Optimization Methods - Assignment 2
CS1.404 (Spring 2024)

Prashant Kumar (2022201058)

April 8, 2024

---

1 Derive the Jacobians and Hessian for the new functions.

1. **Matyas function**

$$f(x) = 0.26(x_1^2 + x_2^2) - 0.48x_1 \times x_2$$

- **Jacobian**

$$\frac{df}{dx_1} = 0.26(2x_1) - 0.48x_2$$
$$\frac{df}{dx_1} = 0.52x_1 - 0.48x_2 \tag{1}$$
$$\frac{df}{dx_2} = 0.26(2x_2) - 0.48x_1$$
$$\frac{df}{dx_2} = 0.52x_2 - 0.48x_1 \tag{2}$$

Therefore, the Jacobian is,

$$\begin{bmatrix} \frac{df}{dx_1} \\ \frac{df}{dx_2} \end{bmatrix}$$

$$\begin{bmatrix} 0.52x_1 - 0.48x_2 \\ 0.52x_2 - 0.48x_1 \end{bmatrix}$$

- **Hessian**

$$\frac{d^2f}{dx_1^2} = 0.52$$
$$\frac{d^2f}{dx_2^2} = 0.52$$
$$\frac{d^2f}{dx_1x_2} = -0.48$$

Therefore, the Hessian is,

$$\begin{bmatrix} \frac{d^2f}{dx_1^2} & \frac{d^2f}{dx_1x_2} \\ \frac{d^2f}{dx_2x_1} & \frac{d^2f}{dx_2^2} \end{bmatrix}$$

$$\begin{bmatrix} 0.52 & -0.48 \\ -0.48 & 0.52 \end{bmatrix}$$

2. **Rotated Hyper-Ellipsoid Function**

$$f(x) = \sum_{i=1}^{d} \sum_{j=1}^{i} x_j^2$$

For simplicity, here I'm taking d as 2, then f(x) can be re-written as

$$f(x) = \sum_{i=1}^{2} \sum_{j=1}^{i} x_j^2$$

$$f(x) = (x_1^2) + (x_1^2 + x_2^2)$$
$$f(x) = 2x_1^2 + x_2^2$$

- **Jacobian**

$$\frac{df}{dx_1} = 4x_1 \tag{3}$$
$$\frac{df}{dx_2} = 2x_2 \tag{4}$$

Therefore, the Jacobian is,

$$\begin{bmatrix} \frac{df}{dx_1} \\ \frac{df}{dx_2} \end{bmatrix}$$

$$\begin{bmatrix} 4x_1 \\ 2x_2 \end{bmatrix}$$

- **Hessian**

$$\frac{d^2f}{dx_1^2} = 4$$
$$\frac{d^2f}{dx_1x_2} = 0$$
$$\frac{d^2f}{dx_2^2} = 2$$

Therefore, the Hessian is,

$$\begin{bmatrix} \frac{d^2f}{dx_1^2} & \frac{d^2f}{dx_1x_2} \\ \frac{d^2f}{dx_2x_1} & \frac{d^2f}{dx_2^2} \end{bmatrix}$$

$$\begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}$$

2 Using the Jacobians and Hessians, calculate the minimas for the new functions.

1. **Matyas function**

$$f(x) = 0.26(x_1^2 + x_2^2) - 0.48x_1 \times x_2$$

The Jacobian is,

$$\begin{bmatrix} 0.52x_1 - 0.48x_2 \\ 0.52x_2 - 0.48x_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Solving for x_1 and x_2 we get:

$$0.52x_1 - 0.48x_2 = 0$$
$$0.52x_2 - 0.48x_1 = 0$$
$$x_1 = 0 \text{ and } x_2 = 0$$

The Hessian is,

$$\begin{bmatrix} 0.52 & -0.48 \\ -0.48 & 0.52 \end{bmatrix}$$

Since the Hessian is positive definite because we can clearly see, all the leading principal minors are positive. Therefore $[0, 0]$ is the minima for the function.

2. **Rotated Hyper-Ellipsoid Function**

$$f(x) = \sum_{i=1}^{2} \sum_{j=1}^{i} x_j^2$$

The Jacobian is,

$$\begin{bmatrix} 4x_1 \\ 2x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Solving for x_1 and x_2 we get:

$$4x_1 = 0$$

$$2x_2 = 0$$

$$x_1 = 0 \text{ and } x_2 = 0$$

The Hessian is,

$$\begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}$$

Since the Hessian is positive definite because we can clearly see, all the leading principal minors are positive. Therefore $[0, 0]$ is the minima for the function.

3 State which algorithms failed to converge and under which circumstances.

The reason for not converging successfully is an inadequate number of iterations, Becoming trapped at local minima, overflow, division by zero, etc...
Rosenbrock function with point [2.0, 2, 2, -2] as the starting point is not converging Fletcher-Reeves because of a slightly smaller number of iterations. The same case holds with points [3.0, 3, 3, 3] with approach SR1, [2.,-2.,-2.,2.] with approach SR1, [2.,-2.,-2.,2.] with approach DFP, and [-2.,2.,2.,2.] with method SR1.

