FEDERAL STATE AUTONOMOUS EDUCATIONAL INSTITUTION

OF HIGHER EDUCATION

ITMO UNIVERSITY

Report

on the practical task No. 3

“Task 3. Algorithms for unconstrained nonlinear optimization. First- and second-order methods”

Performed by

Alexandra Matveeva

J4134c

Accepted by

Dr Petr Chunaev

St. Petersburg

2021

**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.

**Problems**

Generate random numbers and . Furthermore, generate the noisy data , where , according to the following rule:

,

where are values of a random variable with standard normal distribution. Approximate the data by the following linear and rational functions:

1. (linear approximant)
2. (rational approximation)

by means of least squares through the numerical minimization (with precision ) of the following function:

To solve the minimization problem, use the methods of Gradient Descent, Conjugate Gradient Descent, Newton’s method and Levenberg-Marquardt algorithm. If necessary, set the initial approximations and other parameters of the methods. Visualize the data and the approximants obtained in a plot separately for each type of **approximant** so that one can compare the results for the numerical methods used. Analyze the results obtained (in terms of number of iterations, precision, number of function evaluations, etc.) and compare them with those from Task 2 for the same dataset.

**Brief theoretical part**

**Gradient Descent**

The **gradient** of a differentiable function at **a** is the *n*-dimensional **column**-vector whose elements are

Gradient descent method is based on the observation that if *f* is differentiable at **a**, then *f*(**x**) decreases **fastest** in a neighbourhood of **a** in the direction of . One may write down the following formula:

Starting with some initial approximation .

If step size is chosen properly, then and furthermore as , where is a local minimum.

If *f* is convex, is Lipschitz and the choice of is due to Barzilai-Borwein,

,

then the uniform convergence is guaranteed.

**(Nonlinear) Conjugate Gradient Descent**

The conjugate gradient method is an iterative method for unconditional optimization in a multidimensional space. The main advantage of the method is that it solves a quadratic optimization problem in a finite number of steps.

is a differentiable function, is an initial approximation, one start in the steepest descent direction:

.

Step size: and . After this iteration, the following steps with constitute one iteration of moving along a subsequent conjugate direction , where :

- Calculate the steepest direction

- Compute according to certain formulas

- Update the conjugate direction

- Find

- Update the position: **.**

The choice of (to guarantee the uniform convergence as for convex *f* and Lipschitz ) is due to Fletcher-Reeves or Polak-Ribiere:

**Newton’s method**

is a convex and twice differentiable. We should find a root of constructing a sequence from an initial approximation so that as , where

From the Taylor expansion of *f* near ,

.

We use this quadratic function (with respect to ) as an approximant to *f* in a neighbourhood of . The vertex of the corresponding parabola gives us the next point . To find the vertex x-coordinate, we write:

. Incrementing by this gives us a point closer :

It is proved that for the chosen class of *f* one has as .

**Levenberg-Marquardt algorithm**

The Levenberg-Marquardt algorithm combines two numerical minimization algorithms: the gradient descent method and the Gauss-Newton method. Levenberg-Marquardt is a popular alternative to the Gauss-Newton method of finding the minimum of a function that is a sum of squares of nonlinear functions,

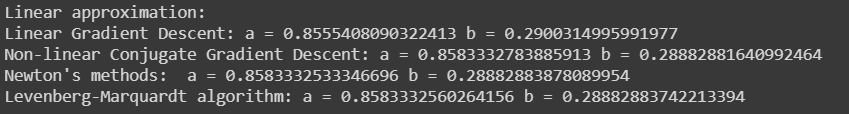
Let the Jacobian of be denoted , then the Levenberg-Marquardt method searches in the direction given by the solution *p* to the equations

,

where are nonnegative scalars and *I* is the identity matrix. The method has the nice property that, for some scalar related to , the vector is the solution of the constrained subproblem of minimizing subject to .

