

Project 2 - Unconstrained Optimization

Instructor: Linwei Wang

1 Overview

In this project, you will experiment with the application of several basic optimization methods to different types of functions. You are given 3 objective functions for optimization. You are expected to implement and apply at least 3 types of basic optimization algorithms: 3 specified and discussed in class (100 points), and 1 optional and of your own choice (bonus of 25 points) on these functions, and analyze/compare the performances of these algorithms on the given functions, such as the strength and weakness of each method on each type of functions, the impact of the parameters in line search or the initial conditions on the solutions, etc.

You will be evaluated by 1) the accuracy of your optimization solutions; and 2) the depth of your analysis on the algorithm behaviors.

2 Problems

You are given the following test objective functions, each of which is representative of different types of optimization functions in terms of whether it is convex, quadratic, large-dimensional, etc.

1. $f(\mathbf{x}) = \sum_{i=1}^n (i \cdot x_i^2)$, where $n = 100$. Global minimum $f(\mathbf{x})_{\min} = 0$.
2. $f(\mathbf{x}) = \mathbf{c}^T \mathbf{x} - \sum_{i=1}^m \log(b_i - \mathbf{a}_i^T \mathbf{x})$ where $m = 500$ and $n = 100$. In matlab, this function and its first and second derivatives can be written as:
 - $f(\mathbf{x}) = \mathbf{c}'\mathbf{x} - \text{sum}(\log(\mathbf{b} - \mathbf{A}\mathbf{x}))$
 - $\nabla f(\mathbf{x}) = \mathbf{c} + \mathbf{A}'(1./(\mathbf{b} - \mathbf{A}\mathbf{x}))$
 - $\mathbf{H}(\mathbf{x}) = \mathbf{A}' \text{diag}(1./((\mathbf{b} - \mathbf{A}\mathbf{x}).^2))\mathbf{A}$ where *diag* means to form a diagonal matrix with the main diagonal from the vector $1./((\mathbf{b} - \mathbf{A}\mathbf{x}).^2)$

$$3. f(\mathbf{x}) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

Except for function 1, we will keep the global optima until after the project. You will be in part evaluated by the accuracy of your solutions compared to the *true* global optima. The global optimum of function 1 is given so that, for this particular function, you can observe and show the convergence curve of the 3 or 4 algorithms (solutions *vs.* true minimum at each iteration). The coefficients in functions 2 are provided in a separate text file. You can read and use the file according to the accompanied "ReadMe.txt" file

3 Task

You will implement the following 3 basic optimization algorithms

- Gradient descent method
- Newton method
- Quasi-Newton method

and for a bonus of 25 points, to implement an additional algorithm

- Method of your choice, such as the Simplex method or the other methods recommended in the further reading.

and apply each of them to the provided 3 test functions.

You will give the solutions of each task (specified in 4 - Deliverables). You will also give analysis of each method on each of the provided functions. As explained earlier, the provided functions are representative in terms of whether they are quadratic, convex, etc. In your analysis, you are expected to relate the behaviors of the experimented methods to these characteristics of the functions (In other words, we are not expecting surface-level analysis such as, for function 3, this method is better than that method, etc). The important characteristics will be discussed after the project.

4 Deliverables

1. Code. Please submit all the code, along with a brief documentation. We should be able to run your code based on these.

2. Report. Three page maximum including figures, single column. Please be clear and concise. You are expected to include (but not restricted to) the following contents:
 - For each test function,
 - (a) Provide the optimization solution obtained from each method
 - (b) Provide the convergence figure (function value at each iteration versus the given global minimum) for function 1.
 - (c) Discuss and compare the advantages and disadvantages of the 4 methods on this type of function, and analyze the possible reasons.
 - (d) You may test the impact of the initial conditions and the parameters of the algorithm, such as the parameters of backtracking linesearch, on the solutions of each method.
 - Overall, you are expected to give insightful discussions and conclusions on the performance of the experimented methods, with regard to the type of functions they are being applied to. You are expected to not only stay on the surface of the expression of each given function, but to relate to the nature/type of the functions. (Hint: for example, whether the given functions are convex, or quadratic, etc).
3. Presentation. You will give a 8 minute presentation as a group. You don't have to describe in detail the first 3 methods discussed in class. If you chose to test another method of you choice for bonus point, please describe the method. Describe what solutions you get, are they expected or not, what might be the reasons if certain methods didn't succeed in certain functions, what you discovered or learned, etc.

5 Grading

- 100 points on the 3 methods discussed in classes
 - Content: 50%
 - Report: 20%
 - Presentation: 30%
- 25 bonus point on one method of your choice.