4 Plot $f(x)$ vs iterations and $f'(x)$ vs iterations.

5 Make a contour plot with arrows indicating the direction of updates for all 2-D functions.

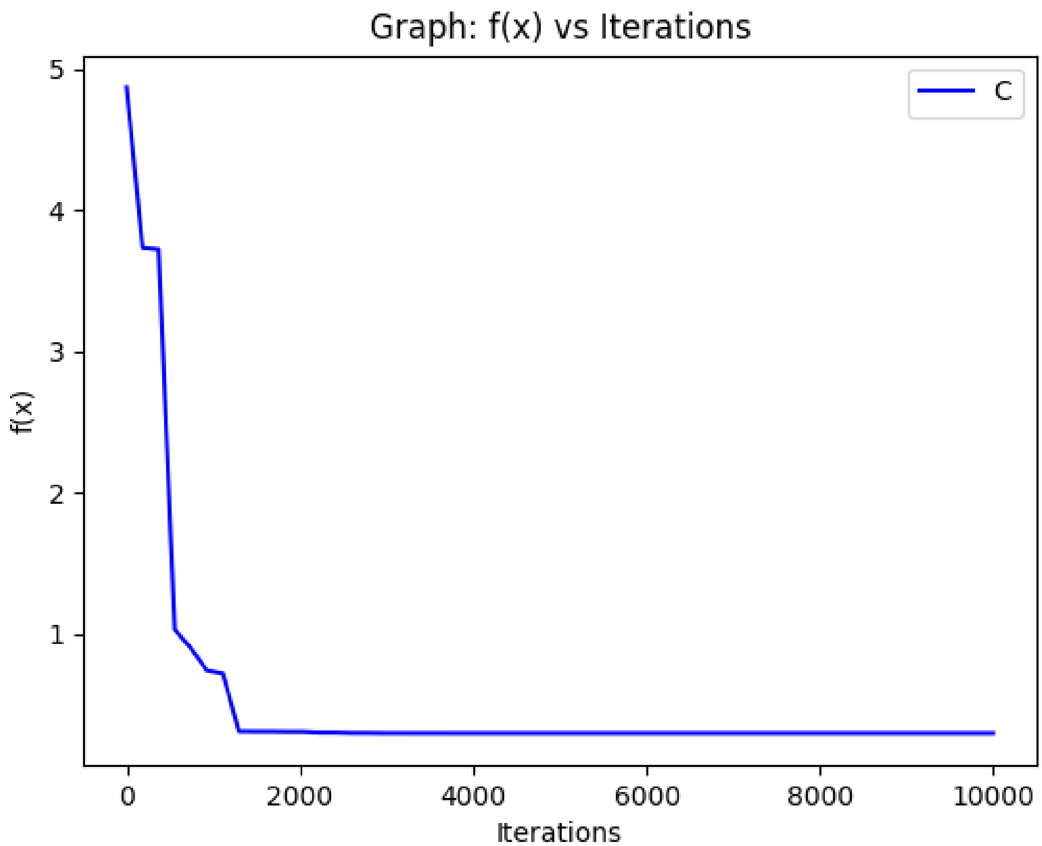

Figure 1: three_hump_camel_function_[2. 1.]_PolakRibiere_vals

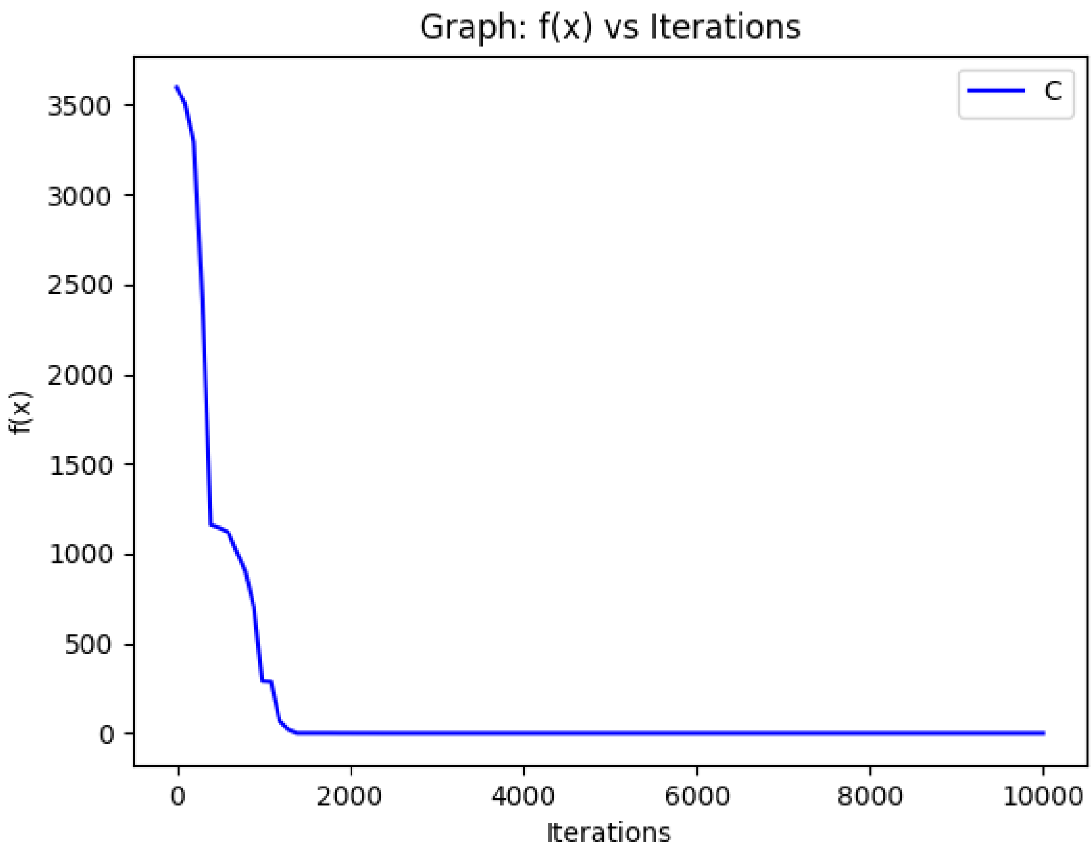

Figure 2: hyperEllipsoid_function_[10. -10. 15. 15. -20. 11. 312.]_PolakRibiere_vals

