Numerical Optimization 2024 - Homework 4

Aral Cimcim, k11720457

Artificial Intelligence, JKU Linz

April 16, 2024

Problem 1.

The contour plot of the Rosenbrock function, points $(1,1), (0,0), (-1,1), (0,1)$ and the function values where $f(x)$ attains the values $0, 1, 4$ are shown in Figure: 1.

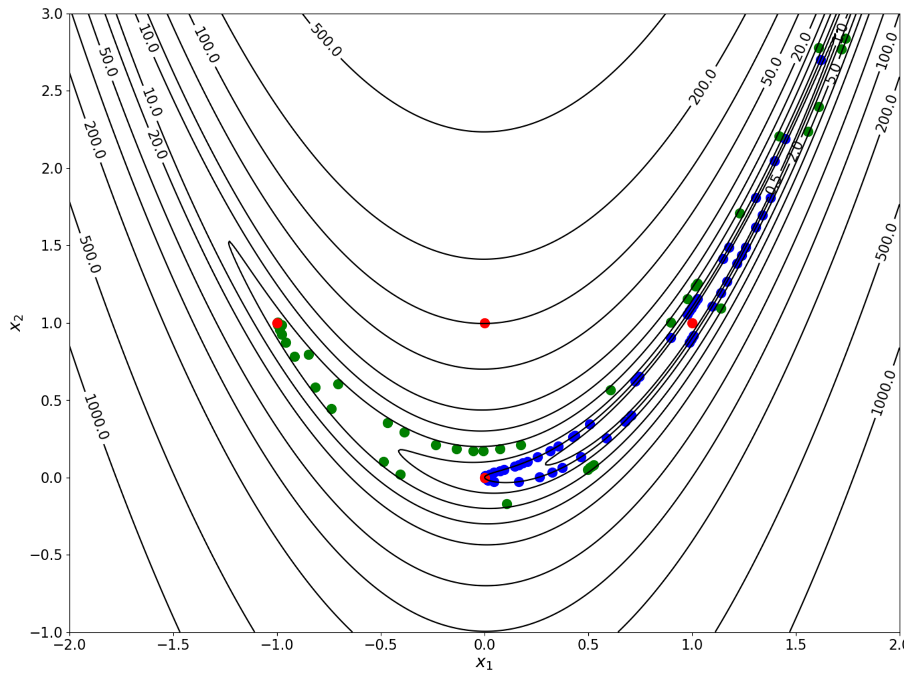

Figure 1: Contour plot of $f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$

Problem 2.

The gradient of the Rosenbrock function is defined as:

$$\nabla f(x) = \begin{bmatrix} -400x_1(x_2 - x_1^2) - 2(1 - x_1) \\ 200(x_2 - x_1^2) \end{bmatrix} \tag{1}$$

The Hessian of the Rosenbrock function is defined as:

$$\nabla^2 f(x) = \begin{bmatrix} -400(x_2 - 3x_1^2) + 2 & -400x_1 \\ -400x_1 & 200 \end{bmatrix} \tag{2}$$

Steepest descent moves along $p_k = -\nabla f(x_k)$ at each step. For α_k as the global minimizer of $\phi(\alpha) = f(x_k + \alpha p_k)$, the following statements hold:

i) I expect the steepest descent method to converge to the global min. of the function at $(1, 1)$ (depending on the starting point and α, the algorithm will find the local min.)

ii) The rate of convergence will be linear. I plotted the graph for the convergence, C is less than 10^{-5} after 5000 iterations, shown in Figure: 3. (starting at $(-1, 1)$ after 10000 iterations, $f(x) = 1.0256665448988478e - 08$, see below for the exact values)

iii) Starting at $x_0 = (-1, 1)$, the trajection does not go directly to the limit but move up and down and follow the deep valley of the function to reach the global min. at $(1, 1)$, shown in Figure: 7.

```
Iteration 0: 4
Iteration 1: 101.0
Iteration 2: 101.0
Iteration 3: 0.980101239863517
Iteration 4: 0.7790048295335187
Iteration 5: 0.7487672432382151
Iteration 6: 0.7281119086406248
Iteration 7: 0.7045080495384312
...
Iteration 9989: 1.0309324143070513e-08
Iteration 9990: 1.0304426606525326e-08
Iteration 9991: 1.0299695480712044e-08
Iteration 9992: 1.0294813298468107e-08
Iteration 9993: 1.0290097502686015e-08
Iteration 9994: 1.0285230626158758e-08
Iteration 9995: 1.0280530111705617e-08
Iteration 9996: 1.0275678491906015e-08
Iteration 9997: 1.0270993209860194e-08
Iteration 9998: 1.0266156798725102e-08
Iteration 9999: 1.0261486700313638e-08
Iteration 10000: 1.0256665448988478e-08
```

Problem 3.

