Linear function:

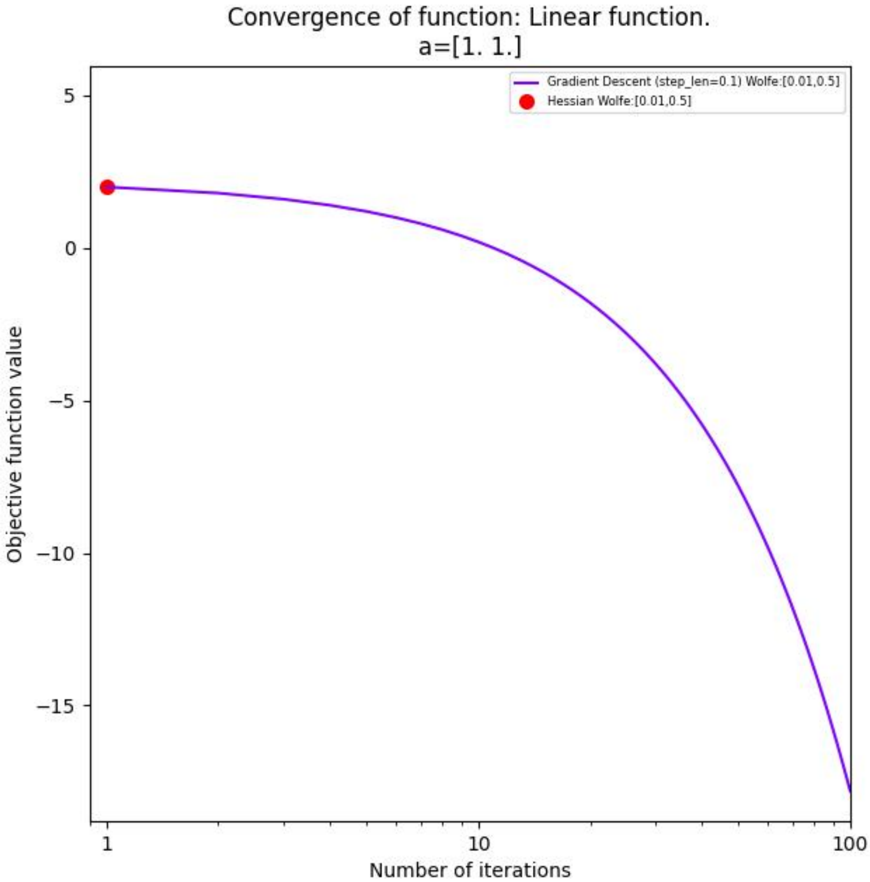
Gradient Descent

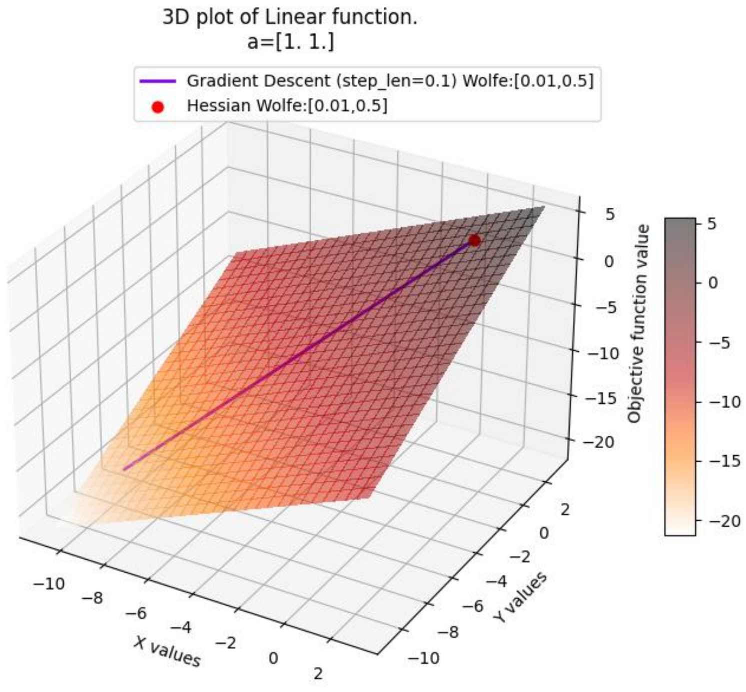
Wolfe C1=0.01 step_len=0.1

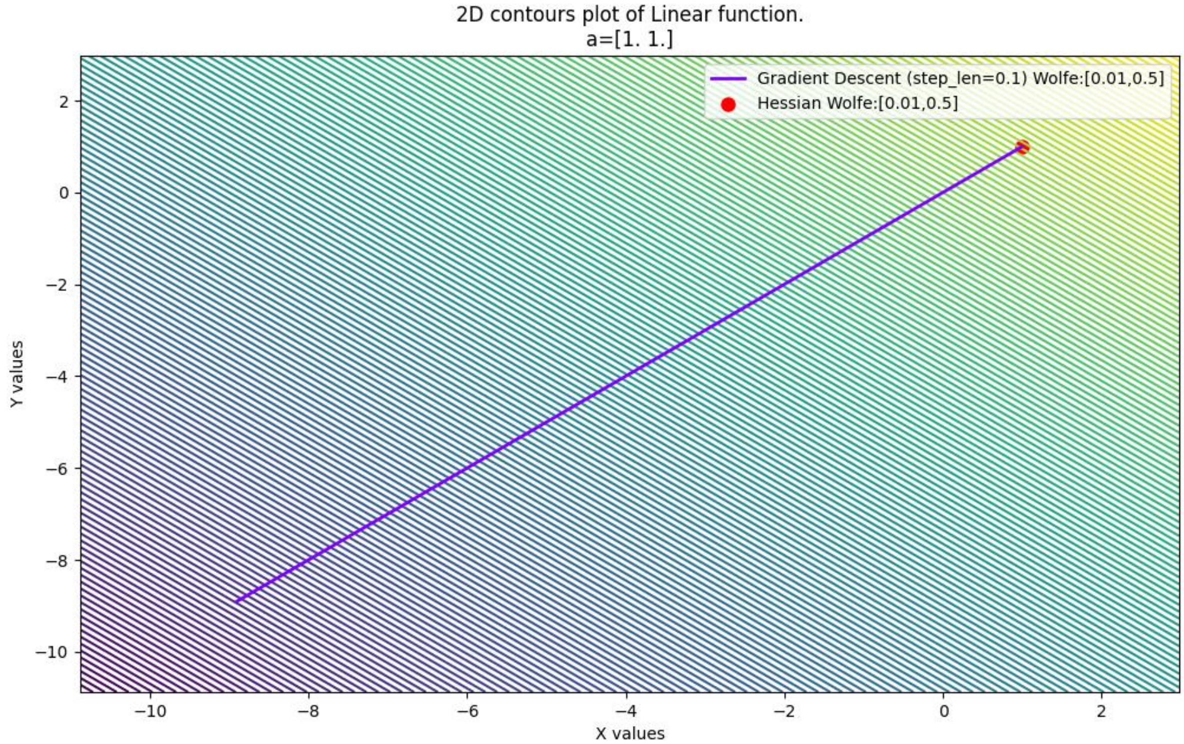
{'location': array([-8.9, -8.9]), 'objective': -17.79999999999997, 'success': False, 'num_iter': 100}

