Assignment 2 Report

Name : Aman Raj

Roll No: 2021101064

1. Hessian and Jacobian of all the functions

Trid function

Hessian of Trid function :-

We have

$$f(x) = \sum_{i=1}^{d} (x_i - 1)^2 - \sum_{i=2}^{d} x_{i-1} x_i$$

So we get

$$\frac{\partial f(x)}{\partial x_i} = 2(x_i - 1) - (x_{i-1} + x_{i+1})$$

$$= 2x_i - 2 - x_{i-1} - x_{i+1}$$

and so we get

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \begin{cases} 2 & \text{if } i = j \\ -1 & \text{if } |i-j| = 1 \\ 0 & \text{otherwise.} \end{cases}$$

Thus the Hessian is given by the matrix

$$H = \nabla^2 f(x)$$

$$\cancel{=} A.$$

where H is $d \times d$ and we have

$$\boxed{H(i,j) = \begin{cases} 2 & \text{if } i = j \\ -1 & \text{if } |i-j| = 1 \\ 0 & \text{otherwise.} \end{cases}}$$

Three Hump Camel Function

We have

$$f(x) = 2x_1^2 - 1.05x_1^4 + \frac{x_1^6}{6} + x_1x_2 + x_2^2$$

So we get

$$\frac{\partial f}{\partial x_1} = 4x_1 - 4.2x_1^3 + x_1^5 + x_2$$

and $$\frac{\partial f}{\partial x_2} = x_1 + 2x_2.$$

So the Jacobian is given by

$$J(x_1, x_2) = \nabla f(x_1, x_2)$$

$$J(x_1, x_2) = \begin{bmatrix} 4x_1 - 4.2x_1^3 + x_1^5 + x_2 \\ x_1 + 2x_2 \end{bmatrix}$$

Now we have the Hessian defined by

$$H(x_1, x_2)_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}.$$

But $$\frac{\partial^2 f}{\partial x_1^2} = 4 - 12.6x_1^2 + 5x_1^4$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = 1, \quad \frac{\partial^2 f}{\partial x_2^2} = 2.$$

So, the Hessian is given by

$$H(x_1, x_2) = \begin{bmatrix} 5x_1^4 - 12.6x_1^2 + 4 & 1 \\ 1 & 2 \end{bmatrix}$$

Styblinski-Tang Function

We have.

$$f(x) = \frac{1}{2} \sum_{i=1}^{d} x_i^4 - 16x_i^2 + 5x_i$$

So we get

$$\frac{\partial f}{\partial x_i} = \frac{1}{2} (4x_i^3 - 32x_i + 5)$$

So, we get the Jacobian as

$$J = \nabla f(x)$$

where $J(i) = \frac{1}{2} (4x_i^3 - 32x_i + 5)$
$\forall i \in \{1, 2, 3, \ldots, d\}$. of the vector

J.

Similarly, we have

$$\frac{\partial^2 f}{\partial x_i^2} = 6x_i^2 - 16$$

and

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = 0 \quad \forall i \neq j.$$

So, the Hessian is given by

$$H(x) = \nabla^2 f(x)$$

where we have

$$H_{ij}(x) = \begin{cases} 6x_i^2 - 16 & \text{if } i=j \\ 0 & \text{otherwise.} \end{cases}$$

Rosenbrock Function

We have

$$f(x) = \sum_{i=1}^{d-1} \left[100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2\right]$$

So we have.

$$\frac{\partial f}{\partial x_i} = 200(x_{i+1} - x_i^2)(-2x_i) + 2(x_i - 1) + 200(x_i - (x_{i-1}^2))$$

when $i > 1$ and $i < d-$

If $i = 1$, we have

$$\frac{\partial f}{\partial x_1} = 200(x_2 - x_1^2) \cdot (-2x_1) + 2(x_1 - 1)$$

and if $i = d$, we have

$$\frac{\partial f}{\partial x_d} = 200(x_d - x_{d-1}^2)$$

So, ~~the~~ the Jacobian is given by:

So, we have

$$J = \nabla f(x)$$

where

$$J_i = \begin{cases} 200(x_{i+1} - x_i^2).(-2x_i) \\ \quad + 2(x_i - 1) \\ \quad + 200(x_i - x_{i-1}^2), \; i \in [2, \dots, d-1] \end{cases}$$

$$J_1 = 200(x_2 - x_1^2).(-2x_1) + 2(x_1 - 1)$$

$$J_d = 200(x_d - x_{d-1}^2).$$

Now we have to find Hessian. We have

$$\frac{\partial^2 f}{\partial x_i^2} = 1200\, x_i^2 + 202 - 400\, x_{i+1}$$

$$\frac{\partial^2 f}{\partial x_i \partial x_{i+1}} = -400 x_i$$

where $i \in [2, \dots, d-1]$

Now we have.

$$\frac{\partial^2 f}{\partial x_1^2} = 2 - 400 x_2 + 1200\, x_1^2$$

and

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = -400 x_1$$

Also,

$$\frac{\partial^2 f}{\partial x_d^2} = 200$$

and $$\frac{\partial^2 f}{\partial x_d \cdot \partial x_{d-1}} = -400x_{d-1}.$$

~~So, we~~ Also, we have $\frac{\partial^2 f}{\partial x_i \partial x_j} = 0$

where i,j are not covered above.