Line search with backtracking is defined as:

$$\text{choose } \bar{\alpha} > 0, \rho, c \in (0,1); \text{Set } \alpha \leftarrow \bar{\alpha};$$

$$\text{repeat until } f(x_k + \alpha p_k) \leq f(x_k) + c\alpha \nabla f_k^T p_k$$

$$\alpha \leftarrow \rho\, \alpha$$

$$\text{end at } \alpha_k = \alpha$$

i) The plot for the backtracking line search method (BLSM) for $\rho = 0.5, c = 0.0001, \alpha = 1$ is shown in Figure: 2, after 5 iterations the function begins to converge. Starting from x_k at $(1.2, 1.2)$, after 5000 iterations the function value is minimized to the order of 10^{-5}, shown in Figure: 3. The corresponding contour and line plots for $\alpha = 1$ and $\alpha = 0.3$ are shown in Figure: 4, Figure: 5 and Figure: 6.

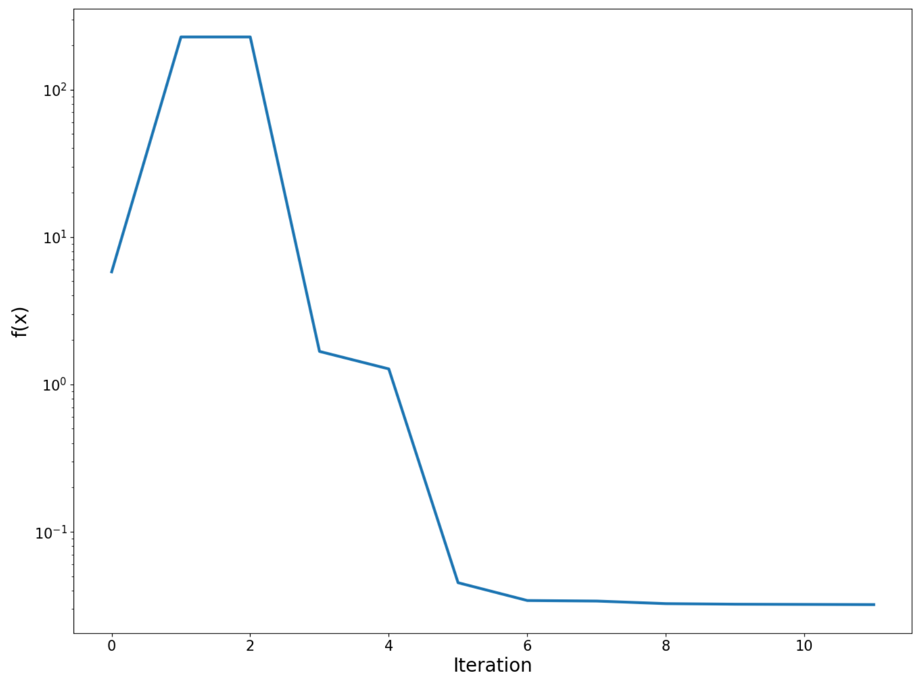

Figure 2: BLSM for $\rho = 0.5, c = 0.0001, \alpha = 1$ for 10 iterations

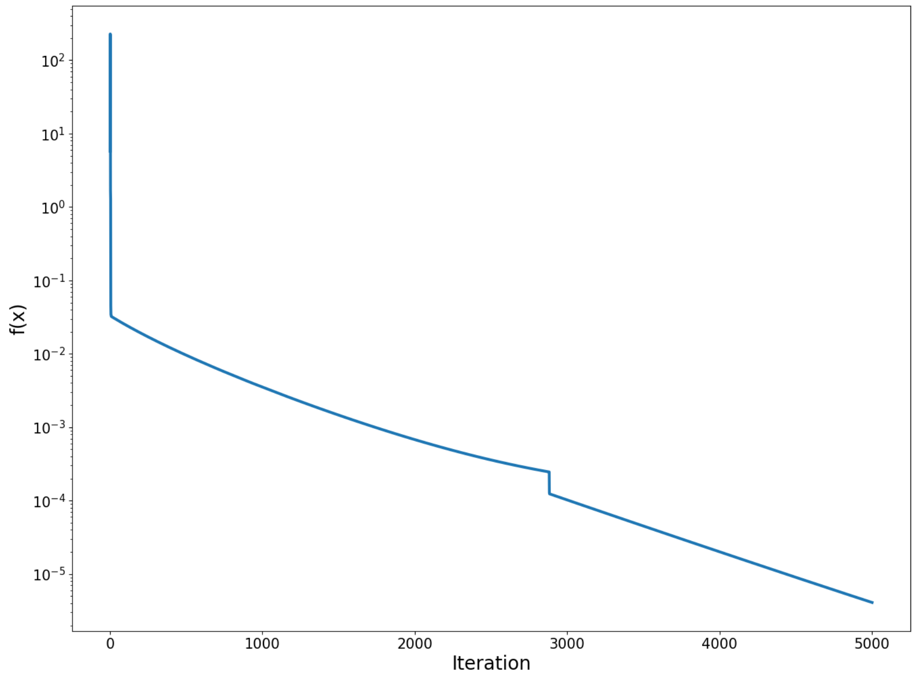

Figure 3: BLSM for hyperparameters $\rho = 0.5, c = 0.0001, \alpha = 1$ for 5000 iterations

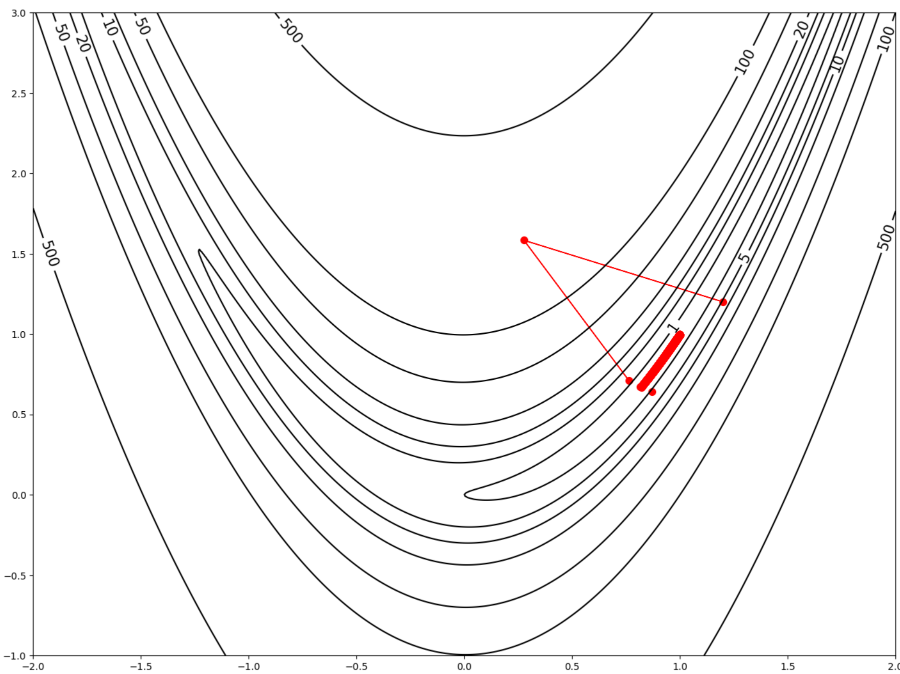

Figure 4: Steepest descent with BLSM after 5000 iterations for hyperparameters ($\rho = 0.5, c = 0.0001, \alpha = 1$)