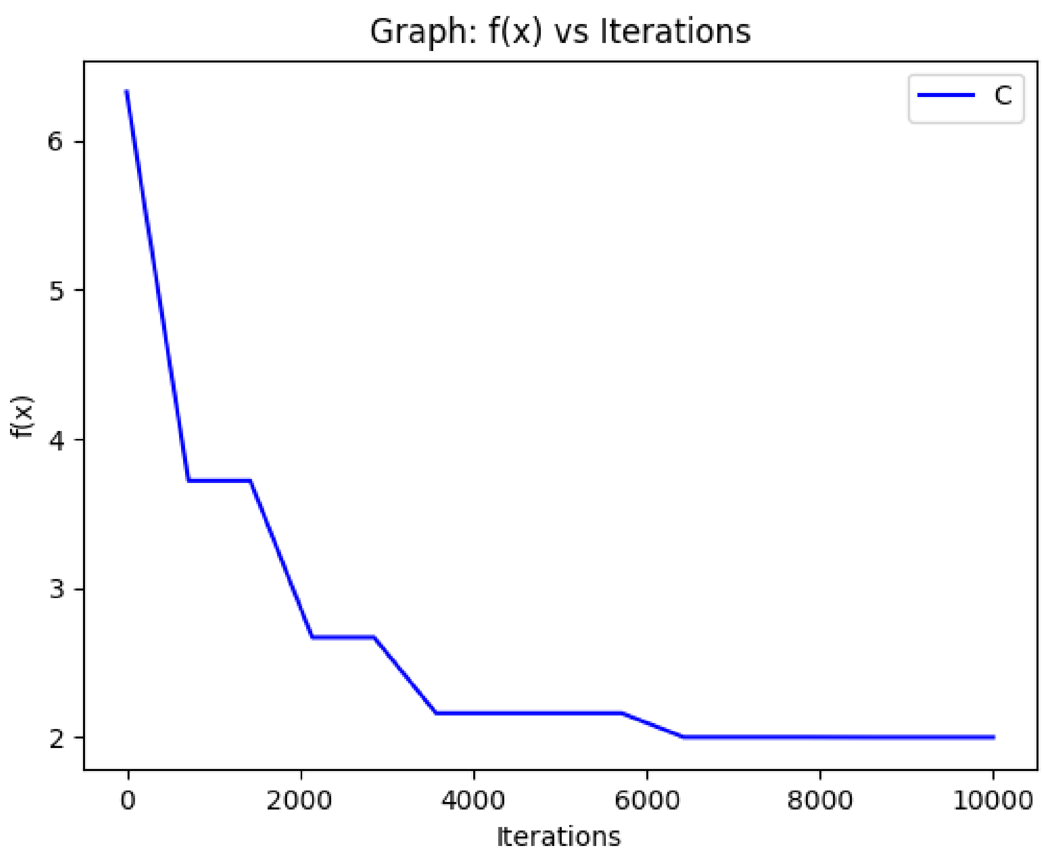

Figure 3: func_1_[3. 3.]_HestenesStiefel_vals

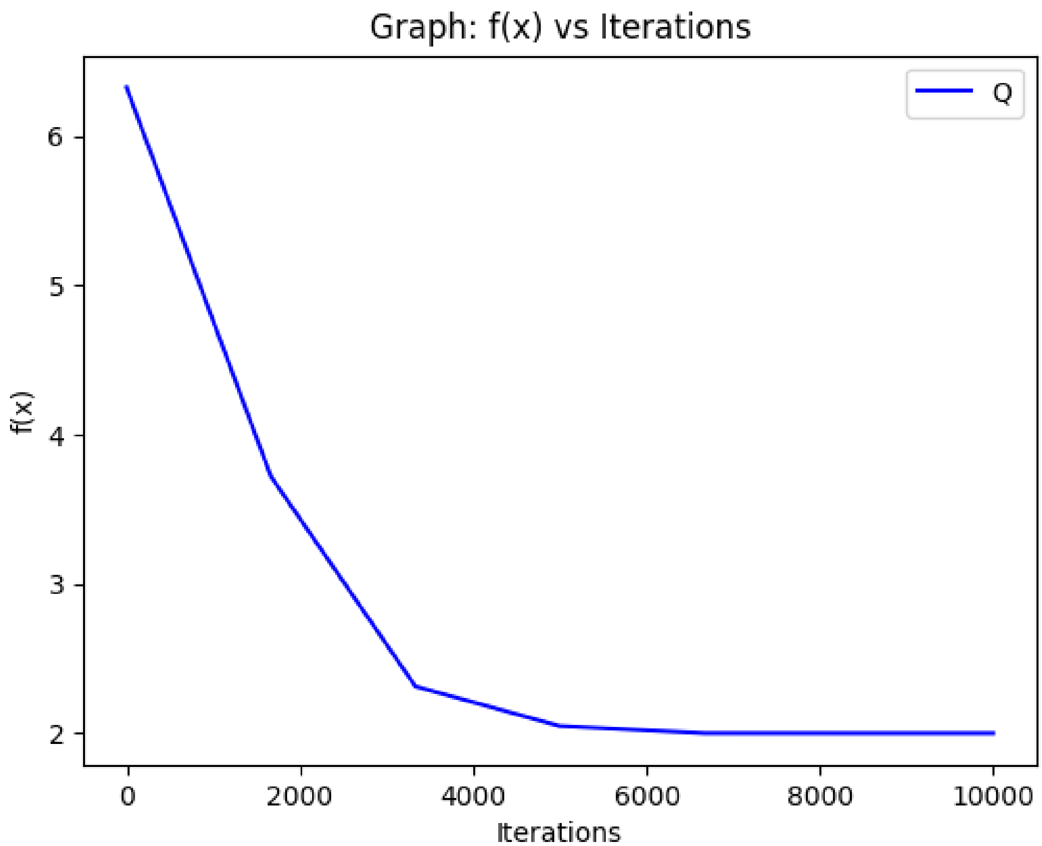

Figure 4: func_1_[3. 3.]_DFP_vals