So, we have Hessian

$$H = \nabla^2 f(x) =$$

where $H_{ij}(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)$.

and all the values are calculated above.

Matyas Function

Rotated Hyper-Ellipsoid Function

2. Minima of All The Functions

Trid Function

For minima, we need.

$$\nabla f(x) = 0$$

$$\Rightarrow \quad \cancel{\nabla f} \; \frac{\partial f}{\partial x_i} = 0$$

$\Rightarrow$ We get

$$2x_i - 2 - x_{i-1} - x_{i+1} = 0 \quad - ①$$

$$\forall i \in \{2, \dots, d\}$$

and $\quad 2(x_1 - 1) = x_2. \quad - ②$

The solution for the above system is given by

$$x_i = \cancel{} \; i \cdot (d - i + 1)$$

Now, we find the Hessian:

$$\nabla^2 f(x_i) = \begin{bmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \dots & 0 \\ 0 & -1 & 2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & -1 & 2 \end{bmatrix}$$

This is positive definite, so we have

$$x = [x_i]_{i=1,\dots,d} \quad \text{as minima}$$

where $x_i = i(d - i + 1)$.

Three Hump Camel Function

For three hump camel, we have

$$\nabla f = 0$$

$$\Rightarrow \quad \frac{\partial f}{\partial x_1} = 0, \ \frac{\partial f}{\partial x_2} = 0$$

$$\Rightarrow \quad 4x_1 - 4.2x_1^3 + x_1^5 + x_2 = 0$$

and $\quad x_1 + 2x_2 = 0$

Solving above two equations, we get

$$x_1 = 0, \ \mp 1.747, \ \mp 1.07$$
$$x_2 = 0, \ \pm 0.874, \ \pm 0.53$$

Now we have Hessian

$$\nabla^2 f = \begin{bmatrix} 4 - 12.6x_1^2 + 5x_1^4 & 1 \\ 1 & 2 \end{bmatrix}$$

Now for (0,0), $\nabla^2 f$ is not positive semidefinite. However it is for $[\mp 1.747 \ \ \mp 0.874]$ and $[\mp 1.07 \ \ \pm 0.53]$

So we see that minimas are:

$$[\mp 1.747 \ \ \mp 0.874] \text{ and } [\mp 1.07 \ \ \pm 0.53]$$

Styblinksi Tang Function

inima of Styblinski Tang Function :-

e have for minima,

$$\Rightarrow \nabla f(x) = 0$$

∴ We get $\dfrac{\partial f}{\partial x_i} = 0$

$$\Rightarrow 4x_i^3 - 32x_i + 5 = 0$$

$\Rightarrow x_i = -2.903, 0.156, 2.747$ are the values for each $i \in \{1,2,\dots,d\}$.

ow, we have for minima,

$$\Rightarrow \nabla^2 f(x) \text{ - positive semidefinite.}$$

$$\Rightarrow \nabla^2 f = H_{d\times d}$$

where $H_{ij} = \begin{cases} (12x_i^2 - 32)/2 & \text{if } i=j \\ 0 & \text{if } i\neq j \end{cases}$

Since $\nabla^2 f$ is diagonal matrix, we see that the eigenvalues are given by

$$\lambda_i = \frac{12x_i^2 - 32}{2}.$$

Now, $\lambda_i > 0$ if $x_i = -2.903, 2.747$ and < 0 otherwise.

Thus, the minimas globally occurs at

$$\boxed{x_i = -2.903 \text{ or } 2.747 \quad \forall\ i \in \{1, \dots, d\}}$$

This ~~is~~ are the only minimas of the function.

Matyas Function

We have

$$\cancel{f(x) = \sum_{i=1}^{d-1}}$$

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

So we have

$$\frac{\partial f}{\partial x_1} = 0.52 x_1 - 0.48 x_2$$

and $$\frac{\partial f}{\partial x_2} = 0.52 x_2 - 0.48 x_1$$

Thus, the Jacobian is given by

$$J(x_1, x_2) = \begin{bmatrix} 0.52 x_1 - 0.48 x_2 \\ 0.52 x_2 - 0.48 x_1 \end{bmatrix}$$