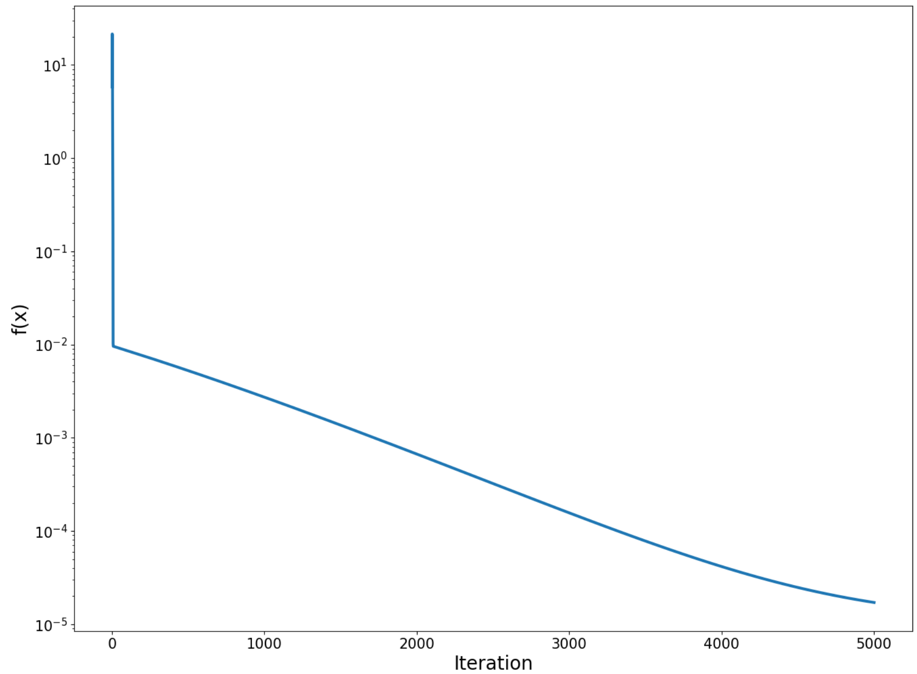

Figure 5: BLSM for hyperparameters $\rho = 0.5, c = 0.0001, \alpha = 0.3$ for 5000 iterations

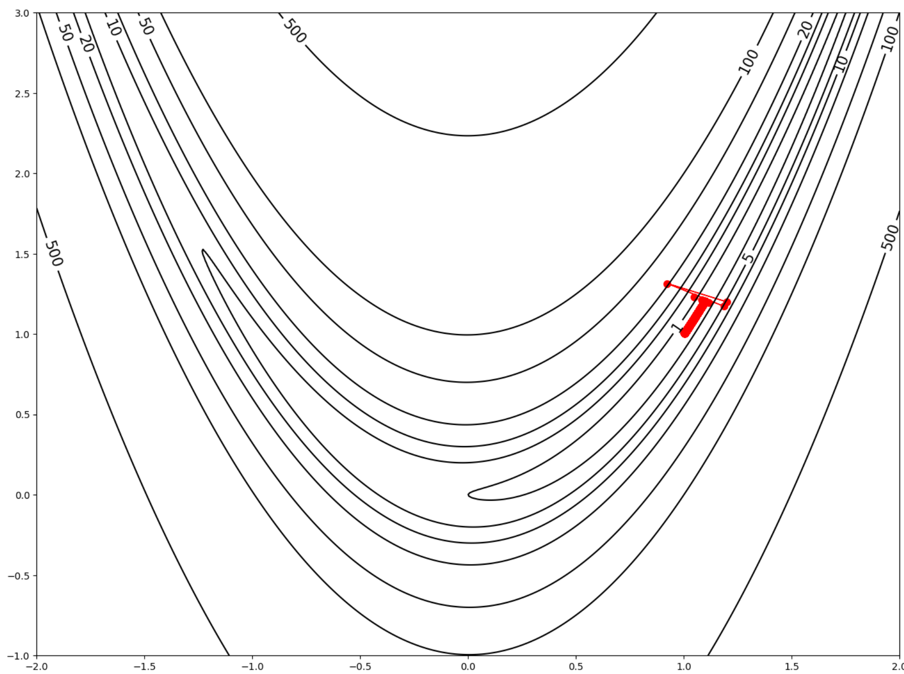

Figure 6: Steepest descent with BLSM after 5000 iterations for hyperparameters ($\rho = 0.5, c = 0.0001, \alpha = 0.3$)

ii) For the starting point $(-1, 1)$, the plot is shown in Figure: 7 (included the starting point $(-1, 1)$ in the submission as it is easier to show the changes). For the starting point $(-1.2, 1)$, the plot is shown in Figure: 8. With increasing α values the convergence rate to reach the global min. is faster.