Hessian

Wolfe C1=0.01 {'location': array([1., 1.]), 'objective': 2.0, 'success': False, 'num_iter': 1}







Quad function 1:

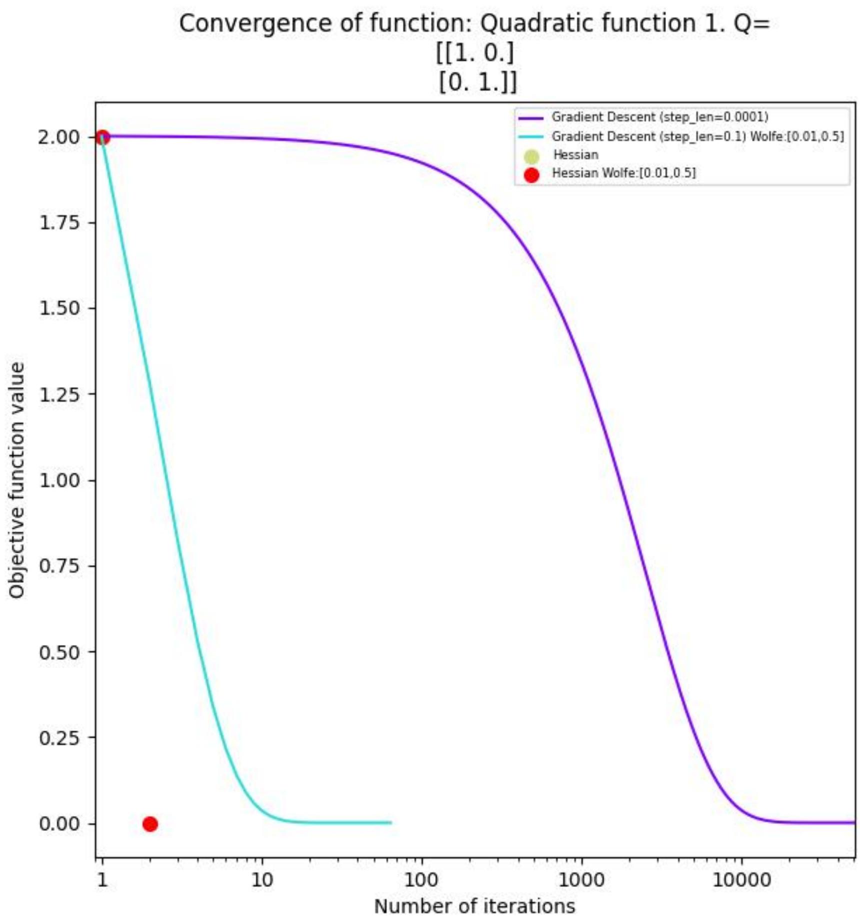
Gradient Descent

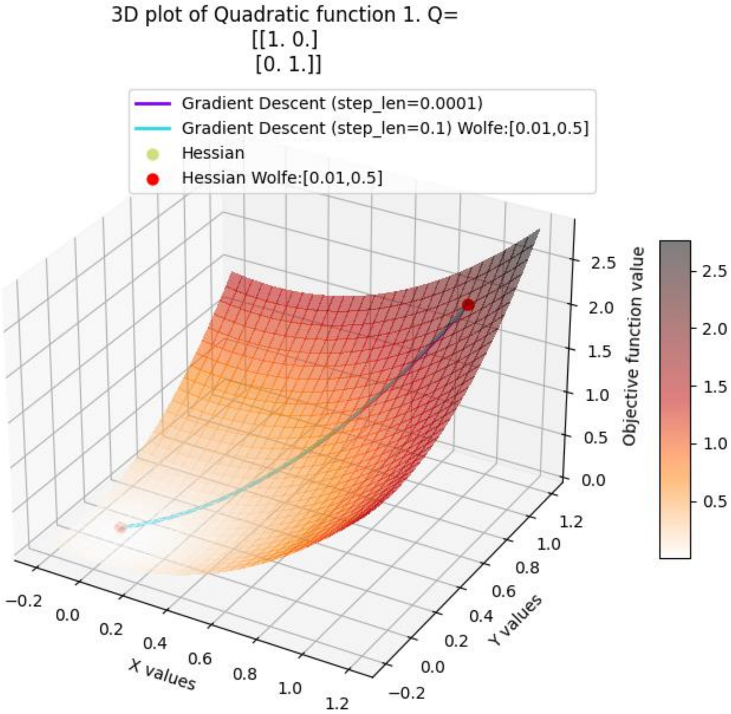
Wolfe=0 step_len=1e-4 {'location': array([3.53495443e-05, 3.53495443e-05]), 'objective': 2.4991805650220l7e-09, 'success': True, 'num_iter': 51247}

