepest	(1,3.14159)	$4,96e^{-7}$	13
Preconditional 1	(1,3.1416)	$1.92e^{-7}$	12
Preconditional 2	(1,3.1416)	$3.33e^{-7}$	13
Preconditional 3	(1,3.14159)	$4.8e^{-7}$	23
Preconditional 4	(1,3.14159)	$3.33e^{-7}$	13

Table 6: Table to show results of the iteration of the various methods

Obsevation and Conclusion: From **Table1 and Graph 1**, it was noticed that the Newton Algorithm method gave a distinctive solution using six iterations with quadratic convergence. The steepest method is also equivalent to the **preconditional steepest 4** for this function.All other methods the same global minimum with different alphas and within the range of 12 to 14 iterations except for **preconditional steepest 3** with a high number of iteration i.e 23. All methods have different speed and paths of convergence except for PS1 and PS2 that has the same which explains why PS2 does not show on the graph. The local minimizers inferred from all Graph(3-6) is (0,1.57) and (1,3.1416). PS: Since the Newton method works best with quadratic equations and this function is not, the newton is unreliable for this function.