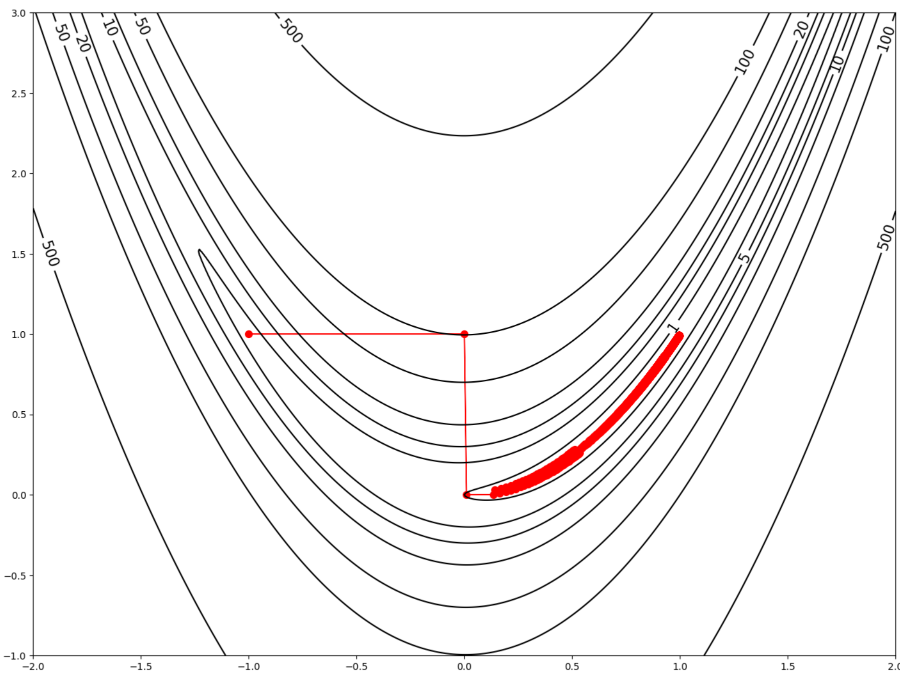

Figure 7: Steepest descent with BLSM after 5000 iterations for stating point $(-1, 1)$ with hyperparameters $(\rho = 0.5, c = 0.0001, \alpha = 1)$

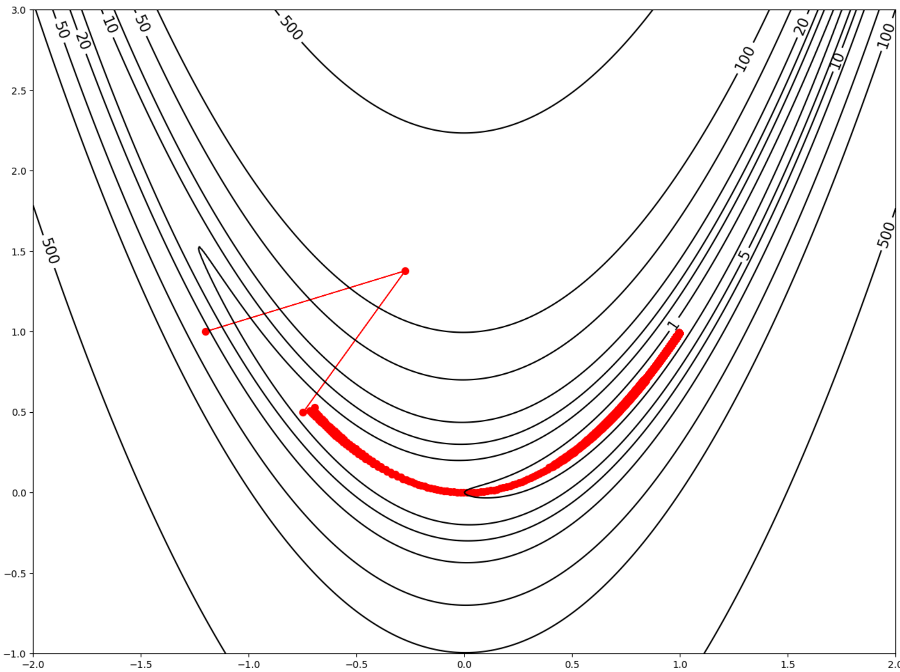

Figure 8: Steepest descent with BLSM after 5000 iterations for stating point $(-1.2, 1)$ with hyper-parameters $(\rho = 0.5, c = 0.0001, \alpha = 1)$