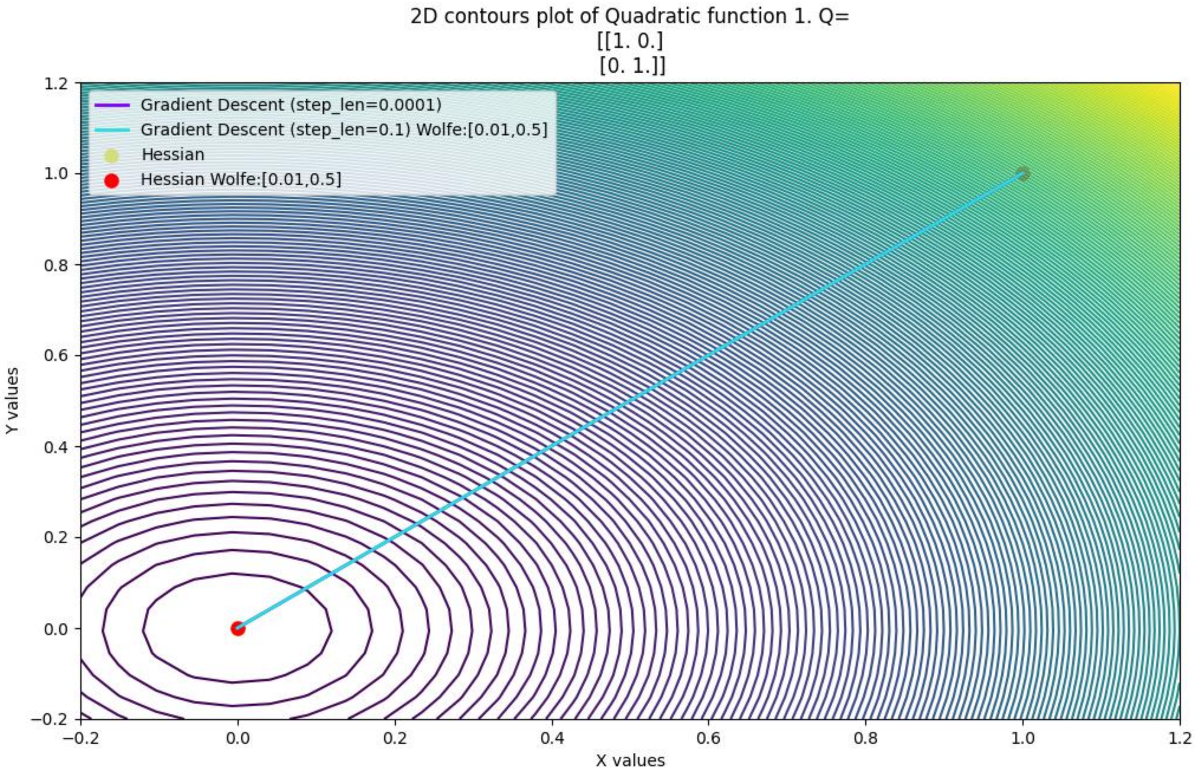
Wolfe=0.01 step_len=0.1 {'location': array([7.84637717e-07, 7.84637717e-07]), 'objective': 1.2313126936373286e-12, 'success': True, 'num_iter': 64}

Hessian

Wolfe=0 or Wolfe=0.01 {'location': array([0., 0.]), 'objective': 0.0, 'success': True, 'num_iter': 2}







Quad function 2:

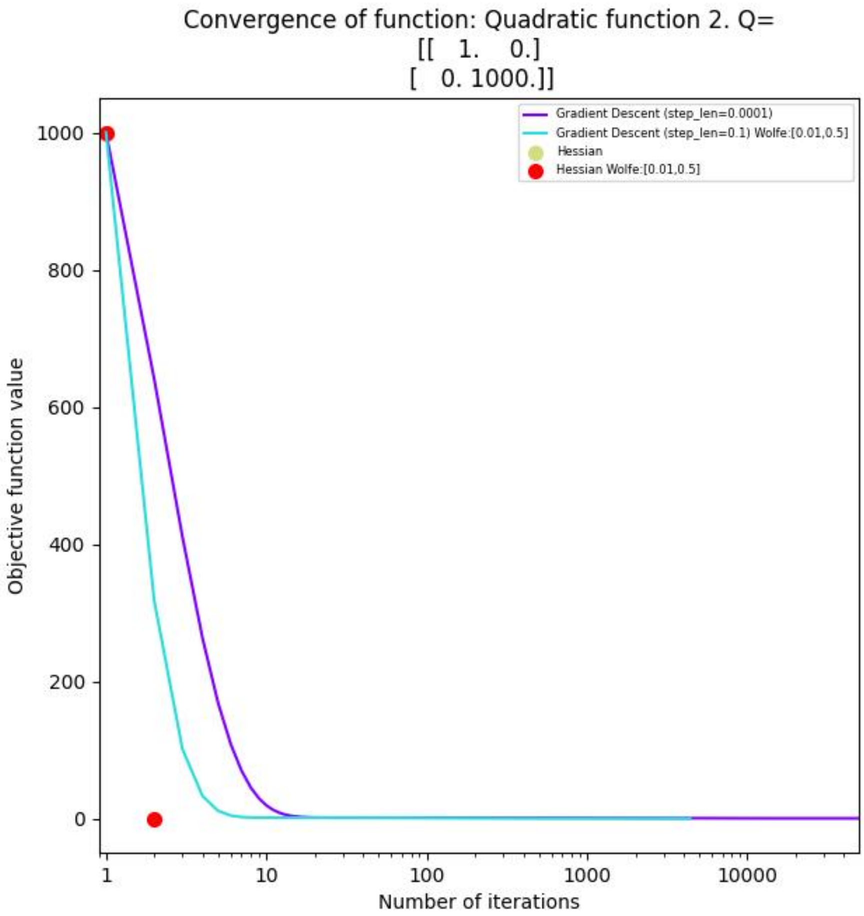
Gradient Descent

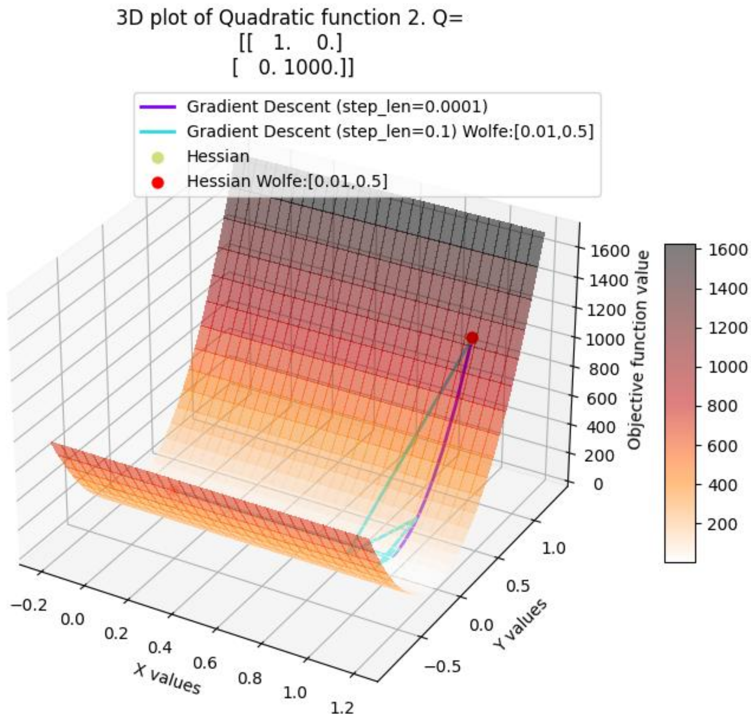
Wolfe=0 step_len=1e-4 {'location': array([4.99948583e-005, 9.88131292e-324]), 'objective': 2.4994858552844077e-09, 'success': True, 'num_iter': 49514}

