FEDERAL STATE AUTONOMOUS EDUCATIONAL INSTITUTION
OF HIGHER EDUCATION
ITMO UNIVERSITY

Report

on the practical task No. 3

Algorithms for unconstrained nonlinear optimization. First- and secondorder methods

Performed by
Carlos Andres Daza Rachen

Academic group
Accepted by
Dr Petr Chunaev

St. Petersburg
2022

Goal

The use of first- and second-order methods (Gradient Descent, Non-linear Conjugate gradient Descent, Newton's method and Levenberg-Marquardt algorithm) in the tasks of unconstrained nonlinear optimization

Formulation of the problem

Compare first and second order methods

Brief theoretical part [1]

First-order methods rely on gradient information to help direct the search for a minimum:

- Gradcient descent: Use the direction of steepest descent.Following the direction of steepest descent is guaranteed to lead to improvement, provided that the objective function is smooth, the step size is sufficiently small, and we are not already at a point where the gradient is zero. The direction of steepest descent is the direction opposite the gradient ∇f, hence the name gradient descent.

- Conjugate gradient: Gradient descent can perform poorly in narrow valleys. The conjugate gradient method overcomes this issue by borrowing inspiration from methods for optimizing quadratic functions:

$$\underset{\mathbf{x}}{\text{minimize}}\, f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{A}\mathbf{x} + \mathbf{b}^\top \mathbf{x} + c$$

  where A is symmetric and positive definite, and thus f has a unique local minimum

Leveraging second-order approximations use the second derivative in univariate optimization or the Hessian in multivariate optimization to direct the search:

- Newton's Method: Second-order information, allows us to make a quadratic approximation of the objective function and approximate the right step size to reach a local minimum as shown. we can analytically obtain the location where a quadratic approximation has a zero gradient. We can then use that location as the next iteration to approach a local minimum.
  The update rule in Newton's method involves dividing by the second derivative. The update is undefined if the second derivative is zero, which occurs when the quadratic approximation is a horizontal line.

- Levenberg-Marquardt algorithm: was developed in the early 1960's to solve nonlinear least squares problems. Least squares problems arise in the context of fitting a parameterized mathematical model to a set of data points by minimizing an objective expressed as the sum of the squares of the errors between the model function and a set of data points.

Results

1. Linear function

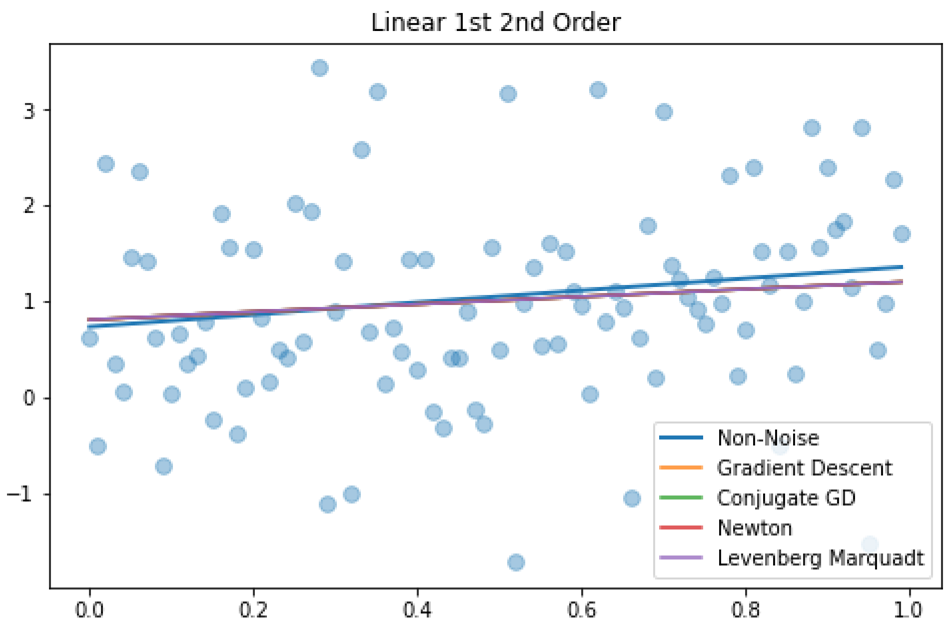
	Iterations	Calls	a	b	value
Gradient descent	41	41	0.395249	0.812475	105.379
Conjugate GD	9	45	0.392559	0.814389	86.3369
Newton	3	15	0.392530	0.814405	86.3369
Levenberg	6	6	0.392530	0.814405	43.1684
Exhaustive Search	119	119	0.682614	0.627814	104.511
Gauss Search	3	101	0.682587	0.627814	104.511
Nelder-Mead	20	38	0.682244	0.628022	104.511