**Results**

1. Linear approximantion

****

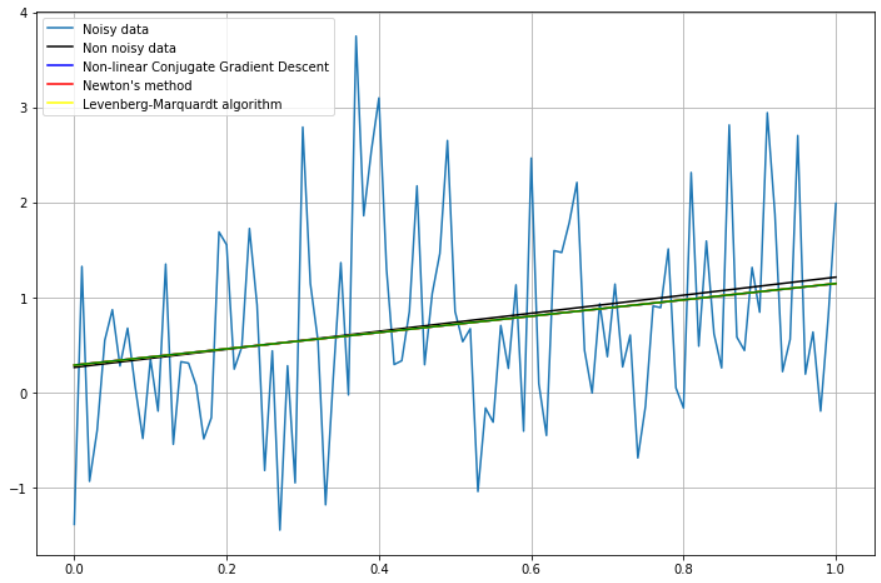
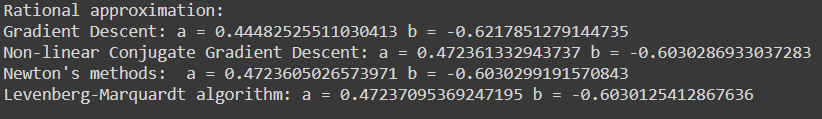


Fig. 1 - The results of the application of Gradient descent, Conjugate gradient method, Newton’s method and the Levenberg-Marquardt algorithm for solving the linear approximation problem

2. Rational approximation



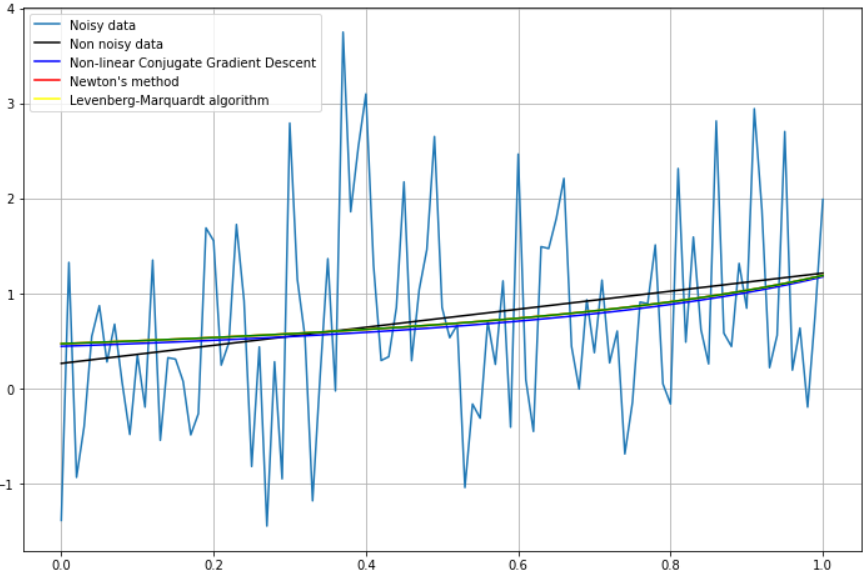


Fig. 2 - The results of the application of Gradient descent, Conjugate gradient method, Newton’s method and the Levenberg-Marquardt algorithm for solving the rational approximation problem

It is well known that the optimization problem associated with linear approximation has a single solution, and therefore it is expected that these methods will give similar optimal values for a and b, regardless of the choice of initial approximations. In the case of rational approximation, significant nonlinearities arise, and therefore the choice of the initial approximation can significantly affect the result.

**Conclusion**

During the execution of practical task, the approximate solutions 𝑥: was found. To complete this task the first- (Gradient descent, Conjugate Gradient) and second-order (Newton’s method, Levenberg-Marquardt algorithm) methods were used. We obtained the same results for each method. A detailed analysis can be found in the results section.

**Appendix**

GitHub Link: <https://github.com/alex-mat-s/Algorithms/blob/main/Lab3.ipynb>