Now we have

$$\frac{\partial^2 f}{\partial x_1^2} = 0.52$$

$$\frac{\partial^2 f}{\partial x_2^2} = 0.52$$

and we have $$\frac{\partial^2 f}{\partial x_1 \partial x_2} = -0.48$$

So, the Hessian is

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

For

$\Rightarrow$

$\Rightarrow$

To co

To ge

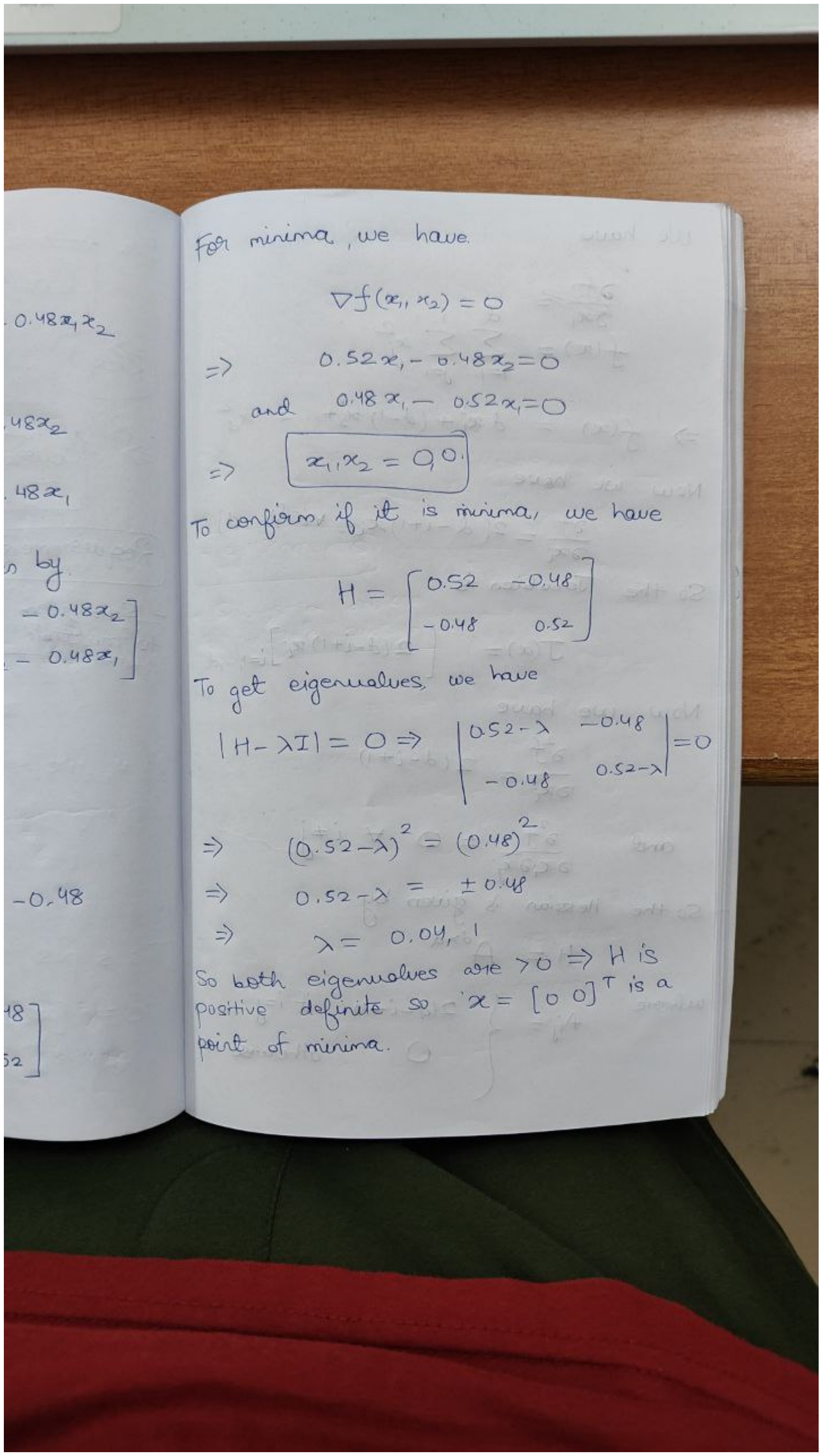
|H

$\Rightarrow$

$\Rightarrow$

$\Rightarrow$

So be
positiv
point



For minima, we have

$$\nabla f(x_1, x_2) = 0$$

$$\Rightarrow \quad 0.52x_1 - 0.48x_2 = 0$$

and $\quad 0.48x_1 - 0.52x_1 = 0$

$$\Rightarrow \quad \boxed{x_1, x_2 = 0, 0.}$$

To confirm if it is minima, we have

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

To get eigenvalues, we have

$$|H - \lambda I| = 0 \Rightarrow \begin{vmatrix} 0.52-\lambda & -0.48 \\ -0.48 & 0.52-\lambda \end{vmatrix} = 0$$

$$\Rightarrow \quad (0.52-\lambda)^2 = (0.48)^2$$

$$\Rightarrow \quad 0.52 - \lambda = \pm 0.48$$

$$\Rightarrow \quad \lambda = 0.04, 1$$

So both eigenvalues are $> 0 \Rightarrow H$ is positive definite so $x = [0\ 0]^T$ is a point of minima.

Rotated Hyper-Ellipsoid Function

We have

$$\cancel{\frac{\partial f}{\partial x_1} =}$$

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

$$\Rightarrow f(x) = d x_1^2 + (d-1) x_2^2 + \dots + x_d^2$$

Now we have

$$\frac{\partial f}{\partial x_i} = 2(d-i+1)x_i \quad \forall i \in \{1, \dots, d\}$$

So the Jacobian is

$$J(x) = \left[2(d-i+1) x_i \right]_{i=1,\dots,d}$$

Now we have

$$\frac{\partial^2 f}{\partial x_i^2} = 2(d-i+1)$$

and

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = 0 \quad \forall i \neq j$$

So the Hessian is given by

$$H = A^{d \times d}$$

where

$$A_{ij} = \begin{cases} 2(d-i+1) & \text{if } i=j \\ 0 & \text{otherwise} \end{cases}$$