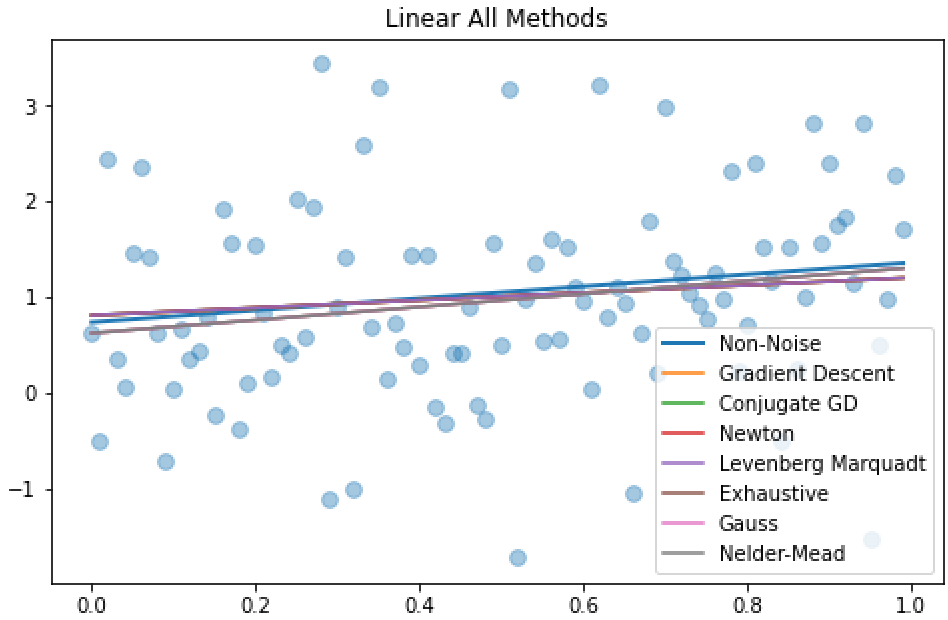
2. Rational function

	Iterations	Calls	a	b	value
Gradient descent	170	170	0.694740	-0.510737	104.159
Conjugate GD	10	84	0.681259	-0.527889	104.149
Newton	10	57	0.681260	-0.527889	104.149
Levenberg	20	20	0.681259	-0.52789	52.0745
Exhaustive Search	119	119	0.681247	-0.527907	104.149
Gauss Search	4	98	0.692410	-0.523534	104.17
Nelder-Mead	35	68	0.681014	-0.528354	104.149