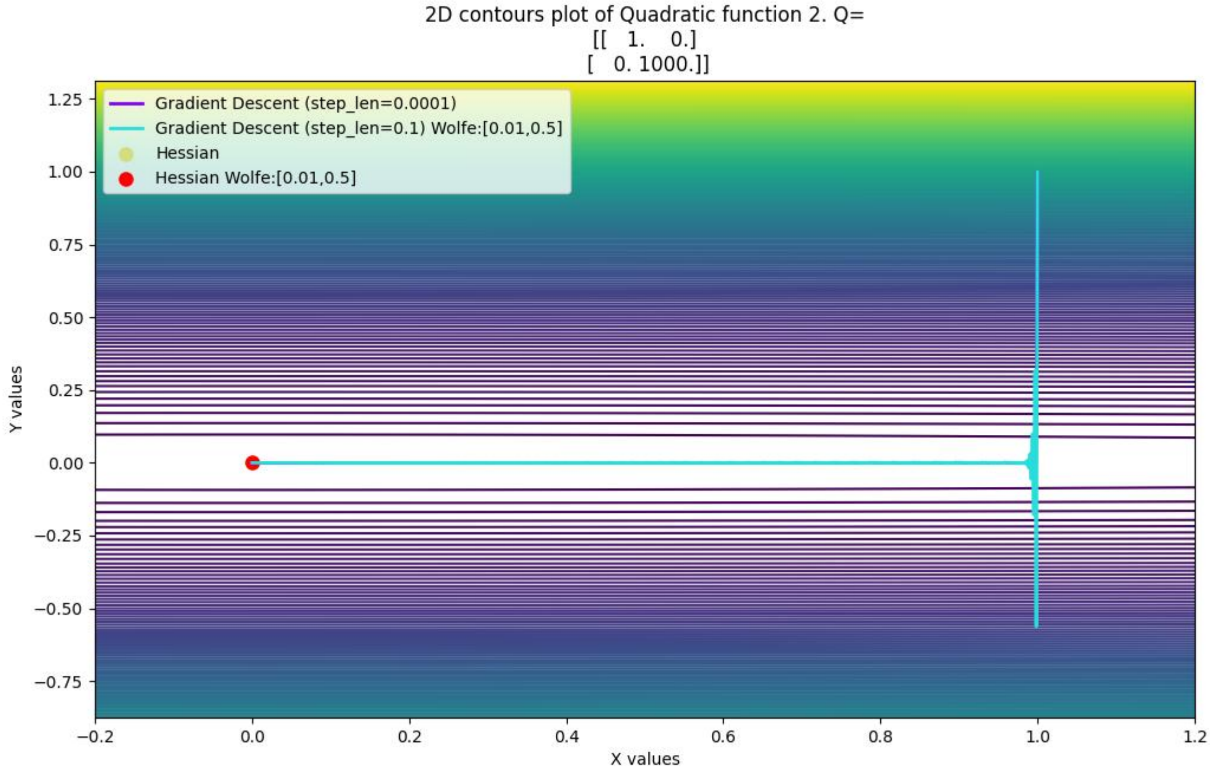
Wolfe=0.01 step_len=0.1 {'location': array([5.32536162e-05, 1.47963971e-07]), 'objective': 2.85784097857743e-09, 'success': True, 'num_iter': 4399}

Hessian

Wolfe=0 Wolfe=0.01 {'location': array([0., 0.]), 'objective': 0.0, 'success': True, 'num_iter': 2}


Convergence of function: Quadratic function 2. Q=
[[1. 0.]
[0. 1000.]]
Gradient Descent (step_len=0.0001)
Gradient Descent (step_len=0.1) Wolfe:[0.01,0.5]
Hessian
Hessian Wolfe:[0.01,0.5]
Objective function value
1000
800
600
400
200
0
1
10
100
1000
10000
Number of iterations


3D plot of Quadratic function 2. Q=
[[1. 0.]
[0. 1000.]]
Gradient Descent (step_len=0.0001)
Gradient Descent (step_len=0.1) Wolfe:[0.01,0.5]
Hessian
Hessian Wolfe:[0.01,0.5]
Objective function value
X values
Y values



Quad function 3:

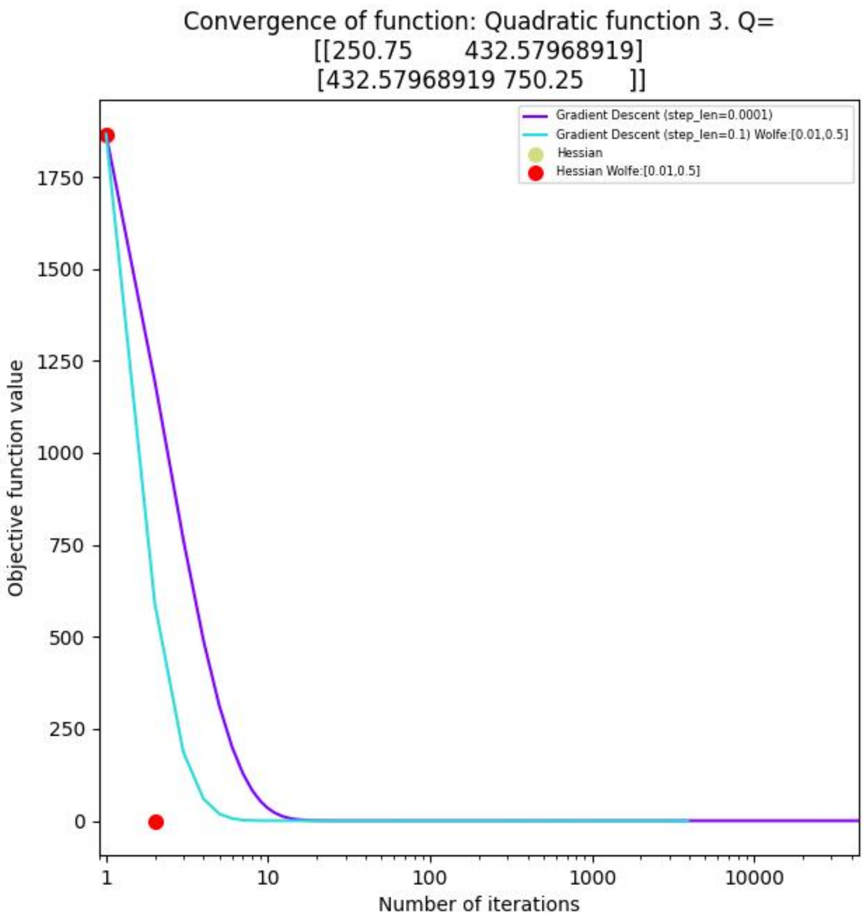
Gradient Descent

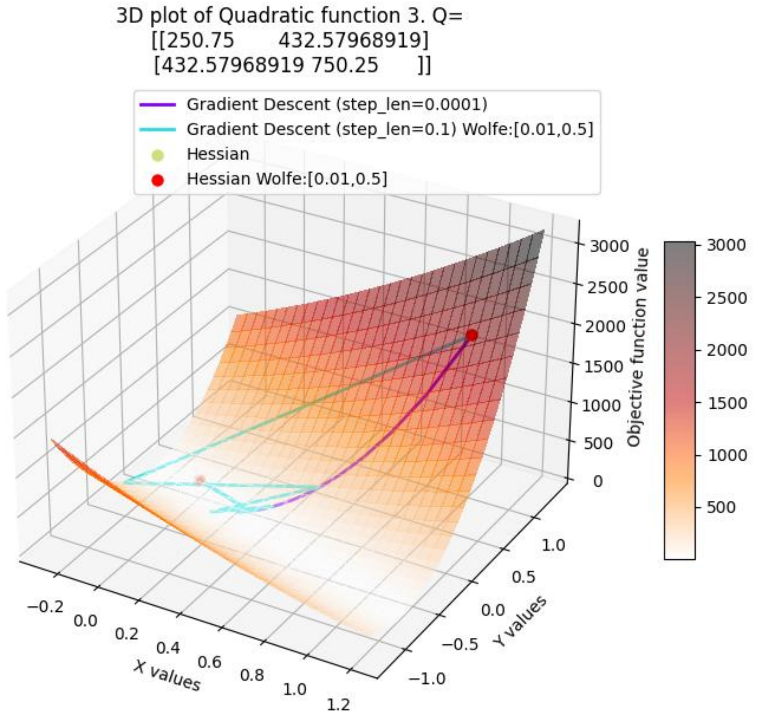
Wolfe=0 step_len=1e-4 {'location': array([4.32988945e-05, -2.49986284e-05]), 'objective': 2.4997256907573186e-09, 'success': True, 'num_iter': 44489}

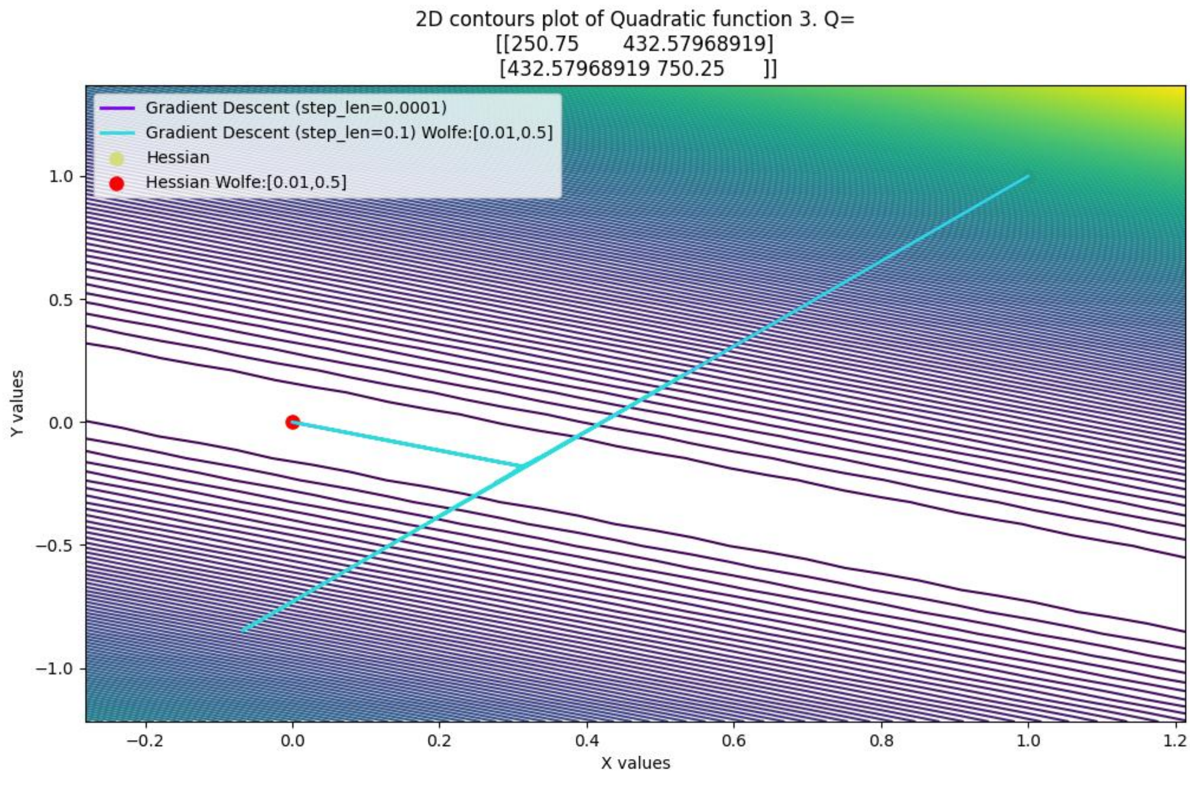
Wolfe=0.01 step_len=0.1 {'location': array([5.11836869e-05, -2.93608231e-05]), 'objective': 3.5089019632397813e-09, 'success': True, 'num_iter': 3905}

Hessian

Wolfe=0 Wolfe=0.01 {'location': array([-4.24105195e-14, 2.45359288e-14]), 'objective': 2.4025524106528963e-27, 'success': True, 'num_iter': 2}







Rosenbrock function:

$$f(x_1, x_2) = 100\left(x_2 - x_1^2\right)^2 + (1 - x_1)^2$$

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 2 \cdot 100\left(x_2 - x_1^2\right) \cdot (-2x_1) - 2(1 - x_1) \\ 2 \cdot 100\left(x_2 - x_1^2\right) \end{pmatrix} = \begin{pmatrix} -400x_1\left(x_2 - x_1^2\right) - 2(1 - x_1) \\ 200\left(x_2 - x_1^2\right) \end{pmatrix} = \begin{pmatrix} 400x_1^3 - 400x_1x_2 + 2x_1 - 2 \\ 200\left(x_2 - x_1^2\right) \end{pmatrix}$$

$$\nabla^2 f = \begin{pmatrix} \frac{\partial^2 f}{\partial {x_1}^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial {x_2}^2} \end{pmatrix} = \begin{pmatrix} 1200x_1^2 - 400x_2 + 2 & -400x_1 \\ -400x_1 & 200 \end{pmatrix}$$