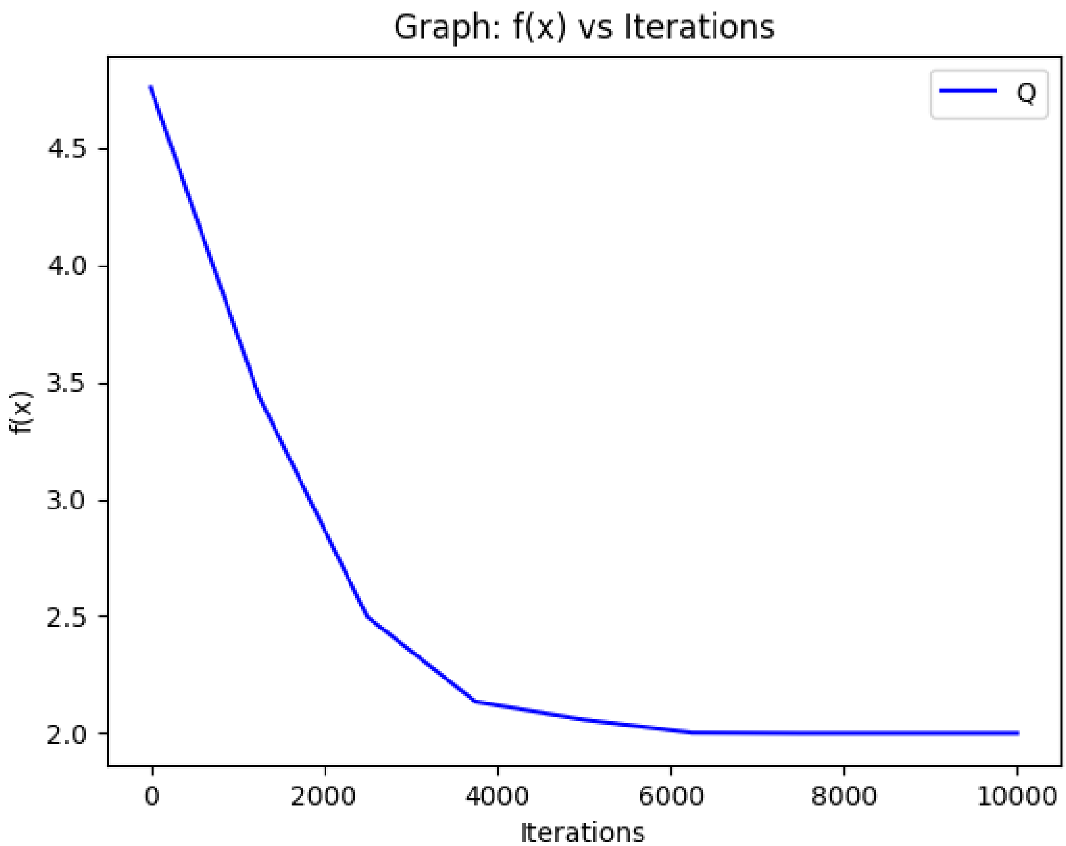

Figure 5: func_1_[-3.5 0.5]_SR1_vals

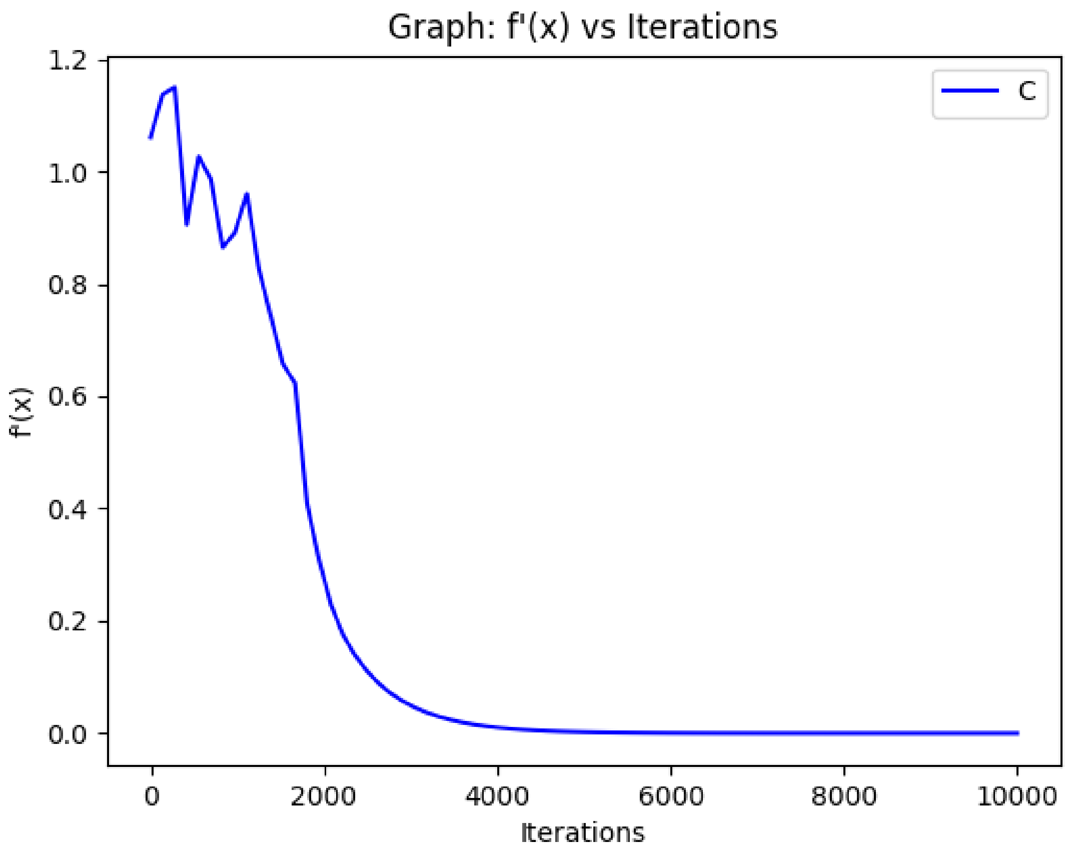

Figure 6: func_1_derivative_[-3.5 0.5]_FletcherReeves_grad

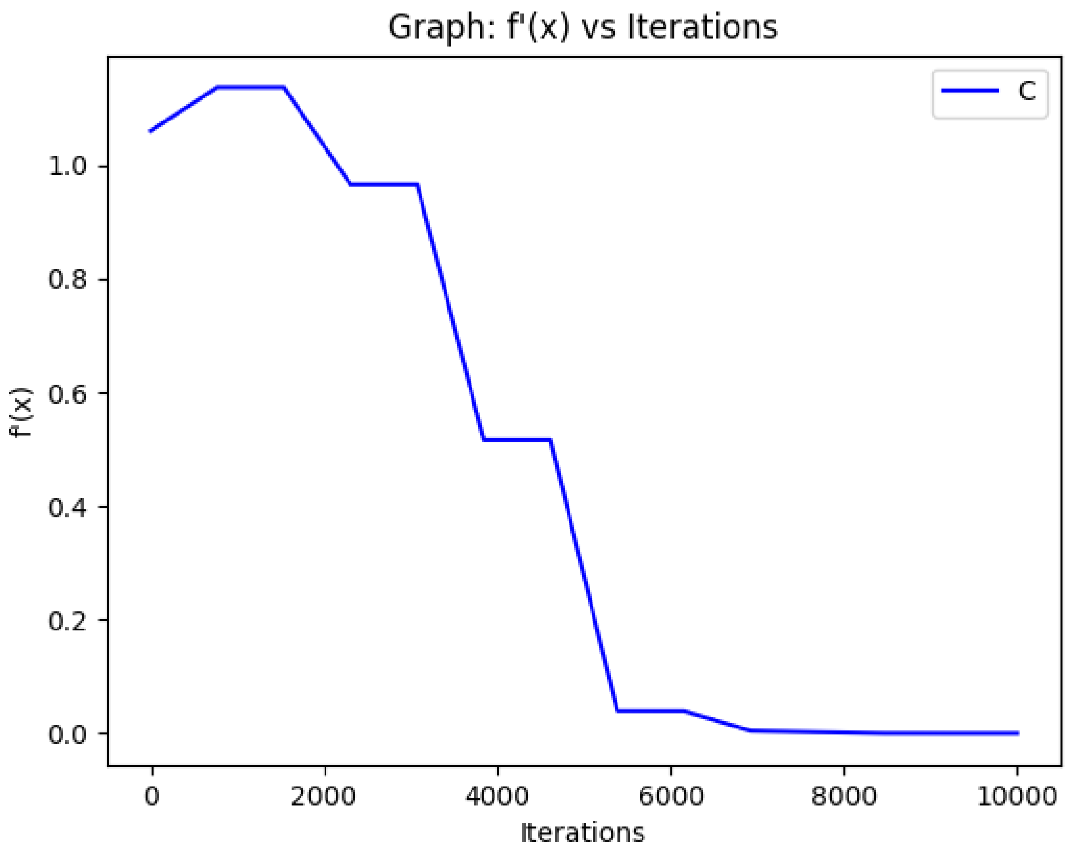

Figure 7: func_1_derivative_[-3.5 0.5]_PolakRibiere_grad

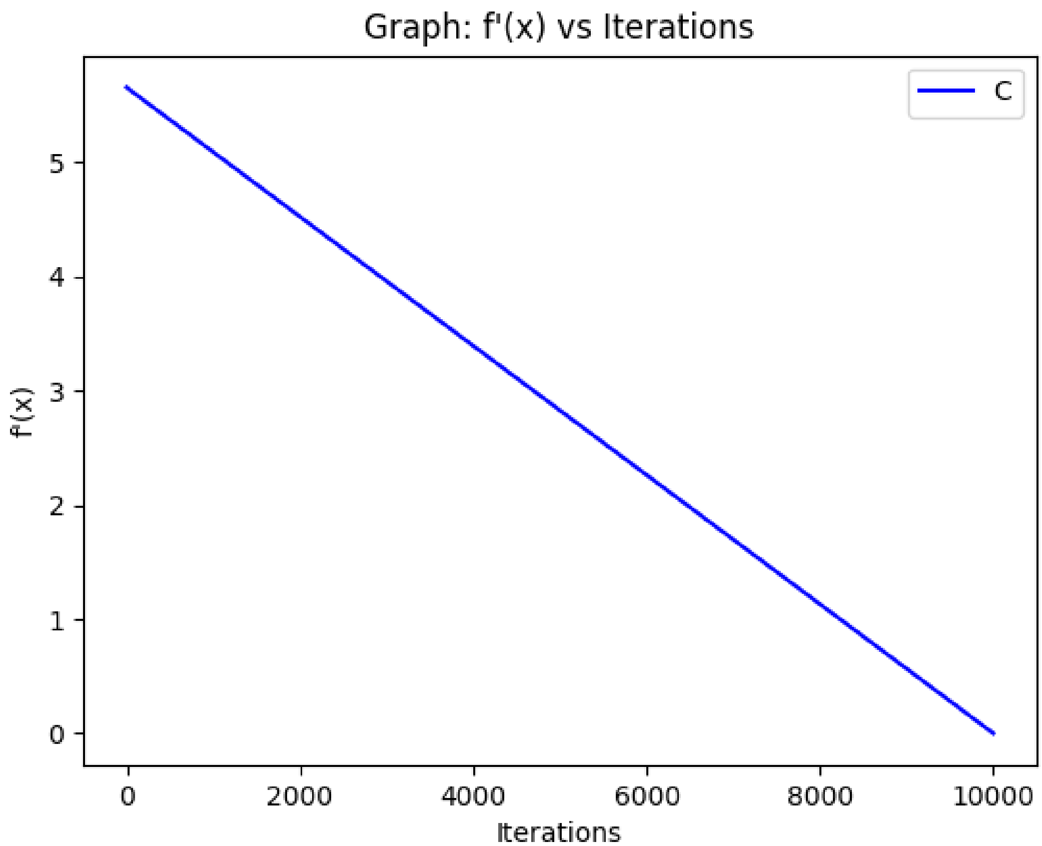

Figure 8: trid_function_derivative_[-2. -2.]_HestenesStiefel_grad

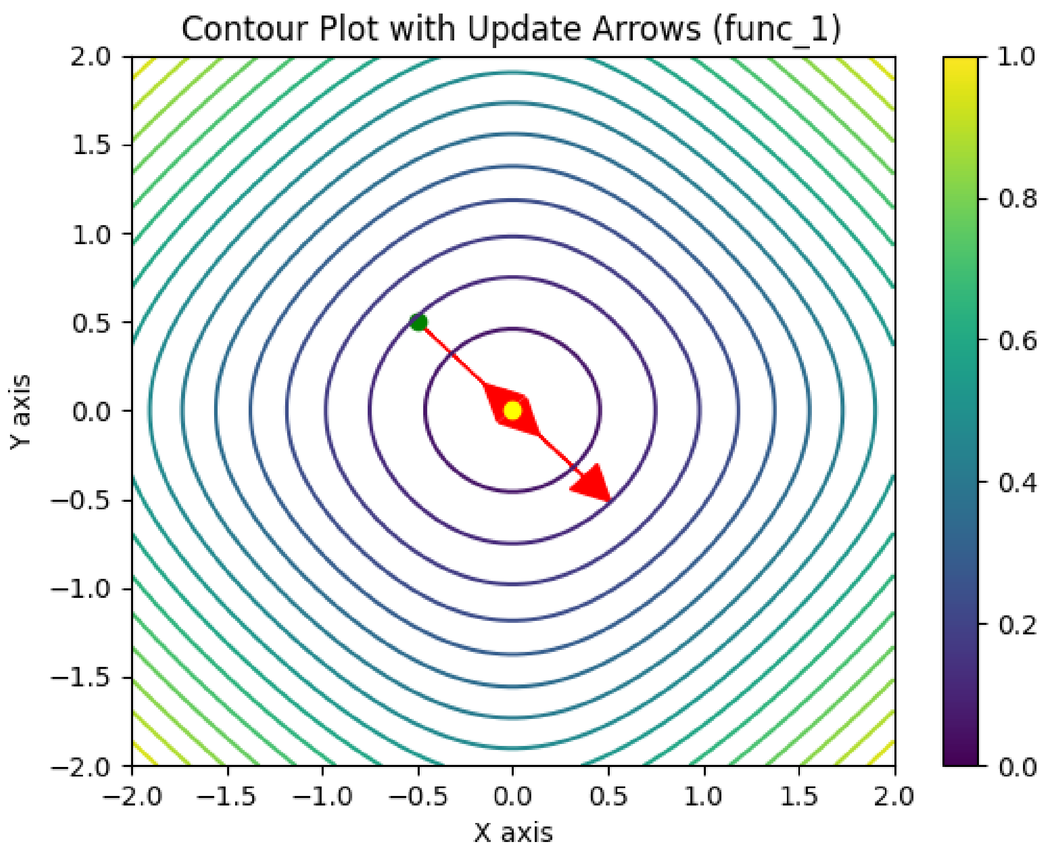

Figure 9: func_1_[-0.5 0.5]_DFP_cont

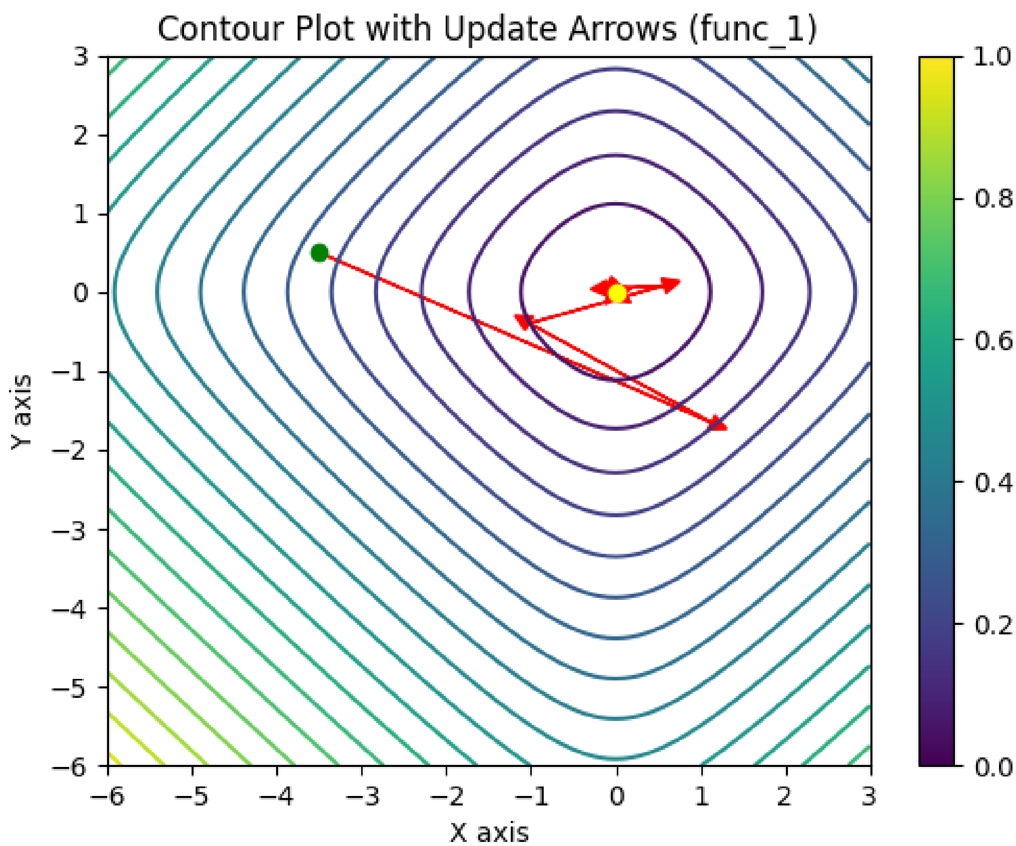

Figure 10: func_1_[-3.5 0.5]_BFGS_cont

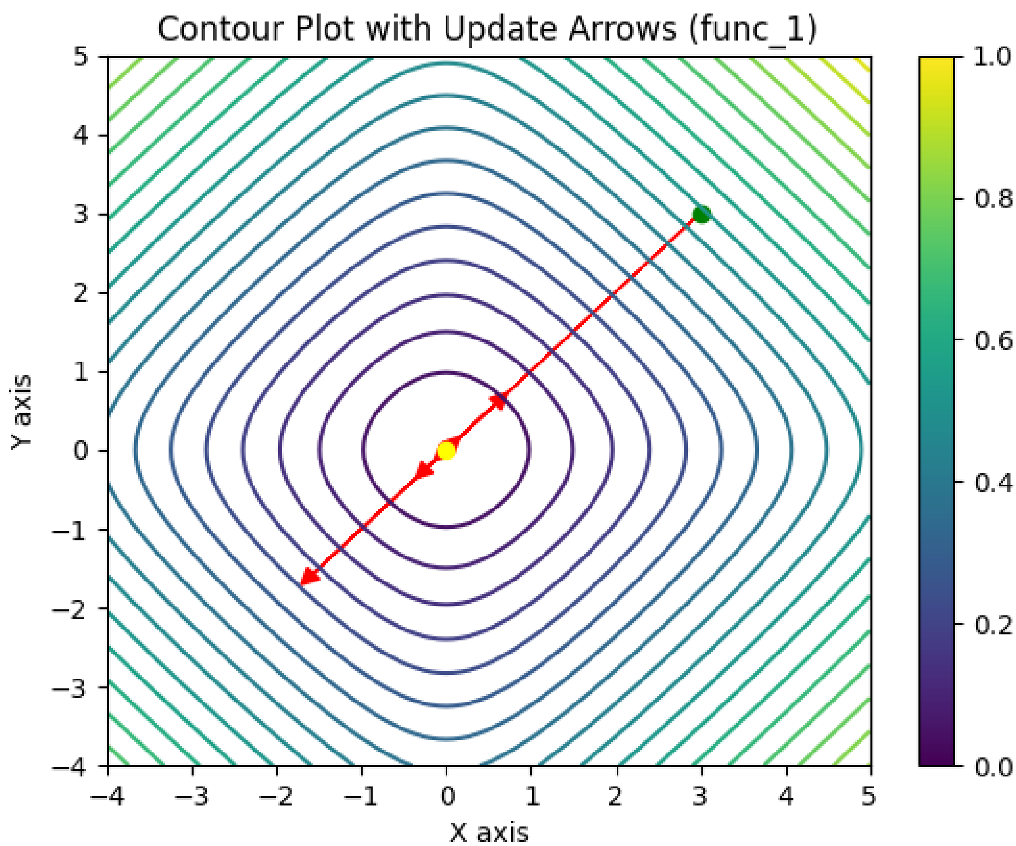

Figure 11: func_1_[3. 3.]_BFGS_cont

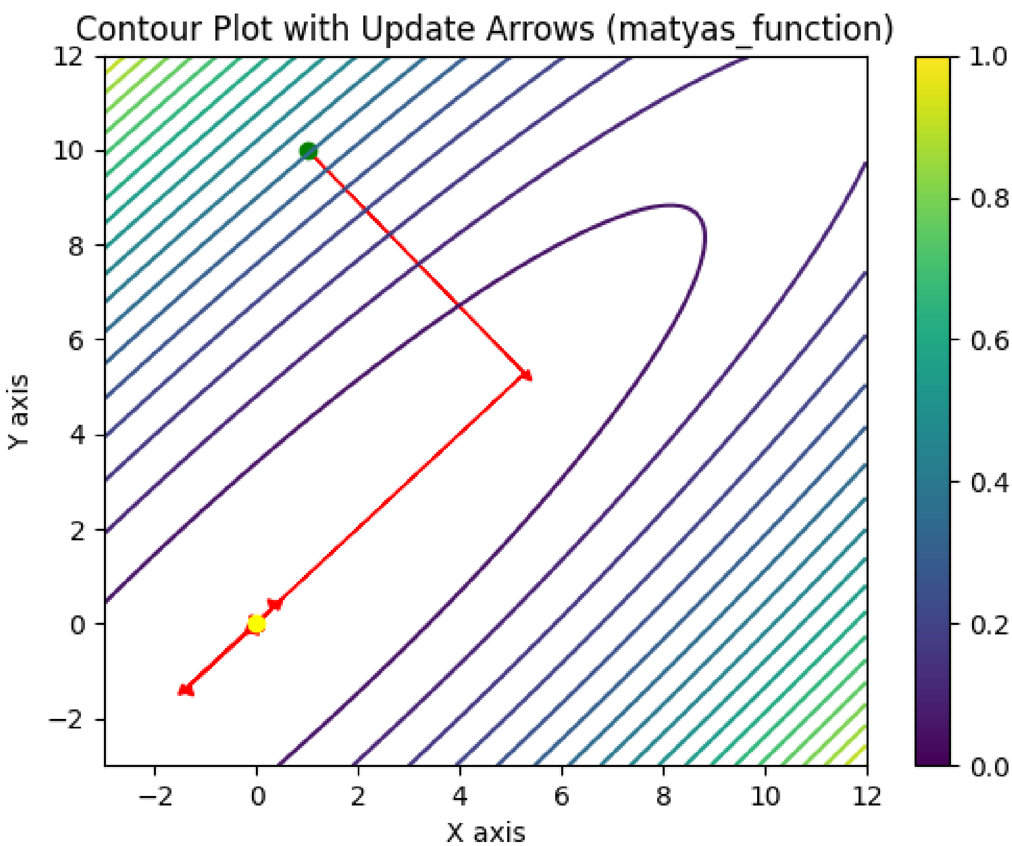

Figure 12: matyas_function_[1. 10.]_HestenesStiefel_cont

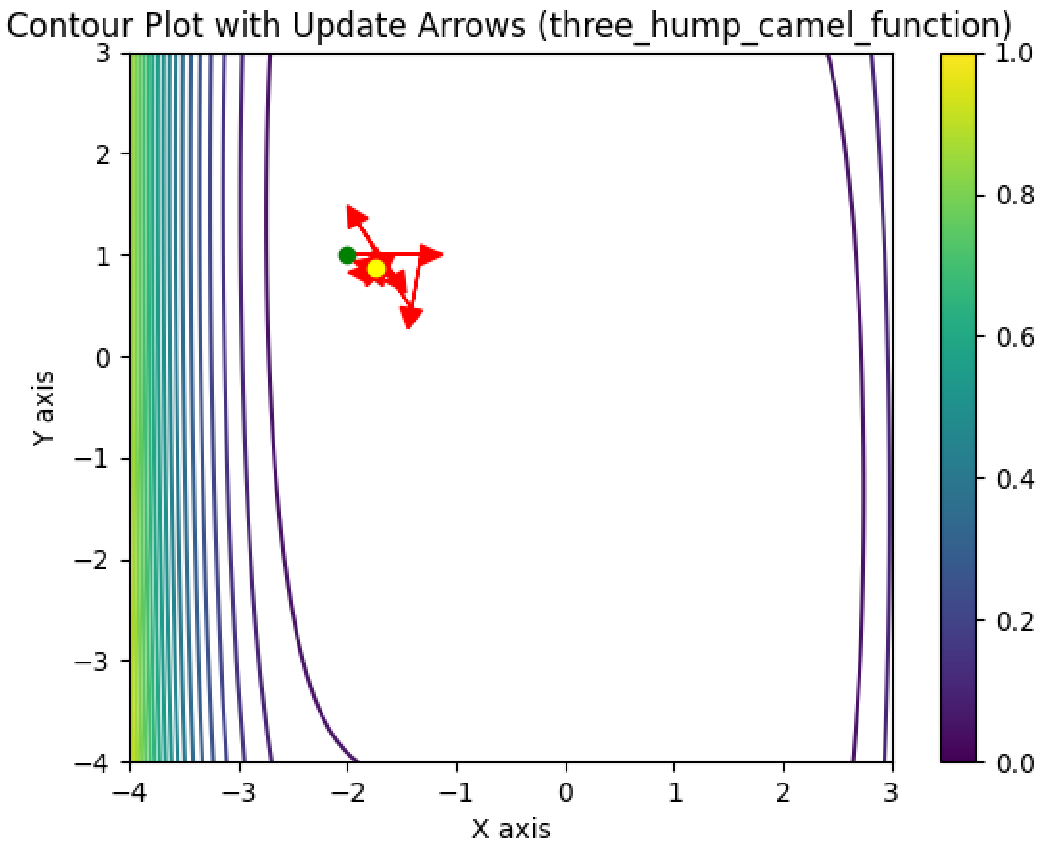

Figure 13: three_hump_camel_function_[-2. 1.]_DFP_cont