For minim

$\Rightarrow$

$\Rightarrow$

To check

Hessian

Since H

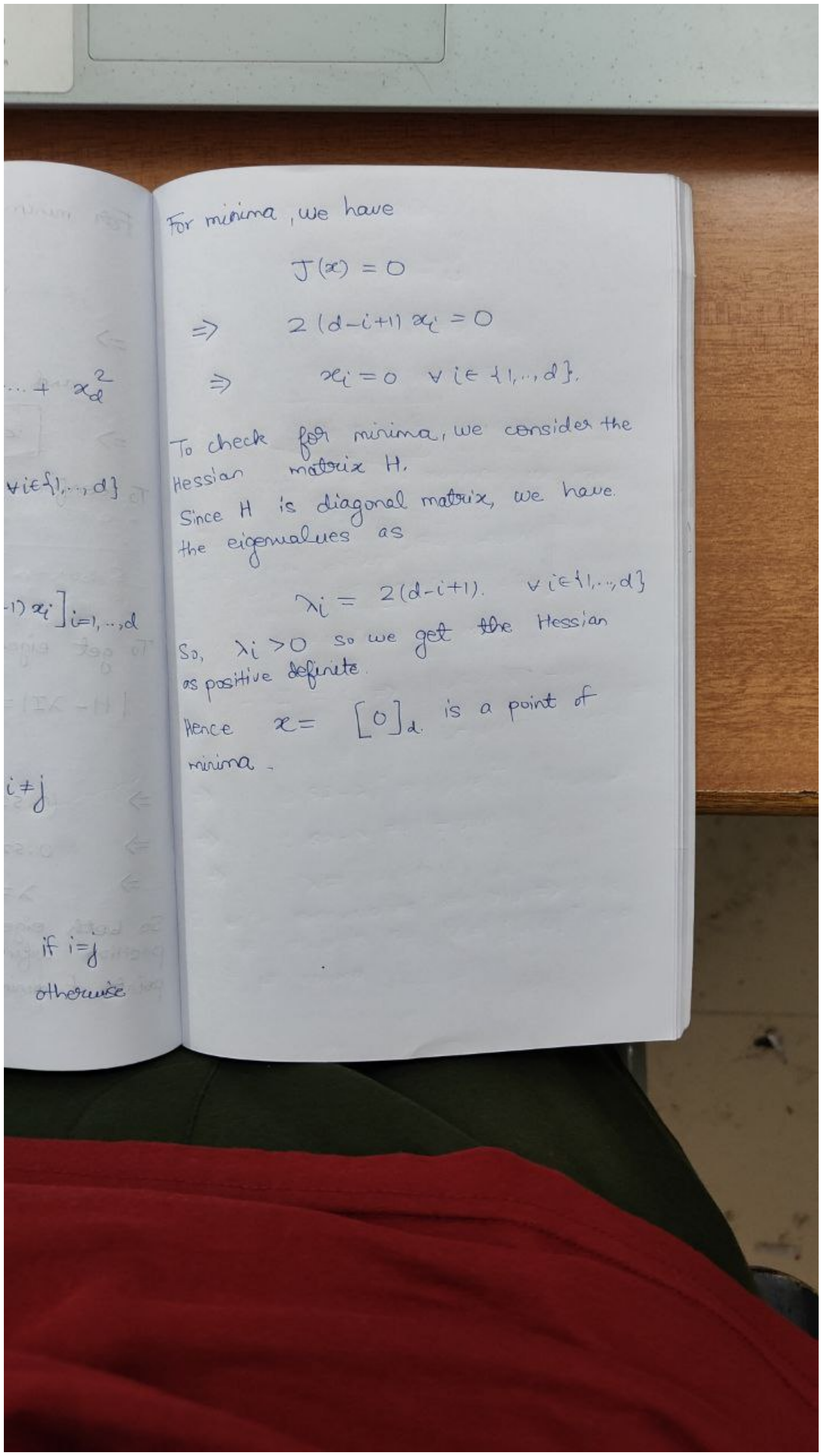
the eig

So, λ_i

as positiv

Hence

minima



For minima, we have

$$J(x) = 0$$

$$\Rightarrow \quad 2(d-i+1)\, x_i = 0$$

$$\Rightarrow \quad x_i = 0 \quad \forall i \in \{1, .., d\}.$$

To check for minima, we consider the Hessian matrix H.
Since H is diagonal matrix, we have the eigenvalues as

$$\lambda_i = 2(d-i+1). \quad \forall i \in \{1, .., d\}$$

So, $\lambda_i > 0$ so we get the Hessian as positive definite.
Hence $x = [0]_d$ is a point of minima.

3. Algorithms Which Failed To Converge & Circumstances:

Hestenes-Stiefel Method:

- Styblinski Tang Function failed to converge in most of the cases in Hestenes Stiefel, except for [3, -3, 3, -3]
- Square Root function failed to converge in most of the cases, except the [=3.5, 0.5]

Polak Ribiere Method:

- Converged on all values

Fletcher - Reeves Method:

- Failed to converge on Rosenbrock function at [2, 2, 2, -2] and [3, 3, 3, 3]

Symmetric Rank One Correction:

- Failed to converge on Rosenbrock function.

DFP

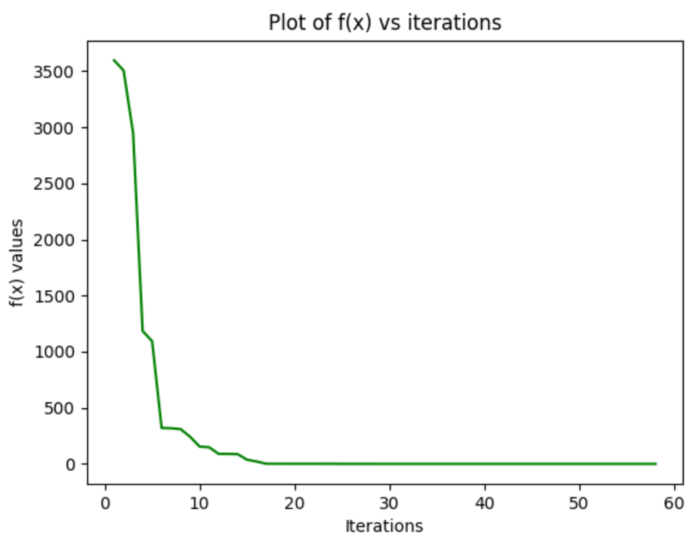
- Failed to converge on Rosenbrock function at [2, -2, -2, 2]

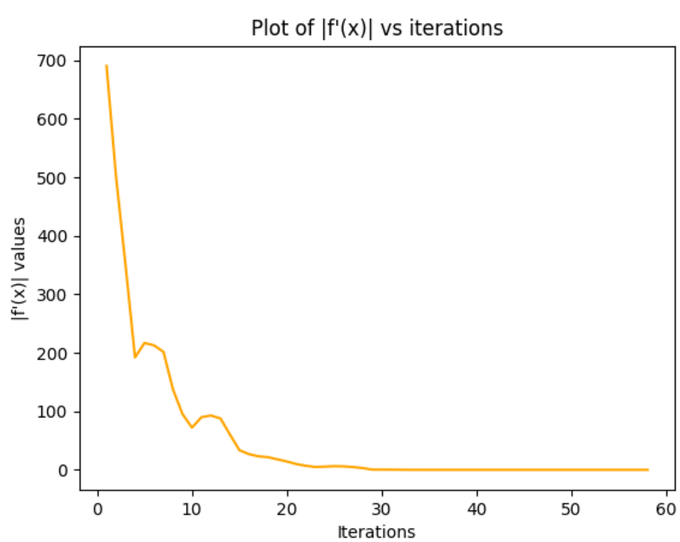
BFGS

- Failed to converge on Rosenbrock function.

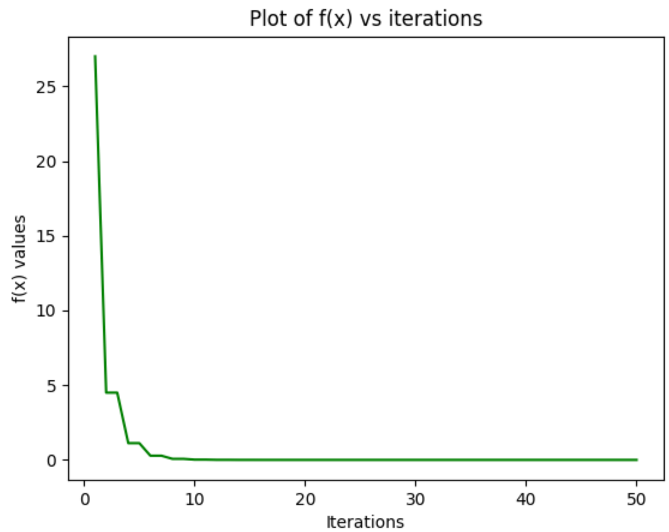
4. Plots of f(x) vs Iterations, |f′(x)| vs iterations

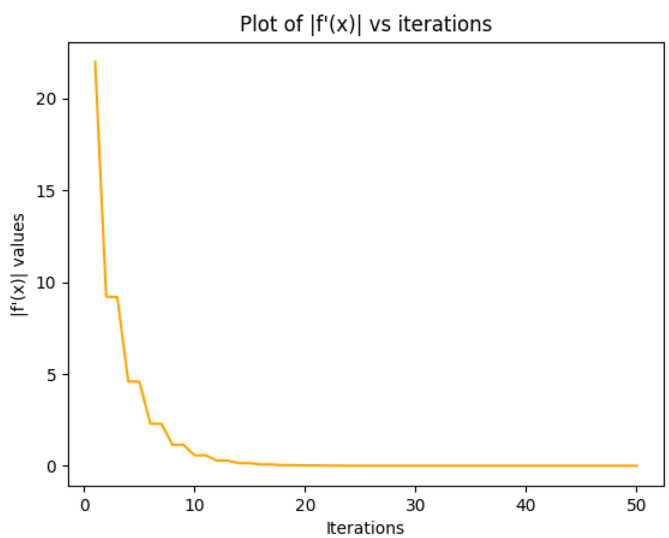
Hestenes-Stiefel Method [Hyperellipsoid curve]