Gradient Descent

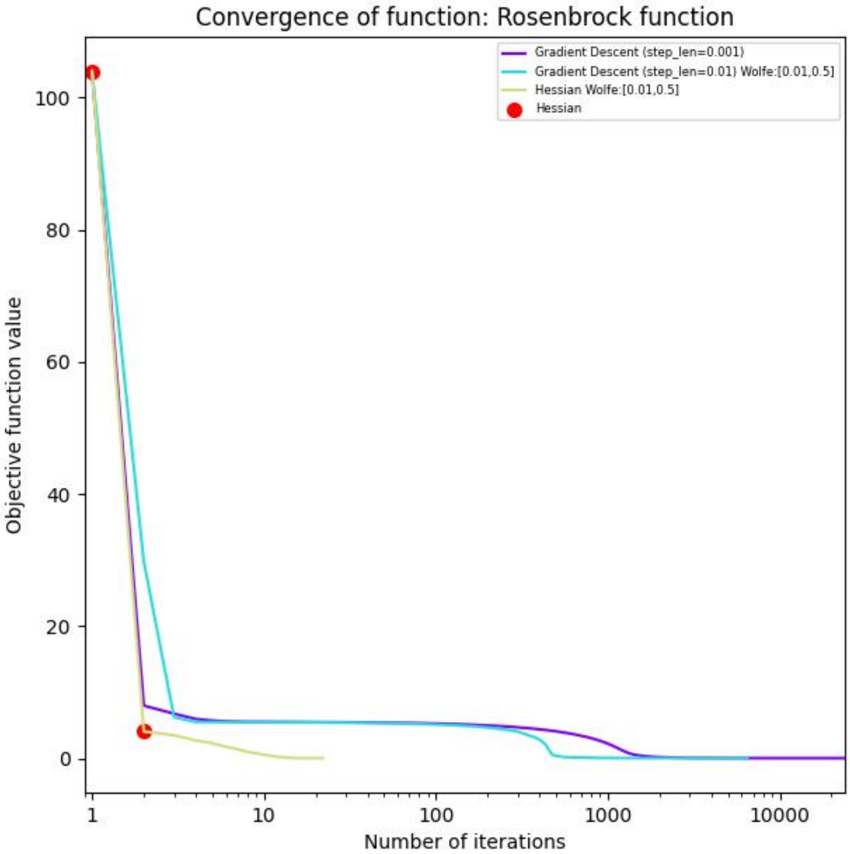
Wolfe=0 step_len=1e-3 {'location': array([0.99996466, 0.99992918]), 'objective': 1.2507997177822004e-09, 'success': True, 'num_iter': 23912}

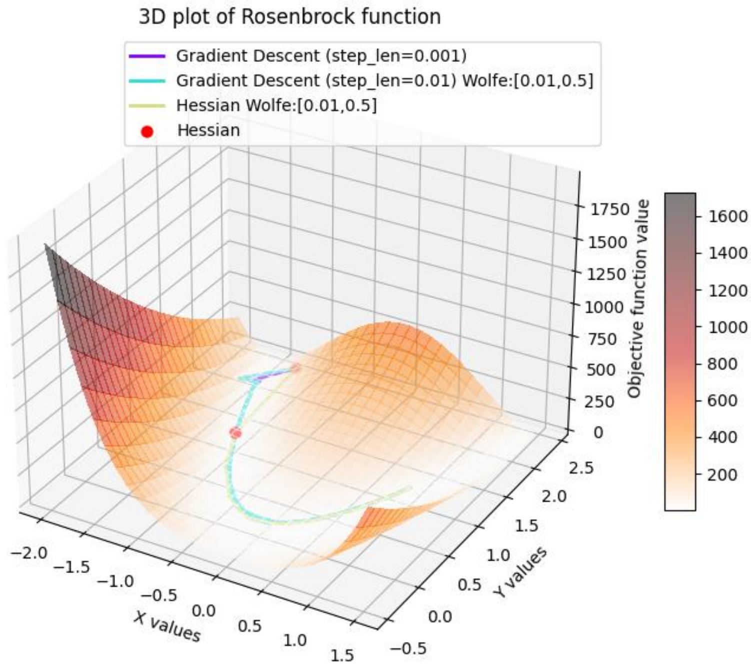
Wolfe=0.01 step_len=0.01 {'location': array([0.99989644, 0.99979314]), 'objective': 1.0731391146827696e-08, 'success': True, 'num_iter': 6488}

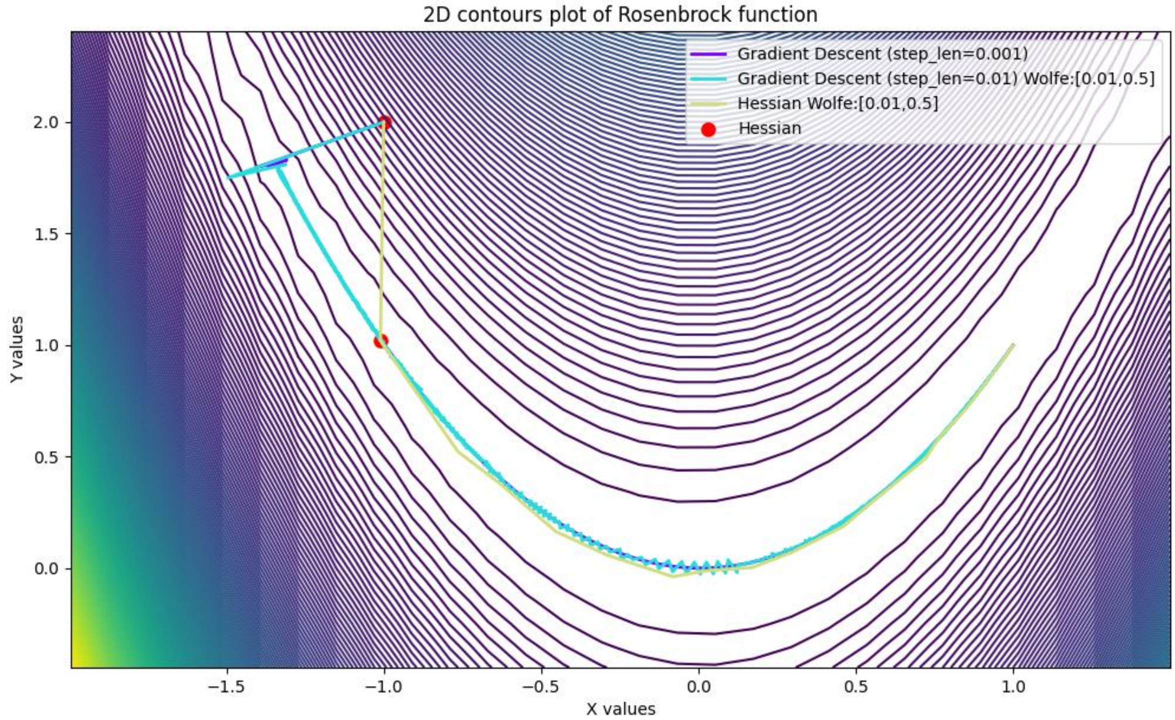
Hessian

Wolfe=0 Success is true because the function couldn't descent any further, although it wasn't close to the minimum {'location': array([-1.01005025, 1.0201005]), 'objective': 4.040303032827961, 'success': True, 'num_iter': 2}

Wolfe=0.01 {'location': array([1., 1.]), 'objective': 9.60931190172345e-29, 'success': True, 'num_iter': 22}







