School of Computer Science and Applied Mathematics

APPM 3017: Optimization III
Assignment 01

Lecturer: *Matthews Sejeso* ***Date:*** *14 September 2021*

INSTRUCTIONS

- The assignment is due on **Friday, October 08, 2021, 23:59:59**. Late submission will not be accepted under any circumstance.
- Work in groups of **no more than five** members. Only one member of the group has to submit the final work. Clearly indicate all the group members and their contributions.
- Your work should be typed and submitted in a single pdf file of seven pages maximum.
- For submission, upload your final work via Ulwazi. Do not submit anything via email.
- There are **40** marks allocated.

---

1 Overview

This assignment aims to experiment with several basic optimization methods discussed in the course content. The objective is to analyze and compare the performance of these methods on the given test functions, such as the strength and weakness of each method, the impact of parameters of line search or initial conditions on the optimal solution, etc. You will be evaluated on the depth of your analysis of the algorithm behaviour.

2 Test functions

The following objective functions represent different types of optimization test functions: quadratic, large-dimensional, etc. Test the performance of the methods on these functions.

1. $f(\mathbf{x}) = \sum_{i=1}^{n}(i \cdot x_i^2)$, where $n = 50$. The optimal solution is $\mathbf{x}^* = \mathbf{0}$.
2. $f(\mathbf{x}) = \sum_{i=1}^{n-1}\left[100(x_{i+1} - x_i^2)^2 + (1 - x_i)^2\right]$, where $n = 10$. The optimal solution is $\mathbf{x}^* = [1, \ldots, 1]^T$
3. $f(\mathbf{x}) = (1.5 - x_1 + x_1x_2)^2 + (2.25 - x_1 + x_1x_2^2)^2 + (2.625 - x_1 + x_1x_2^3)^2$. The optimal solution $\mathbf{x}^* = [3, 0.5]^T$

3 Task

Your task is to implement the following optimization algorithms with backtracking line search:

- Steepest descent method
- Newtons method
- Quasi-Newton method (DFP or BFGS)
- Conjugate gradient method (Polak-Ribiere or Fletcher-Reeves)

Analyze each method on each of the given test functions. In your analysis, relate the behaviour of the methods to the characteristics of the functions.

4 Grading

Write a report of seven pages maximum, including figures, single column. Please be clear and concise. In your report, be sure to address the following contents:

1. Provide a brief description of each of the methods. Also, provide the details of your implementations, for example, state the stopping criterion used, etc. [10 marks]

2. For all three test functions and all methods, create a table reporting the number of function calls and iterations required for convergence. You may include any other metric that can help analyse the methods. It may be helpful to test the impact of the initial condition by trying more than one starting point. Sometimes just using one point can be misleading. Discuss your results. [10 marks]

3. Provide the convergence plot for all three test functions. For each test function, *plot the results of all methods in the same set of axis.* A convergence plot should show iterations on the x-axis and a convergence metric with a log scale on the y-axis. You may use the convergence metric $|f(\mathbf{x})-f(\mathbf{x}^*)|$, or any other metric such as $\|\nabla f(\mathbf{x}^k)\|$ (since $f(\mathbf{x}^*)$ is usually not available.) [10 marks]

4. Discuss and compare the advantages and disadvantages of all the methods on the test functions and analyse the possible reasons. [10 marks]

Overall, you are expected to give insightful discussion and conclusion on the performance of the methods on the given test functions. Relate the performance to the nature of the functions. You are not required to submit the code; however, *I may request a meeting to discuss your code and results when marking.*