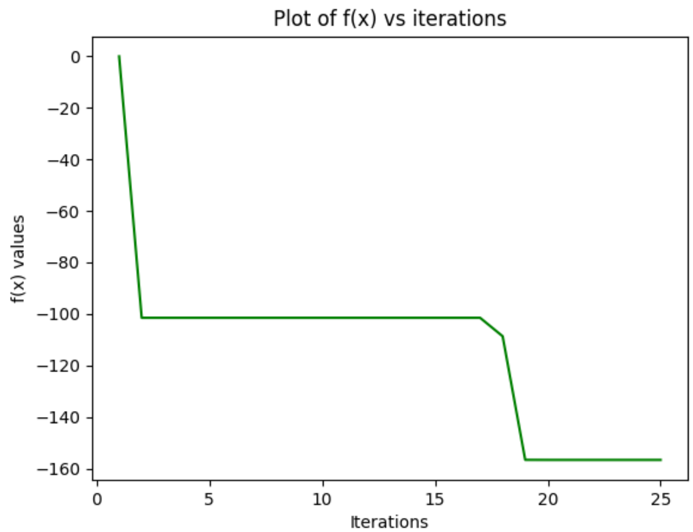


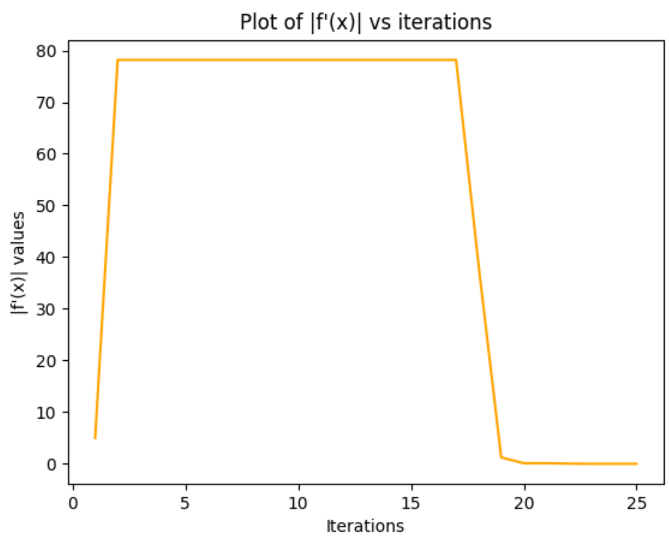
Polak Ribiere Method [Hyperellipsoid curve]



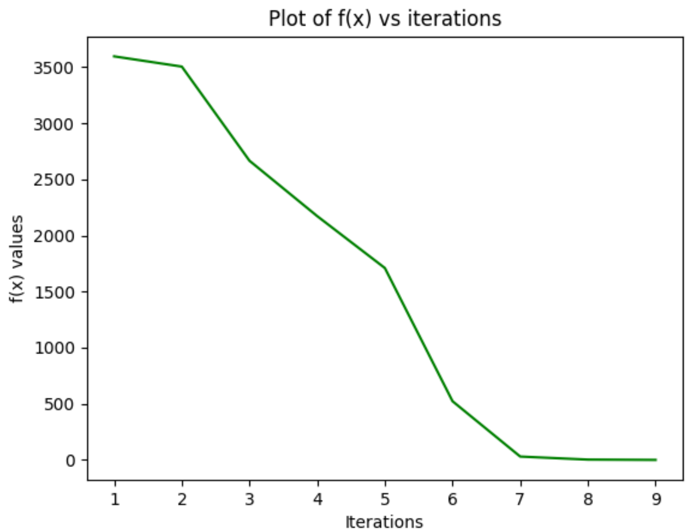


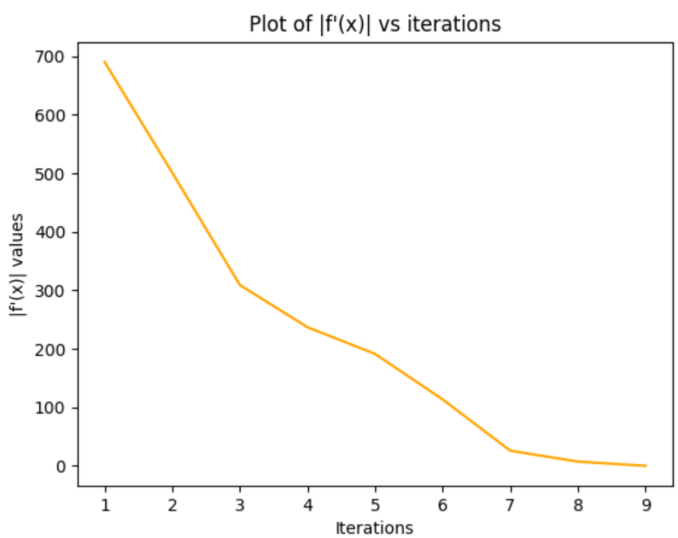
Fletcher Reeves Method [Styblinksi Tang at 0,0,0,0]



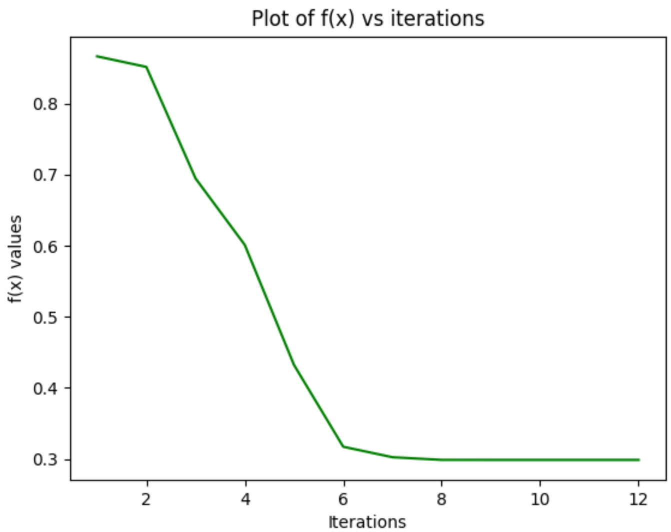


Symmetric Rank One Method [Hyperellipsoid curve]

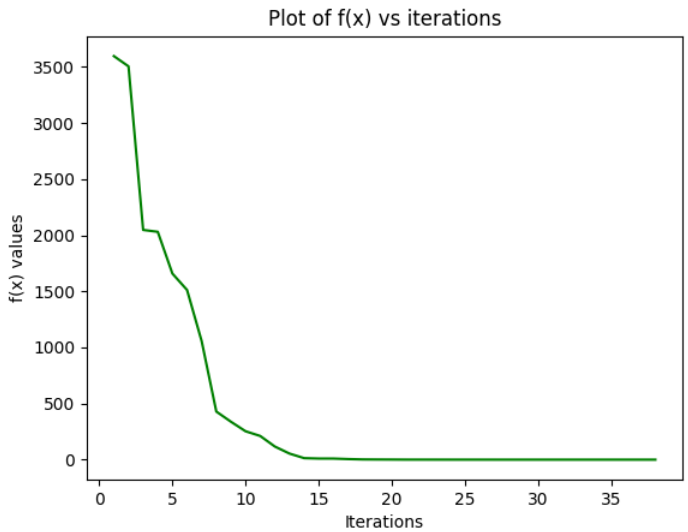


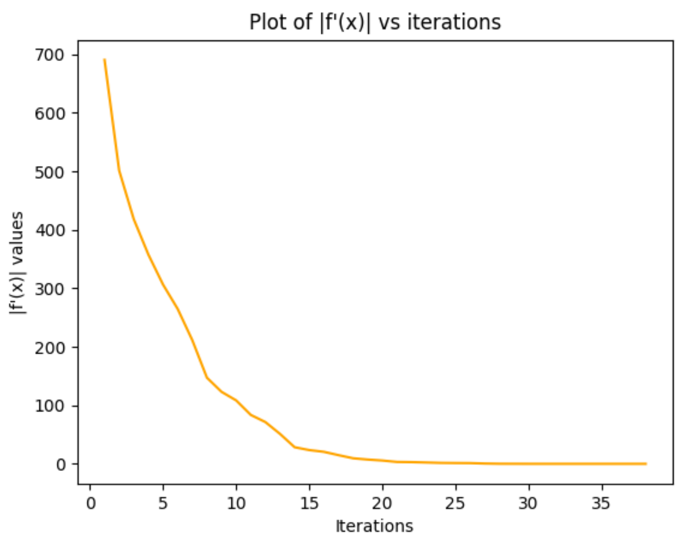


DFP Method [Three hump camel function]



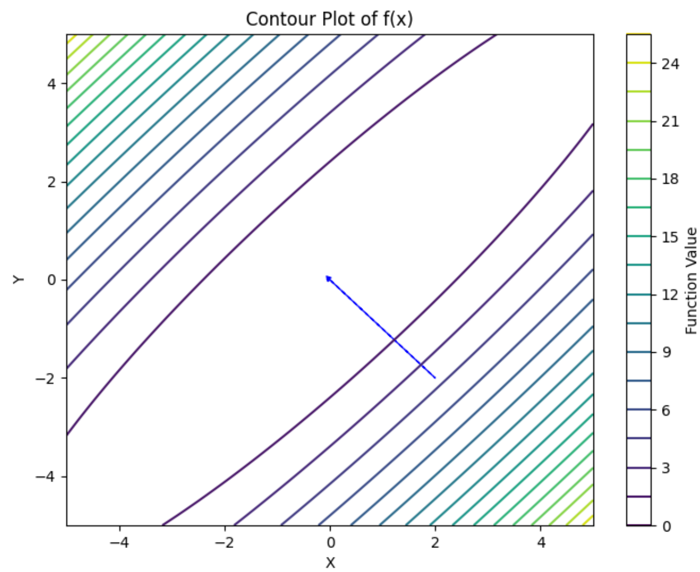
BFGS Method [Hyperellipsoid function]



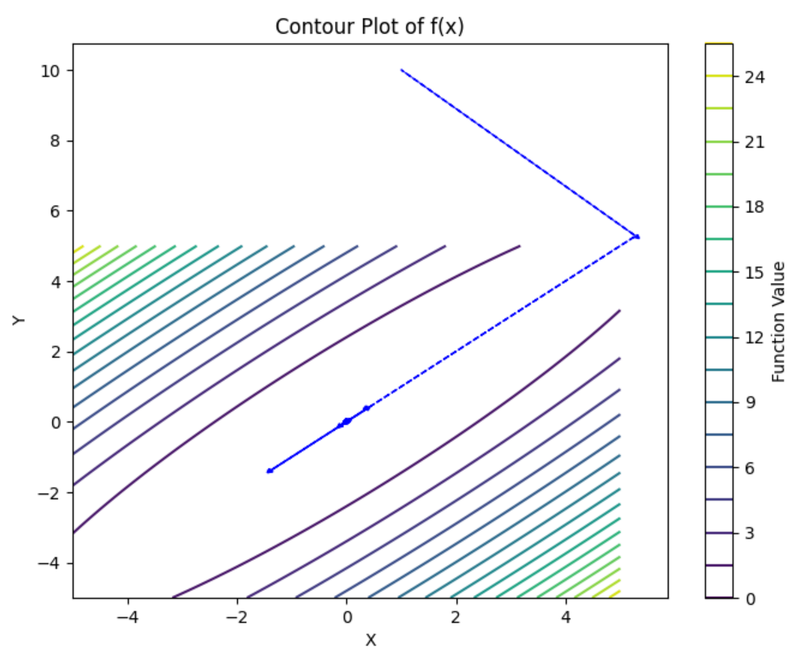


4. Contour Plot of f(x)

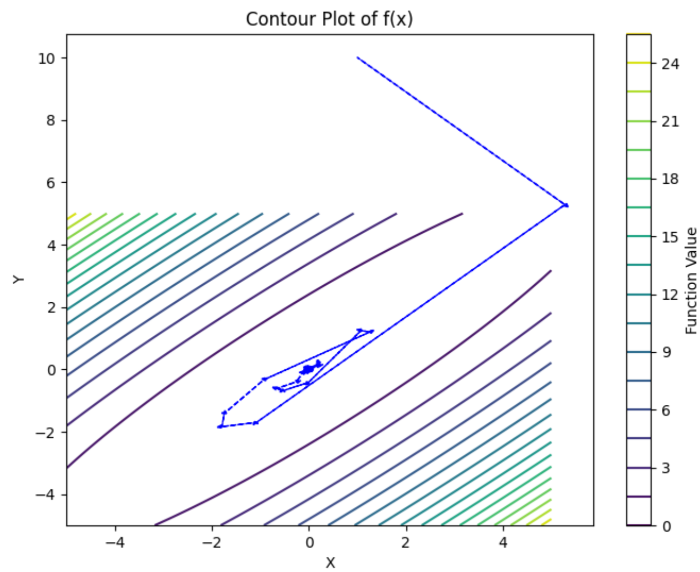
Hestenes-Stiefel Method [Matyas curve at 2,-2]



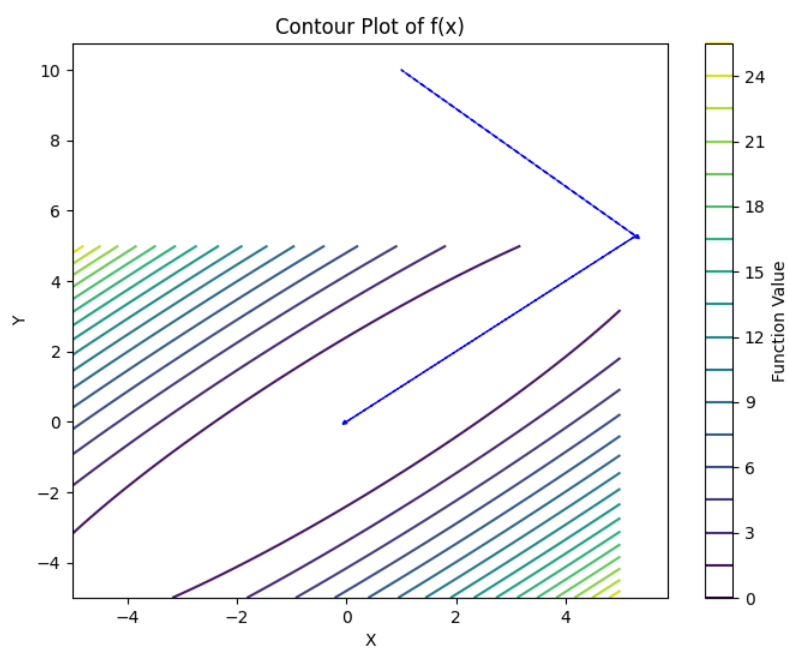
Polak Ribiere Method [Matyas function at 1,10]



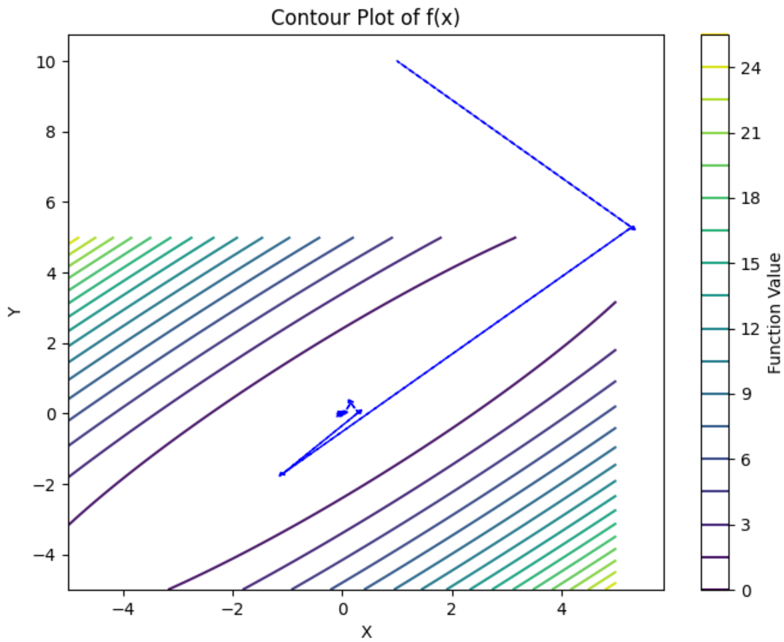
Fletcher Reeves Method [Matyas curve at 1,10]



Symmetric Rank One Method



DFP Method [Matyas function at 1,10]



BFGS Method [Matyas function at 1,10]