Boyd Function (3g):

$$f\left(x_1, x_2\right) = e^{x_1+3x_2-0.1} + e^{x_1-3x_2-0.1} + e^{-x_1-0.1} = f_1\left(x_1, x_2\right) + f_2\left(x_1, x_2\right) + f_3\left(x_1, x_2\right)$$

$$f_1\left(x_1, x_2\right) = e^{x_1+3x_2-0.1}$$

$$\nabla f_1 = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} \\ \frac{\partial f_1}{\partial x_2} \end{pmatrix} = \begin{pmatrix} e^{x_1+3x_2-0.1} \\ 3e^{x_1+3x_2-0.1} \end{pmatrix} = \begin{pmatrix} f_1\left(x_1, x_2\right) \\ 3f_1\left(x_1, x_2\right) \end{pmatrix}$$

$$\nabla^2 f_1 = \begin{pmatrix} \frac{\partial^2 f_1}{\partial {x_1}^2} & \frac{\partial^2 f_1}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f_1}{\partial x_1 \partial x_2} & \frac{\partial^2 f_1}{\partial {x_2}^2} \end{pmatrix} = \begin{pmatrix} f_1\left(x_1, x_2\right) & 3f_1\left(x_1, x_2\right) \\ 3f_1\left(x_1, x_2\right) & 9f_1\left(x_1, x_2\right) \end{pmatrix}$$

$$f_2\left(x_1, x_2\right) = e^{x_1-3x_2-0.1}$$

$$\nabla f_2 = \begin{pmatrix} \frac{\partial f_2}{\partial x_1} \\ \frac{\partial f_2}{\partial x_2} \end{pmatrix} = \begin{pmatrix} e^{x_1-3x_2-0.1} \\ -3e^{x_1-3x_2-0.1} \end{pmatrix} = \begin{pmatrix} f_2\left(x_1, x_2\right) \\ -3f_2\left(x_1, x_2\right) \end{pmatrix}$$

$$\nabla^2 f_2 = \begin{pmatrix} \frac{\partial^2 f_2}{\partial {x_1}^2} & \frac{\partial^2 f_2}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f_2}{\partial x_1 \partial x_2} & \frac{\partial^2 f_2}{\partial {x_2}^2} \end{pmatrix} = \begin{pmatrix} f_2\left(x_1, x_2\right) & -3f_2\left(x_1, x_2\right) \\ -3f_2\left(x_1, x_2\right) & 9f_2\left(x_1, x_2\right) \end{pmatrix}$$

$$f_3\left(x_1, x_2\right) = e^{-x_1-0.1}$$

$$\nabla f_3 = \begin{pmatrix} \frac{\partial f_3}{\partial x_1} \\ \frac{\partial f_3}{\partial x_2} \end{pmatrix} = \begin{pmatrix} -e^{-x_1-0.1} \\ 0 \end{pmatrix} = \begin{pmatrix} -f_3\left(x_1, x_2\right) \\ 0 \end{pmatrix}$$

$$\nabla^2 f_3 = \begin{pmatrix} \frac{\partial^2 f_3}{\partial {x_1}^2} & \frac{\partial^2 f_3}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f_3}{\partial x_1 \partial x_2} & \frac{\partial^2 f_3}{\partial {x_2}^2} \end{pmatrix} = \begin{pmatrix} f_3\left(x_1, x_2\right) & 0 \\ 0 & 0 \end{pmatrix}$$

$$\nabla f = \nabla f_1 + \nabla f_2 + \nabla f_3 = \begin{pmatrix} f_1 + f_2 - f_3 \\ 3f_1 - 3f_2 \end{pmatrix}$$

$$\nabla^2 f = \nabla^2 f_1 + \nabla^2 f_2 + \nabla^2 f_3 = \begin{pmatrix} f_1 + f_2 + f_3 & 3f_1 - 3f_2 \\ 3f_1 - 3f_2 & 9f_1 + 9f_2 \end{pmatrix}$$

$$\nabla f = 0 \Rightarrow \left\{ \begin{matrix} f_1 = f_2 \\ 2f_1 - f_3 = 0 \end{matrix} \right\} \Rightarrow \left\{ \begin{matrix} x_2 = 0 \\ 2e^{2x_1} = 1 \end{matrix} \right\} \Rightarrow x_1 = \tfrac{1}{2} ln \tfrac{1}{2}$$

$$f\left(\nabla f = 0\right) = e^{\frac{1}{2}ln\frac{1}{2}-0.1} + e^{\frac{1}{2}ln\frac{1}{2}-0.1} + e^{-\frac{1}{2}ln\frac{1}{2}-0.1} = e^{-0.1}\left(e^{\frac{1}{2}ln\frac{1}{2}} + e^{\frac{1}{2}ln\frac{1}{2}} + e^{-\frac{1}{2}ln\frac{1}{2}}\right) = e^{-0.1}\left(2e^{\frac{1}{2}ln\frac{1}{2}} + e^{-\frac{1}{2}ln\frac{1}{2}}\right) = e^{-0.1}\left(\frac{e^{ln2+ln\frac{1}{2}}+1}{e^{\frac{1}{2}ln\frac{1}{2}}}\right) = e^{-0.1}\frac{1+1}{\sqrt{\frac{1}{2}}} = e^{-0.1}\left(2\sqrt{2}\right)$$