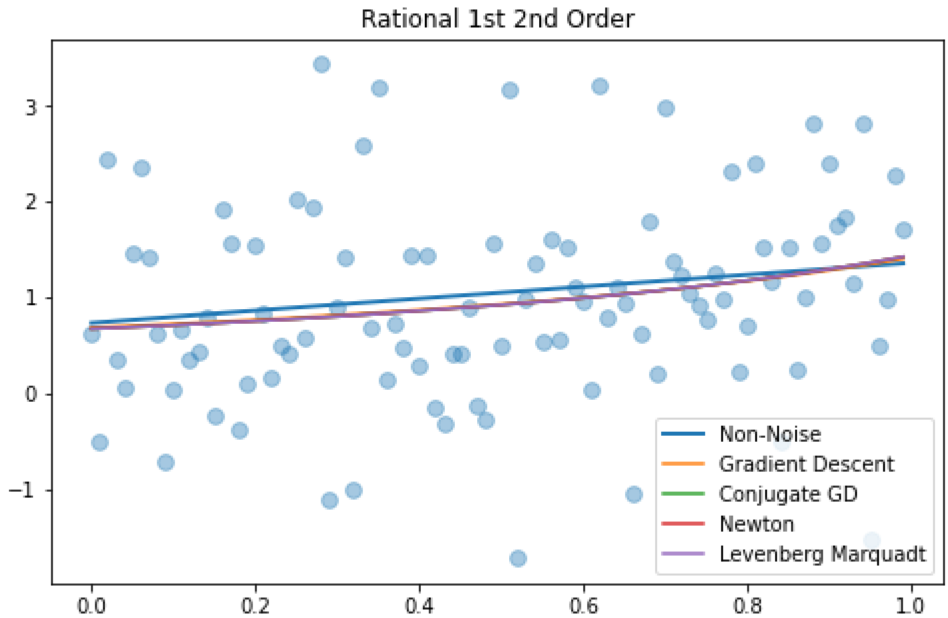
3. Plot

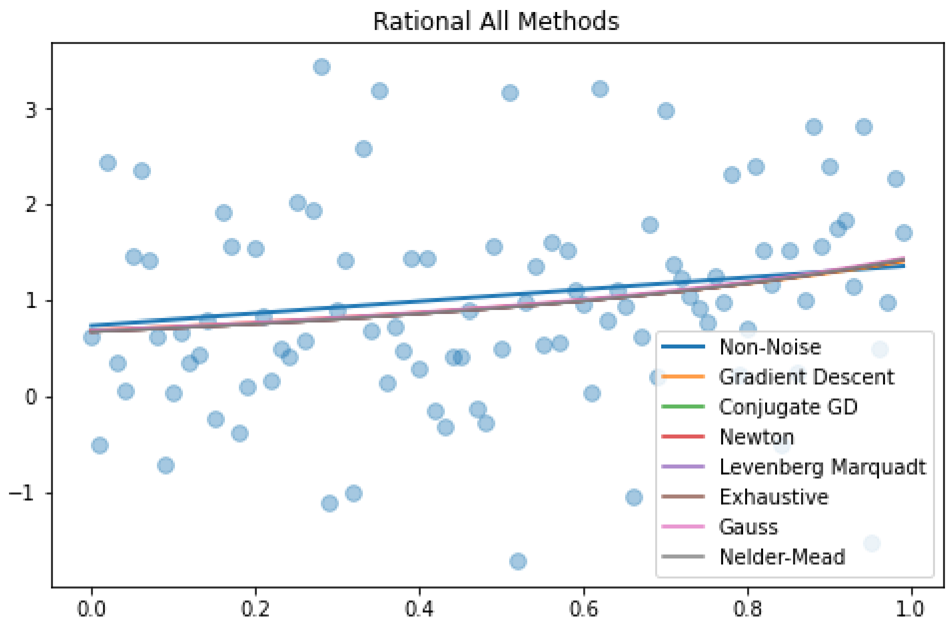
3.1 Linear




Linear All Methods
3
2
1
0
−1
0.0
0.2
0.4
0.6
0.8
1.0
Non-Noise
Gradient Descent
Conjugate GD
Newton
Levenberg Marquadt
Exhaustive
Gauss
Nelder-Mead

3.2 Rational


Rational 1st 2nd Order
3
2
1
0
−1
0.0
0.2
0.4
0.6
0.8
1.0
Non-Noise
Gradient Descent
Conjugate GD
Newton
Levenberg Marquadt



Conclusions

- Gauss search is the speedest in accordance to the less number of iterations that it needs and also have decent results.
- Gradient descent for rational function even took more time than exahustive, it's the most demanding in number of iterations.
- The linear approximations had a result closer to the true dependence. The methods of task 2 have a similar behavior and those of task 3 also mark the same graph.

Bibliography

[1] Kochenderfer, Algorithms for optimization
[2] Henri P. Gavin, The Levenberg-Marquardt algorithm for nonlinear least squares curve-fitting problems, 2020