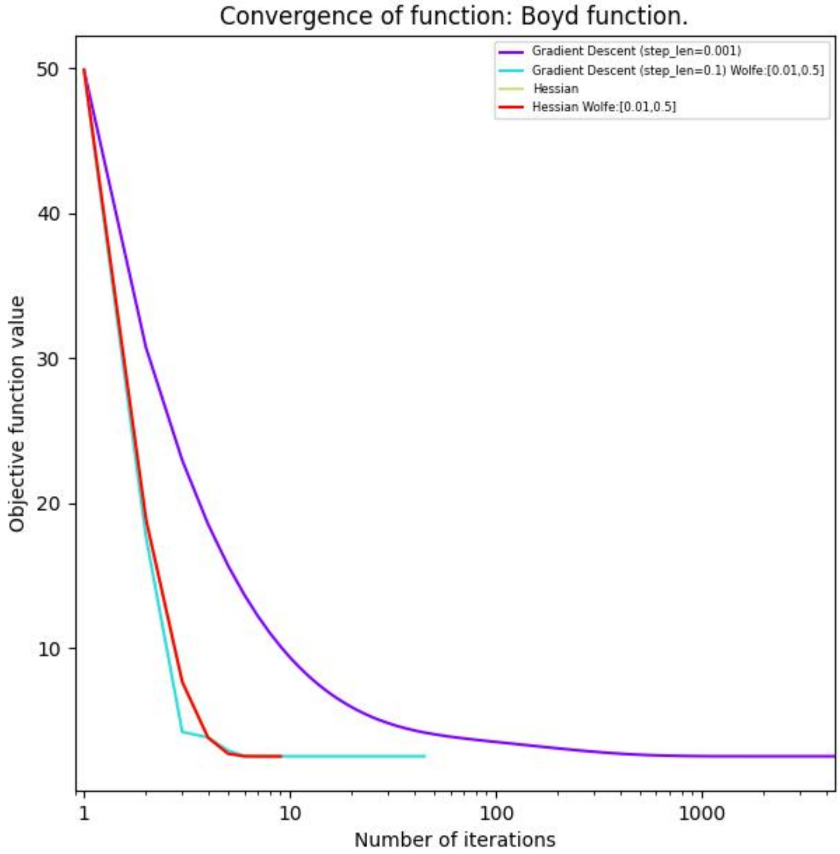
Gradient Descent

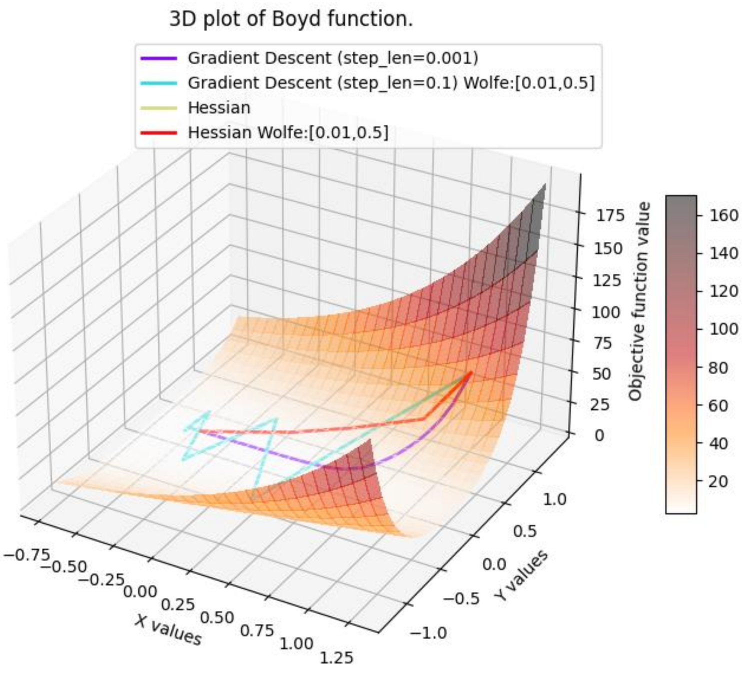
Wolfe=0 step_len=1e-3 {'location': array([-3.46561260e-01, 9.10667386e-18]), 'objective': 2.5592666968527538, 'success': True, 'num_iter': 4425}

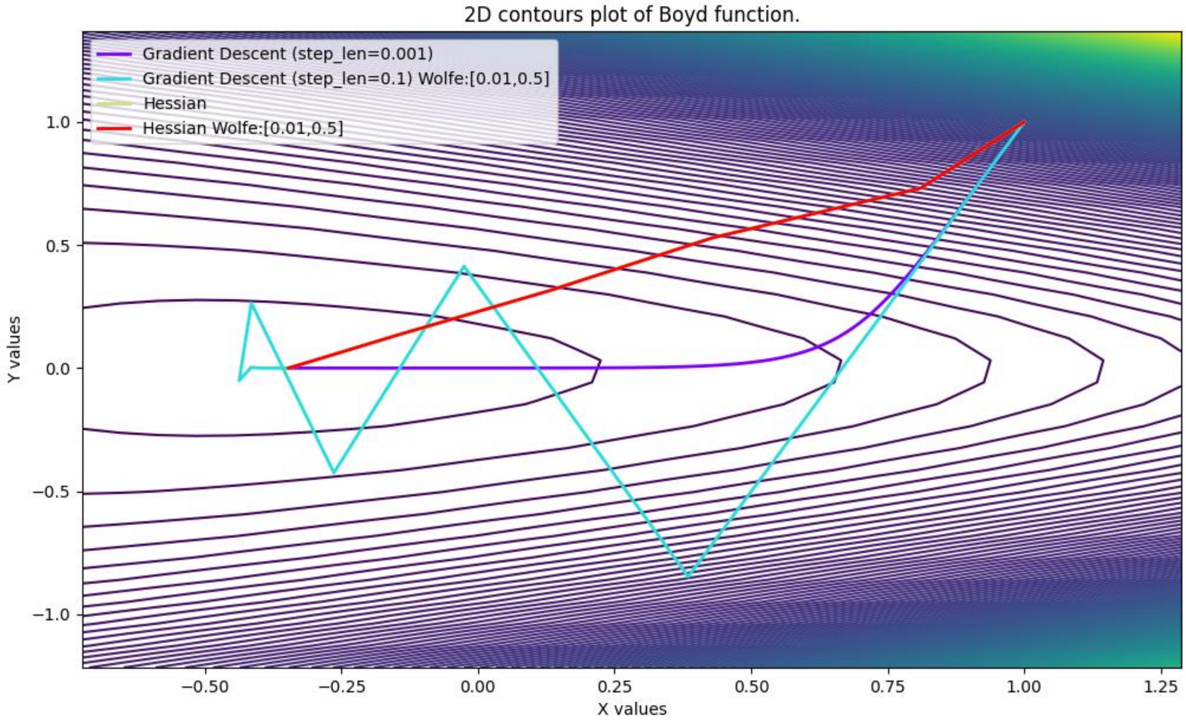
Wolfe=0.01 step_len=0.1 {'location': array([-3.46574494e-01, -7.26961280e-18]), 'objective': 2.559266696659261, 'success': True, 'num_iter': 45}

Hessian

Wolfe=0 Wolfe=0.01 {'location': array([-3.46573590e-01, 6.80690347e-12]), 'objective': 2.5592666966582156, 'success': True, 'num_iter': 9}


Convergence of function: Boyd function.
Gradient Descent (step_len=0.001)
Gradient Descent (step_len=0.1) Wolfe:[0.01,0.5]
Hessian
Hessian Wolfe:[0.01,0.5]
Objective function value
Number of iterations
50
40
30
20
10
1
10
100
1000


3D plot of Boyd function.
Gradient Descent (step_len=0.001)
Gradient Descent (step_len=0.1) Wolfe:[0.01,0.5]
Hessian
Hessian Wolfe:[0.01,0.5]
Objective function value
X values
Y values
−0.75
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00
1.25
−1.0
−0.5
0.0
0.5
1.0
0
25
50
75
100
125
150
175
20
40
60
80
100
120
140
160


2D contours plot of Boyd function.
Gradient Descent (step_len=0.001)
Gradient Descent (step_len=0.1) Wolfe:[0.01,0.5]
Hessian
Hessian Wolfe:[0.01,0.5]
1.0
0.5
0.0
−0.5
−1.0
Y values
−0.50
−0.25
0.00
0.25
0.50
0.75
1.00
1.25
X values

In [ ]: