Report
Assignment 1
Name: Udrasht Pal
Roll No.: 2022201020

1) **Trid Function**

$$f(X) = \sum_{i=1}^{d} (x_i - 1)^2 - \sum_{i=2}^{d} x_{i-1}x_i$$
$$X = [x_1, x_2, \dots x_d]$$

$$f(x) = (x_1 - 1)^2 + (x_2 - 1)^2 + \cdots + (x_d - 1)^2 - [x_1x_2 + x_2x_3 + \dots\dots + x_{d-1}x_d)$$
$$\frac{\partial f}{\partial x_1} = 2(x_1 - 1) - x_2$$
$$\frac{\partial f}{\partial x_2} = 2(x_2 - 1) - (x_1 + x_3)$$
$$\frac{\partial f}{\partial x_3} = 2(x_3 - 1) - (x_2 + x_4)$$

$$at\ i = 1 \quad \frac{\partial f}{\partial x_1} = 2(x_1 - 1) - x_2$$

$$\text{i} \geq 2 \text{ and } i < \text{d} \quad \frac{\partial f}{\partial x_{\text{i}}} = 2(x_i - 1) - (x_{i-1} + x_{i+1})$$

$$\frac{df}{dx_{\text{i}}} = \begin{cases} 2(x_i - 1) - x_{i+1} & \text{at } i = 1, \\ 2(x_i - 1) - (x_{i-1} + x_{i+1}) & \text{at } i > 1 \text{ and } i < d, \\ 2(x_i - 1) - x_{i-1} & \text{at } i = d \end{cases}$$

$$\frac{\partial^2 f}{\partial x_{\text{i}}^2} = \begin{cases} 2 & \text{at } i = 1 \\ 2 & \text{at } i > 1 \text{ and } i < d \\ 2 & \text{at } i = d, \end{cases}$$
$$\frac{\partial f}{\partial x_{\text{i}} x_i} = \begin{cases} -1 & \text{at } i = 1 \text{ and } j = i + 1 \\ -1 & \text{at } i \in [2, d) \text{ and } (j = i - 1 \text{ or } j = i + 1) \\ -1 & i = d \text{ and } j = i - 1 \end{cases}$$
$$0\ otherwise$$

$$H(X) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_2 x_1} & \cdots & \cdots & \frac{\partial^2 f}{\partial x_\partial x_1} \\ \cdots & \cdots & \cdots & \cdots & - \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 f}{\partial x_1\, \partial x_d} & \frac{\partial^2 f}{\partial x_1 x_d} & \cdots & \cdots & \frac{\partial^2 f}{\partial x_d^2} \end{bmatrix}$$
$$H(X) = \begin{bmatrix} 2 & -1 & 0 & 0 & \cdots & \cdots & 0 \\ -1 & 2 & -1 & 0 & \cdots & \cdots & 0 \\ 0 & -1 & 2 & -1 & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & \cdots & 2 \end{bmatrix}$$

Find Stationary point,
$\nabla f(\bar{x}) = 0$

$$\left[\frac{\partial f}{\partial r_1}, \frac{\partial f}{\partial r_L} \dots \frac{\partial f}{\partial x_a}\right]^{\top} = [0,0 \dots 0]^{\top}$$

i=1

$$2(x_1 - 1) - x_2 = 0 \qquad - \qquad (1)$$

$i = 2$ to $i = d - 1$

$$2(x_i - 1) - x_i + 1 - x_{i-1} = 0$$

summing of all such $(d - 1 - 2 + 1) = d - 2$ equations

$$\sum_{i=2}^{d-1} 2(x_i - 1) - \sum_{i=2}^{d-1} x_{i+1} - \sum_{i=2}^{d-1} x_{i-1} = 0$$

$$2\sum_{i=2}^{d-1} x_i - \sum_{i=3}^{d} x_i - \sum_{i=1}^{d-2} x_i - 2(d-2) = 0 \qquad - \qquad (2)$$

at $i = d$

$$2(x_d - 1) - x_{d-1} = 0 \qquad - \qquad (3)$$

Add eq. (2) & (3)

$$2\sum_{i=2}^{d-1} x_i + 2x_d - \sum_{i=3}^{d} x_i - \sum_{i=1}^{d-2} x_i - x_{d-1} - 2 - 2(d-2) = 0$$

$$\Rightarrow 2\sum_{i=2}^{d} x_i - \sum_{i=3}^{d} x_i - \sum_{i=1}^{d-1} x_i - 2(d-1) = 0$$

$$\Rightarrow 2x_2 + 2x_d + 2\sum_{i=1}^{d-x_i} x_i - 2\sum_{i=3}^{d-1} x_i - (x_1 + x_2 + x_d) - 2(d-1) = 0.$$

$$\Rightarrow 2(x_2 + x_d) - (x_1 + x_2 + x_d) = 2(d-1)$$

put $x_2 = 2(x_1 - 1)$ using equation (1)

$$= 2(2x_1 - 2) + 2x_d - x_1 - x_2 - x_d = 2(d-1)$$

$$= 3x_1 - 4 + x_d - 2x_1 + 2 = 2d - 2$$

$$x_1 + x_d = 2d$$

$$x_d = 2d - x_1 \qquad - \qquad (4)$$

$$\nabla f = \begin{bmatrix} 2(x_1 - 1) - x_2 \\ 2(x_2 - 1) - x_1 - x_3 \\ 2(x_1 - 1) - x_2 - x_4 \\ \vdots \\ 2(x_d - 1) - x_{d-1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$x_2 = 2(x_1 - 1)$
$x_3 = 2(x_2 - 1) - x_1$
$x_3 = 4(x_1 - 1) - 2 - x_1$
$x_3 = 3x_1 - 6$
$x_3 = 3(x_1 - 2)$

Similarly,

$x_4 = 4(x_1 - 3)$
$x_i = i(x_i - (i - 1))$
$x_d = d(x_1 - (d - 1))$

Now by eq. (4) we have

$x_d = 2d - x_1$

Using (3) & (4)

$$2d - x_1 = d(x_1 - (d - 1))$$
$$2d - x_1 = dx_1 - d^2 + d$$

$$d^2 - d - dx_1 + 2d - x_1 = 0$$
$$d^2 + d - dx_1 - x_1 = 0$$

$$d^2 + d - x_1(d + 1) = 0$$
$$d(d + 1) - x_1(d + 1) = 0$$
$$(d + 1)(d - x_1) = 0$$

$$\Rightarrow x_1 = d \quad \text{at} \quad d \neq -1$$

We already computed x_2 ... x_d in terms of x_1.

$$x_i = i(x_1 - (i - 1))$$

put $x_1 = d$, we gat

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

Therefore , the stationary point for this function is -

$$\bar{x} \quad = [x_1 \quad x_2 \quad \cdots \quad x_{\mathrm{d}}]^{\mathsf{T}}$$

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

Now we need to prove that Hessian is +ve definite.

$$H(X) = \begin{bmatrix} 2 & -1 & 0 & 0 & 0 & \cdots & & 0 \\ -1 & 2 & -1 & 0 & 0 & \cdots & & 0 \\ 0 & -1 & 2 & -1 & 0 & \cdots & & 0 \\ \vdots & & & & & & & \vdots \\ 0 & 0 & 0 & \cdots & \cdots & 0 & -1 & 2 \end{bmatrix}$$

$$\text{Trace (H)} = 2 \times d > 0$$

We need to prove that |H(X)|>0
We prove this by induction.
Claim: Let $n \times n$ be the size of Hessian. If Δ_H denote determinant of $n \times n$ Hessian, the $\Delta_n > 0$
and $\Delta_n > \Delta_{n-1} > \Delta_{n-2} \dots \Delta_1$

Prove: Base care.

$$H_1 = [2]$$
$$\Delta_1 > 0$$
$$H_2 = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$
$$\Delta_2 = 3 > 0$$

$\Delta H_2 > 0$ and $\Delta H_2 > \Delta H_1 \cdots$ True.
Induction hypothesis is:

$$\Delta H_n > 0$$
$$\Delta H_n > \Delta H_{n+1} > \Delta H_{n-1} \cdots \Delta H_1$$

To prove:

$$\Delta H_{n+1} > 0$$
$$\Delta H_{n+1} > \Delta H_n > \Delta H_{n-1} \dots > \Delta H_1$$

Consider the hessian of size $(n+1) \times (n+1)$.

$$\Delta H_{n+1} = \begin{bmatrix} 2 & -1 & 0 & 0 & 0 & \cdots & & 0 \\ -1 & 2 & -1 & 0 & 0 & \cdots & & 0 \\ 0 & -1 & 2 & -1 & 0 & \cdots & & 0 \\ \vdots & & & & & & & \vdots \\ 0 & 0 & 0 & \cdots & \cdots & 0 & -1 & 2 \end{bmatrix}$$

size $(n+1) \times (n+1)$.

Find ΔH_{n+1} by expending along R1

$$\Delta H_{n+1} = 2 * \begin{bmatrix} 2 & -1 & 0 & 0 & 0 & \cdots & & 0 \\ -1 & 2 & -1 & 0 & 0 & \cdots & & 0 \\ 0 & -1 & 2 & -1 & 0 & \cdots & & 0 \\ \vdots & & & & & & & \vdots \\ 0 & 0 & 0 & \cdots & \cdots & 0 & -1 & 2 \end{bmatrix} + (-1) * \begin{bmatrix} -1 & -1 & 0 & 0 & \cdots & & 0 \\ 0 & 2 & -1 & 0 & \cdots & & 0 \\ 0 & -1 & 2 & 0 & \cdots & & 0 \\ \vdots & & & & & & \vdots \\ 0 & 0 & \cdots & \cdots & 0 & -1 & 2 \end{bmatrix}$$

size $(n-1) \times (n-1)$.

$$= 2 \times \Delta H_n + 1 \times |A\ |$$

$$|A| = \begin{bmatrix} -1 & -1 & 0 & 0 & \cdots & & 0 \\ 0 & 2 & -1 & 0 & \cdots & & 0 \\ 0 & -1 & 2 & 0 & \cdots & & 0 \\ \vdots & & & & & & \vdots \\ 0 & 0 & \cdots & \cdots & 0 & -1 & 2 \end{bmatrix}$$

$$= (-1) * \begin{bmatrix} 2 & -1 & 0 & 0 & \cdots & & 0 \\ -1 & 2 & -1 & 0 & \cdots & & 0 \\ \vdots & & & & & & \vdots \\ 0 & 0 & \cdots & \cdots & 0 & -1 & 2 \end{bmatrix} + 1 * \begin{bmatrix} 0 & -1 & 0 & 0 & \cdots & & 0 \\ 0 & 2 & -1 & 0 & \cdots & & 0 \\ \vdots & & & & & & \vdots \\ 0 & 0 & \cdots & \cdots & 0 & -1 & 2 \end{bmatrix}$$

size $(n-2) \times (n-2)$.

$$|A| = -\Delta H_{n-2} + 0$$

$$\Delta H_{n+1} = 2 \times \Delta H_n - \Delta H_{n-2}$$

$$\Delta H_{n+1} = \Delta H_n + (\Delta H_n - \Delta H_{n-2)}$$

By hypothesis

$$\Delta H_n > 0$$
$$\Delta H_n > \Delta H_{n+1} > \Delta H_{n-1} \cdots \Delta H_1$$

$$\Delta H_{n+1} > \Delta H_n$$

Hence prove

Hessian of tried function has +ve trace and +ve determinant

So Hessian is +ve definite

For d=2

$$x_1 = 1(2+1-1) = 2$$

$$x_2 = 2(2+1-2) = 2$$

$\boldsymbol{f(X)}$ is minima at

$$(\boldsymbol{x_1}, \boldsymbol{x_2}) = (\mathbf{2}, \mathbf{2}) \quad \textbf{global minima}$$

2) **Three Hump Camel**

$$f(x) = 2x_1^2 - 1.05x_1^4 + \frac{x_1^6}{6} + x_1x_2 + x_2^2$$
$$\cdot \nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \end{bmatrix}^{\top}$$
$$\frac{\partial f}{\partial x_1} = 4x_1 - 4.2x_1^3 + \frac{6x_1^5}{6} + x_1 = 4x_1 - 4 \cdot 2x_1^3 + x_1^5 + x_2$$

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

$\nabla f(\mathrm{X}) = [4x_1 - 4.2x_1^3 + x_1^5 + x_2 \quad , \quad x_1 + 2x_2]^{\top} =$ Jacobian

$$H(X) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1{}^2} & \frac{\partial^2 f}{\partial x_1 x_2} \\ \frac{\partial^2 f}{\partial x_2 x_1} & \frac{\partial^2 f}{\partial x_2{}^2} \end{bmatrix}$$

$$\frac{\partial f}{\partial x_1^2} = 4 - 12 \cdot 6x_1^2 + 5x_1^4, \quad \frac{\partial f}{\partial x_1 r_2} = 1$$
$$\frac{\partial f}{\partial x_1 x_1} = 1 \qquad , \frac{\partial f}{\partial x_2^2} = 2$$

$$H(x) = \nabla^2 f(x) = \begin{bmatrix} 4 - 12 \cdot 6x_1{}^2 + 5x_1^4 & 1 \\ 1 & 2 \end{bmatrix}$$

Find Stationary point.

$$\begin{bmatrix} 4x_1 - 4 \cdot 2x_1{}^3 + x_1{}^5 + x_2 \\ x_1 + 2x_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$4x_1 - 4 \cdot 2x_1{}^3 + x_1{}^5 + x_2 = 0 \quad - \quad (1)$$
$$x_1 = -2x_2 \quad - \quad (2)$$

but the value of from x_1 from equation (2) to (1)

$$4(-2x_2) - 4.2(-2x_2)^3 + (-2x_2)^5 + x_2 = 0$$
$$-8x_2 + 4.2 \times 8x_2^3 - 32x_2^5 + x_2 = 0$$
$$-32x_2^5 + 4.2 \times 8x_2^3 - 7x_2 = 0$$
$$x_2(-32x_2^4 + 4.2 \times 8x_2^2 - 7) = 0$$
$$x_2 = 0, \text{ at } x_2 = 0 - x_1 = 0$$
$$0 \text{ or}$$
$$-32x_2^4 + 4.2 \times 8x_2^2 - 7 = 0$$

for simple simplicity put $x_2^2 = t$

$$-32t^2 + 4.2 \times 8t - 7 = 0$$
$$32t^2 - 33.6t + 7 = 0$$
$$\text{apply } \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$
$$= \frac{33.6 \pm \sqrt{(33.6)^2 - 4 \times 7 \times 32}}{2 * 32}$$

$$t = \frac{33.6 \pm \sqrt{232.96}}{64}$$
$$t = 0.7634 \text{ or } t = 0.2865$$
$$t = x_2^2$$
$$x_2^2 = 0.7634$$
$$x_2 = \pm\sqrt{0.7634}$$
$$x_2 = \pm 0.8737 \text{ or } x_2 = \pm 0.5352$$

stationary points

at $x_2 = 0.8734$	at $x_2 = -0.8734$	at $x_2 = 0.5352$	at $x_2 = -0.5352$	at $x_2 = 0$
$x_1 = -1.7474$	$x_1 = 1.7474$	$x_1 = -1.0704$	$x_1 = 1.0704$	$x_1 = 0$

We need to check at which points the function $F(X)$ is minimum. For that, we have to prove that $H(X)$ is semi-positive definite at that point.

$at\ (x_1, x_2) = (0,0)$
$H(X) = \begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix}$ **It is semi $+ve$ definite because Trace(H(X)) > 0 and |H(X)| > 0**

at $(x_1, x_2) = (-1.7474\ , 0.8735)$

$H(X) = \begin{bmatrix} 12.1435 & 1 \\ 1 & 2 \end{bmatrix}$ **It is semi $+ve$ definite because Trace(H(X)) > 0 and |H(X)| > 0**

at $(x_1, x_2) = (1.7474\ , -0.8735)$

$H(X) = \begin{bmatrix} 12.1435 & 1 \\ 1 & 2 \end{bmatrix}$ **It is semi $+ve$ definite because Trace(H(X)) > 0 and |H(X)| > 0**

at $(x_1, x_2) = (1.0704, -0.5352)$

$H(X) = \begin{bmatrix} -3.872 & 1 \\ 1 & 2 \end{bmatrix}$ **It is not semi $+ve$ definite because Trace(H(X)) < 0 and |H(X)| < 0**

at $(x_1, x_2) = (-1.0704,\ 0.5352)$

$H(X) = \begin{bmatrix} -3.872 & 1 \\ 1 & 2 \end{bmatrix}$ **It is not semi $+ve$ definite because Trace(H(X)) < 0 and |H(X)| < 0**

Three Hump Camel
$f(X)$ is minima at

$$(x_1, x_2) = (0, 0), \qquad (-1.7474, 0.8734), \\ (1.7474, -0.8737)$$

at $(x_1, x_2) = (0, 0)$

$$f(X) = 0$$

at $(x_1, x_2) = (-1.7474, 0.8734)$

$$f(X) = 0.2986$$

at $(x_1, x_2) = (1.7474, -0.8737$

$$f(X) = 0.2986$$

Global Minima $(x_1, x_2) = (0, 0)$

3) **Styblinski-Tang Function**

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

$$\frac{\partial f}{\partial x_1} = \frac{1}{2}(4x_1^3 - 32x_1 + 5).$$

$$\frac{\partial f}{\partial x_2} = \frac{1}{2}(4x_2^3 - 32x_2 + 5)$$

$$\vdots$$

$$\frac{\partial f}{\partial x_\mathrm{d}} = \frac{1}{2}(4x_d^3 - 32x_d + 5)$$

$$J = \nabla f = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3} \cdots \frac{\partial f}{\partial x_\mathrm{d}}\right]^\top$$

$$\mathrm{J} = \begin{bmatrix} \frac{1}{2}(4x_1^3 - 32x_1 + 5) \\ \frac{1}{2}(4x_2^3 - 32x_2 + 5) \\ \vdots \\ \frac{1}{2}(4x_\mathrm{d}^3 - 32x_\mathrm{d} + 5) \end{bmatrix}$$

$$H(X) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_2 x_1} & \cdots & \cdots & \frac{\partial^2 f}{\partial x_\partial x_1} \\ \cdots & \cdots & \cdots & \cdots & - \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 f}{\partial x_1\,\partial x_d} & \frac{\partial^2 f}{\partial x_1 x_d} & \cdots & \cdots & \frac{\partial^2 f}{\partial x_d^2} \end{bmatrix}$$

$$H(\mathrm{X}) = \begin{bmatrix} \frac{1}{2}(12x_1^2 - 32) & 0 & \cdots & \cdots & 0 \\ 0 & \frac{1}{2}(12x_1^2 - 32) & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \cdots & \frac{1}{2}(12x_1^2 - 32) \end{bmatrix}$$

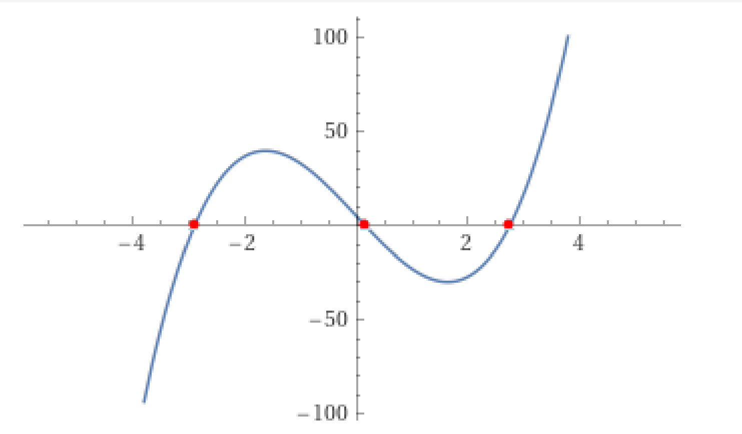
Find Stationary point.

$$\nabla f(\bar{x}) = 0$$

$$\begin{bmatrix} \frac{1}{2}(4x_1^3 - 32x_1 + 5) \\ \frac{1}{2}(4x_2^3 - 32x_2 + 5) \\ \vdots \\ \frac{1}{2}(4x_\mathrm{d}^3 - 32x_\mathrm{d} + 5) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$4x_1^3 - 32x_1 + 5 - 0\,, -\infty < x < \infty$$

By using graph method.



root of the equation $4x_1^3 - 32x_1 + 5 = 0$ is

$x_1 = -2.9035$ or 0.1567, or 2.746

similarly

$$\begin{cases} x_2 = -2.9035,0.1567,2.746 \\ \vdots \\ x_d = -2.9035,0.1567,2.746 \end{cases}$$

$x_1, x_2 . . . x_\mathrm{d}$ have same value

at

$(x_1, x_{2.\,.\,.}x_\mathrm{d}) = (-2.9035, -2.9035, .\ .\ ., -2.9035)$

$$H(X) = \begin{bmatrix} 34.5818 & 0 & \cdots & \cdots & 0 \\ 0 & 34.5818 & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \cdots & 34.5818 \end{bmatrix}$$

It is semi $+ve$ definite because Trace$(\mathbf{H(X)}) > \mathbf{0}$ and $|\mathbf{H(X)}| > \mathbf{0}$

at

$(x_1, x_{2.\,.\,.}x_\mathrm{d}) = (0.1567,0.1567, .\ .\ .,0.1567)$

$$H(X) = \begin{bmatrix} -15.8526 & 0 & \cdots & \cdots & 0 \\ 0 & -15.8526 & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \cdots & -15.8526 \end{bmatrix}$$

It is not semi $+ve$ definite because Trace$\big(\mathbf{H(X)}\big) < \mathbf{0}$

at

$(x_1, x_{2.\,.\,.}x_\mathrm{d}) = (2.746,2.746, .\ .\ .,2.746)$

$$H(X) = \begin{bmatrix} 29.2694 & 0 & \cdots & \cdots & 0 \\ 0 & 29.2694 & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \cdots & 29.2694 \end{bmatrix}$$

It is semi $+ve$ definite because Trace$(\mathbf{H}(\mathbf{X})) > \mathbf{0}$ and $|\mathbf{H}(\mathbf{X})| > \mathbf{0}$

$\boldsymbol{f}(\boldsymbol{X})$ **is minima at**

$$(x_1, x_{2. . .} x_{\mathrm{d}}) = (-2.9035, -2.9035, . . ., -2.9035), \qquad (2.746, 2.746, . . ., 2.746)$$

Conclusion
$\bar{x}_1^* = [2.7468 \quad ... \quad 2.7468]^\top$ is a minimizes

$$f(\bar{x}_1^*) = \frac{1}{2}[(2.7468)^4 - 16 \times (2.7418)^2 + 5 \times 2.7468] \times d$$
$$= \frac{1}{2}(-50.0588d) = -25.02944d$$

$$\bar{x}_2^* = [-2.90353 \ldots - 2.90353]^\top$$
$$f(\bar{x}_2^*) = \frac{1}{2}[(-2.90353)^4 - 16 \times (-2.9353)^2 + 5 \times (-2.90353)] \times d$$
$$= -39.166165d$$

$\therefore$ $\bar{x}_2^*$ is the global minimizes.

4) **Rosenbrock Function**

$$f(X) = \sum_{i=1}^{d-1} [100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2]$$
$$\frac{\partial f}{\partial x_1} = 200(x_2 - x_1^2)(-2x_1) + 2(x_1 - 1)$$
$$= -400x_1(x_2 - x_1^2) + 2(x_1 - 1)$$
$$\frac{\partial f}{\partial x_2} = -400x_2(x_3 - x_2^2) + 2(x_2 - 1) + 200(x_2 - x_1^2)$$
$$\frac{\partial f}{\partial x_{d-1}} = -400x_{d-1}(x_d - x_{d-1}^2) + 2(x_{d-1} - 1) + 200(x_{d-1} - x_{d-2}^2)$$
$$\frac{\partial f}{\partial x_d} = 200(x_d - x_{d-1}^2)$$
$$\therefore \nabla f(\mathrm{X}) = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3} \ldots \frac{\partial f}{\partial x_{\mathrm{d}}}\right]^\top$$
$$J = \begin{bmatrix} -400x_1(x_2 - x_1^2) + 2(x_1 - 1) \\ -400x_2(x_3 - x_2^2) + 2(x_2 - 1) + 200(x_2 - x_1^2) \\ \vdots \\ 200(x_d - x_{d-1}^2) \end{bmatrix}$$

$$H(X) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_2 x_1} & \cdots & \cdots & \frac{\partial^2 f}{\partial x_\partial x_1} \\ \cdots & \cdots & \cdots & \cdots & - \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_d} & \frac{\partial^2 f}{\partial x_1 x_d} & \cdots & \cdots & \frac{\partial^2 f}{\partial x_d^2} \end{bmatrix}$$

$$H(X) = \begin{bmatrix} -400(x_2 - 3x_1^2) + 2 & -400x_1 & \cdots & \cdots & 0 \\ -400x_1 & \cdots & \cdots & \cdots & - \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \cdots & \frac{\partial^2 f}{\partial x_d^2} \end{bmatrix}$$

5） **Root of Square Function**

$$f(X) = \sqrt{1 + x_1^2} + \sqrt{1 + x_2^2}$$

Jacobian $= J = \nabla f(x) = \left[\frac{\partial f}{\partial x_1} \quad \frac{\partial f}{\partial x_2}\right]^\top$

$$= \left[\frac{2x_1}{2\sqrt{1 + x_1^2}}, \quad \frac{2x_2}{2\sqrt{1 + x_2^2}}\right]^\top$$

$$= \left[\frac{x_1}{\sqrt{1 + x_1^2}}, \frac{x_2}{\sqrt{1 + x_2^2}}\right]^\top.$$

$$\frac{\partial^2 f}{\partial x_1^2} = \frac{1}{\sqrt{1 + x_1^2}} + x_1 * \frac{1}{2}(1 + x_1)^{-\frac{3}{2}} * 2x_1$$

$$= \frac{1 + x_1^2 - x^2}{\left(\sqrt{1 + x_1}\right)(1 + x_1^2)}$$

$$= \frac{1}{\left(\sqrt{1 + x_1}\right)(1 + x_1^2)}$$

$$\frac{\partial^2 f}{\partial x_1^2} = \frac{1}{(1 + x_1)^{3/2}}$$

similarly

$$\frac{\partial^2 f}{\partial {x_2}^2} = \frac{1}{(1 + x_2^2)^{\frac{3}{2}}}$$

$$\frac{\partial^2 f}{\partial x_1 \, \partial x_2} = 0 \qquad , \frac{\partial^2 f}{\partial x_2 \, \partial x_1} = 0$$

$$H(X) = \begin{bmatrix} \frac{1}{(1 + x_1^2)^{3/2}} & 0 \\ 0 & \frac{1}{(1 + x_2^2)^{3/2}} \end{bmatrix}$$

Find Stationary point.

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

$$= \begin{bmatrix} \dfrac{x_1}{\sqrt{1 + x_1^2}} \\ \dfrac{x_2}{\sqrt{1 + x_2^2}} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$x_1 = 0, x_2 = 0$$

at $x_1 = 0, x_2 = 0$

$$H(X) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\textbf{It is semi } + \boldsymbol{ve} \textbf{ definite because Trace}\big(\mathbf{H(X)}\big) < \mathbf{0} \textbf{ and } |\mathbf{H(X)}| < \mathbf{0}$$

$$\textbf{at } (\boldsymbol{x_1}, \boldsymbol{x_2}) = (\mathbf{0}, \mathbf{0})$$

$f(X)$ is minima at

$$(x_1, x_2) = (0, 0),$$

Global Minima $(\boldsymbol{x_1}, \boldsymbol{x_2}) = (\mathbf{0}, \mathbf{0})$

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

The algorithms were run for the following test cases. For some of them they converged but from some, they not converged due to several issues:-

Singular Hessian matrices

Inadequate number of iterations.

Becoming trapped at local minima

Divergent series leading to overflow

Result:

Test case	Backtracking	Bisection	Pure	Damped	Levenberg-Marquardt	Combined
0	[2. 2.]	[2. 2.]	[2. 2.]	[2. 2.]	[2. 2.]	[2. 2.]
1	[2. 2.]	[2. 2.]	[2. 2.]	[2. 2.]	[2. 2.]	[2. 2.]
2	[-1.748 0.874]	[-1.748 0.874]	[-1.748 0.874]	[-1.748 0.874]	[-1.748 0.874]	[0. -0.]
3	[1.748 -0.874]	[1.748 -0.874]	[1.748 -0.874]	[1.748 -0.874]	[1.748 -0.874]	[-0. 0.]
4	[0. 0.]	[1.748 -0.874]	[-1.748 0.874]	[-1.748 0.874]	[-1.748 0.874]	[-1.748 0.874]
5	[-0. -0.]	[-1.748 0.874]	[1.748 -0.874]	[1.748 -0.874]	[1.748 -0.874]	[1.748 -0.874]
6	[1. 1. 1. 0.999]	[1. 1. 1. 1.]	[1. 1. 1. 1.]	[1. 1. 1. 1.]	[1. 1. 1. 1.]	[1. 1. 1. 1.]
7	[1. 1. 1. 0.999]	[1. 1. 1. 1.]	[1. 1. 1. 1.]	[1. 1. 1. 1.]	[1. 1. 1. 1.]	[1. 1. 1. 1.]
8	[1. 1. 1. 0.999]	[1. 1. 1. 1.]	[-0.776 0.613 0.382 0.146]	[-0.776 0.613 0.382 0.146]	[-0.776 0.613 0.382 0.146]	[-0.776 0.613 0.382 0.146]
9	[1. 1. 1. 0.999]	[1. 1. 1. 1.]	[1. 1. 1. 1.]	[1. 1. 1. 1.]	[1. 1. 1. 1.]	[1. 1. 1. 1.]
10	[-2.904 -2.904 -2.904 -2.904]	[-2.904 -2.904 -2.904 -2.904]	[0.157 0.157 0.157 0.157]	[0. 0. 0. 0.]	[-2.904 -2.904 -2.904 -2.904]	[-2.904 -2.904 -2.904 -2.904]
11	[-2.904 -2.904 -2.904 -2.904]	[2.747 2.747 2.747 2.747]	[2.747 2.747 2.747 2.747]	[2.747 2.747 2.747 2.747]	[2.747 2.747 2.747 2.747]	[2.747 2.747 2.747 2.747]
12	[-2.904 -2.904 -2.904 -2.904]	[-2.904 -2.904 -2.904 -2.904]	[-2.904 -2.904 -2.904 -2.904]	[-2.904 -2.904 -2.904 -2.904]	[-2.904 -2.904 -2.904 -2.904]	[-2.904 -2.904 -2.904 -2.904]
13	[2.747 -2.904 2.747 -2.904]	[2.747 -2.904 2.747 -2.904]	[2.747 -2.904 2.747 -2.904]	[2.747 -2.904 2.747 -2.904]	[2.747 -2.904 2.747 -2.904]	[2.747 -2.904 2.747 -2.904]
14	[0. 0.]	[0. 0.]	None	[-0. -0.]	[-8.71896425e+115 -8.71896425e+115]	[0. 0.]
15	[0. -0.]	[-0. 0.]	[0. -0.]	[0. -0.]	[0. -0.]	[0. -0.]
16	[0. -0.]	[0. 0.]	None	[-0. 0.]	[1.61634765e+132 0.00000000e+000]	[0. -0.]

Test case table

Test case	Function	Initial points
0	trid_function	[-2.0, -2]
1	trid_function	[-2.0, -2]
2	three_hump_camel_function	[-2.0, 1]
3	three_hump_camel_function	[2.0, -1]
4	three_hump_camel_function	[-2.0, -1]
5	three_hump_camel_function	[2.0, 1]
6	rosenbrock_function	[2.0, 2, 2, -2]
7	rosenbrock_function	[2.0, -2, -2, 2]
8	rosenbrock_function	[-2.0, 2, 2, 2]
9	rosenbrock_function	[3.0, 3, 3, 3]
10	styblinski_tang_function	[0.0, 0, 0, 0]
11	styblinski_tang_function	[3.0, 3, 3, 3]
12	styblinski_tang_function	[-3.0, -3, -3, -3]
13	styblinski_tang_function	[3.0, -3, 3, -3]
14	func_1	[3.0, 3]

15	func_1	[-0.5, 0.5]
16	func_1	[-3.5, 0.5]

Three Hump camel function:

(0,0) is the global minima for the function.

(-1.7474, 0.8737) and (1.7474, -0.8737) are local minima for the function

As show in the result table their are many test case where global minima is not occur due to stuck at local minima.

Only 4 test case where the global minima achieve and 20 test case where the global minima is not achieve

Fail to achieve global minima

- Fail at test case 2, 3 at conditions Backtracking, bisection, pure, Damped ,L-M
- Fail at test case 4, 5 at conditions bisection ,pure, Damped ,L-M,combine

Rosenbrock function:

Minima for the function is [1,1,1,….1]

As show in table there are many test cases where the minima is not occur because iteration required for convergence exceeded 10000.

8 test case where the global minima is not achieve

Fail to achieve global minima

- Fail at test case 6,7,9 at conditions Backtracking
- Fail at test case 8, at conditions Backtracking, bisection ,pure, Damped ,L-M,combine

Styblinski Tang function:

- X1 = [-2.90353 -2.90353 -2.90353 … -2.90353] ➔ THIS IS THE GLOBAL MINIMA
- X3 = [2.7468 2.7468 2.7468 …. 2.7468] ➔ THIS IS A LOCAL MINIMA

Test case 10 damp method stuck on initial point and pure method is also stuck on point

Test case 11 stuck in local minima for bisection, damped, pure L-M combine

In Newton's method, starting from a point far from the actual minimum can cause the algorithm to become stuck and fail to converge. Likewise, the table illustrates additional test cases that failed to converge due to one of the aforementioned reasons.

Fail to achieve global minima

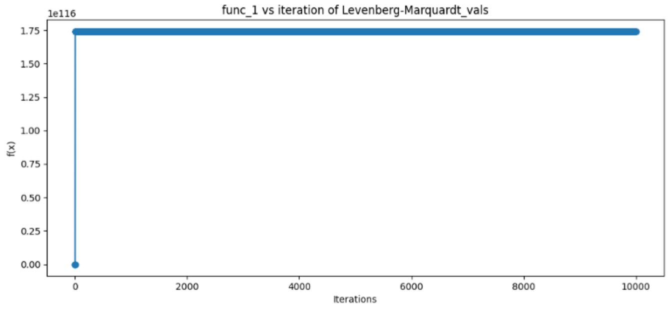
- Fail at test case 10 at conditions pure and damped
- Fail at test case 11 at conditions bisection ,pure, Damped ,L-M,combine
- Fail at test case 12 at conditions Backtracking ,bisection ,pure, Damped ,L-M,combine

Root of square function:

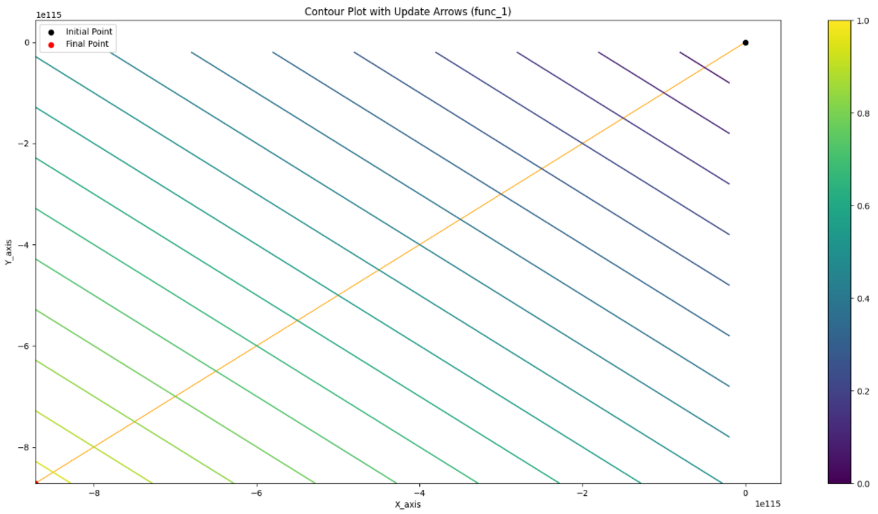
One stationary point and it is global minima that is (0,0)

I test case 14,16 the pure condition the overflow encountered RuntimeWarning: overflow encountered in power return np.diag(1 / ((1 + point**2) ** 1.5))

This is the reason for test case 14 and test case 16, you see such huge values in Levenberg-Marquard column.



initial point [-3, 3] Levenberg-Marquardt condition the function is not Converge



initial point [-3, 3] Levenberg-Marquardt condition the function is not Converge

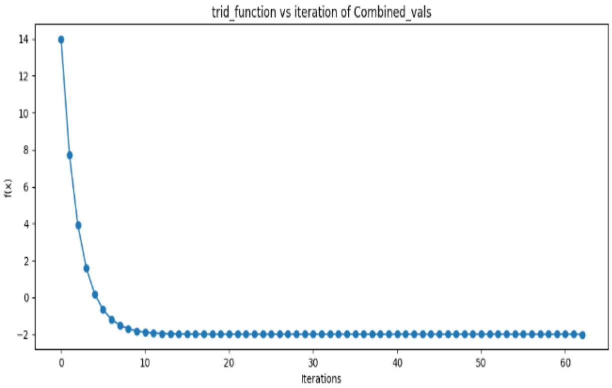
Fail to achieve global minima

- Fail at test case 14 at conditions pure and L-M
- Fail at test case 16 at conditions pure and L-M

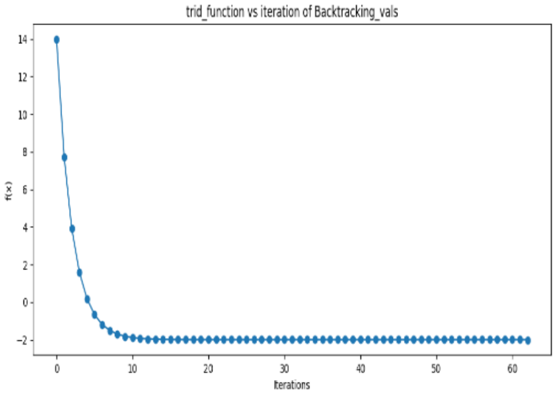
Plot f(x) vs iterations and |f′ (x)| vs iterations

f(x) vs iterations

Plot f(x) vs iterations of Trid Function.

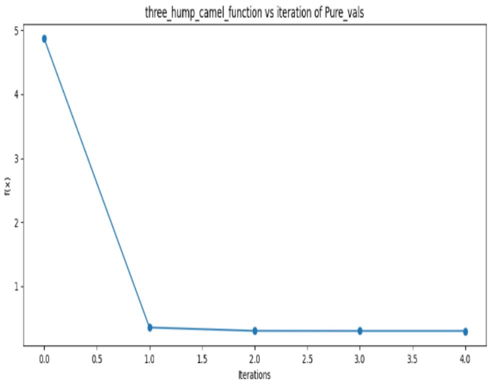


initial point [-2,-2]

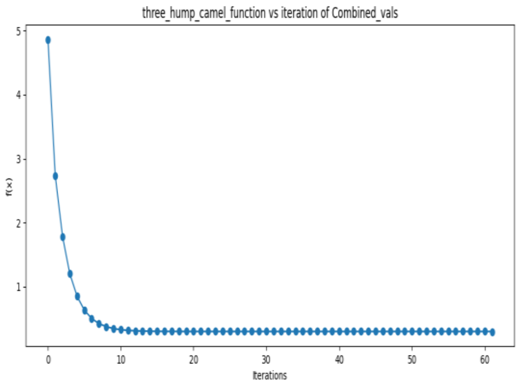


initial point [-2,-2]

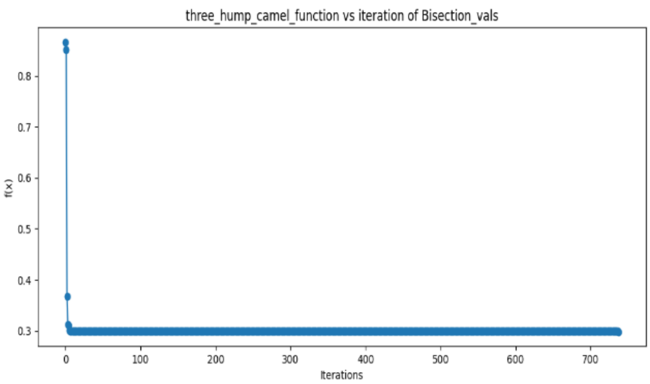
Plot f(x) vs iterations of Three hump camel function



initial point [2,1]

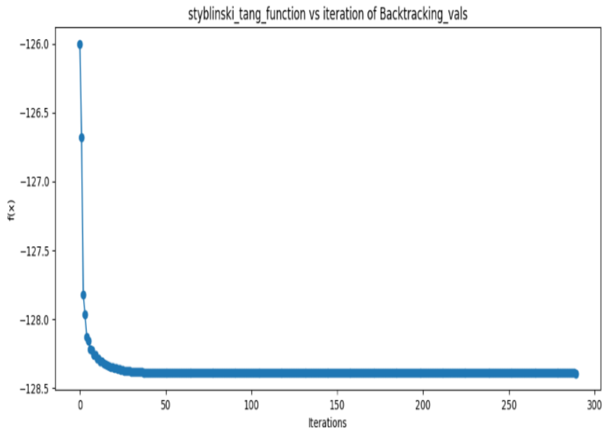


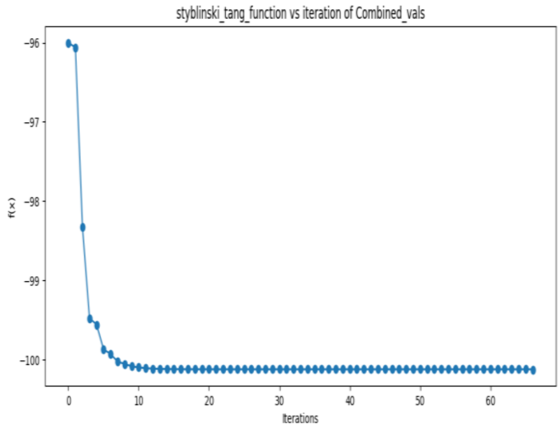
initial point [-2,-1]



initial point [2,- 1]

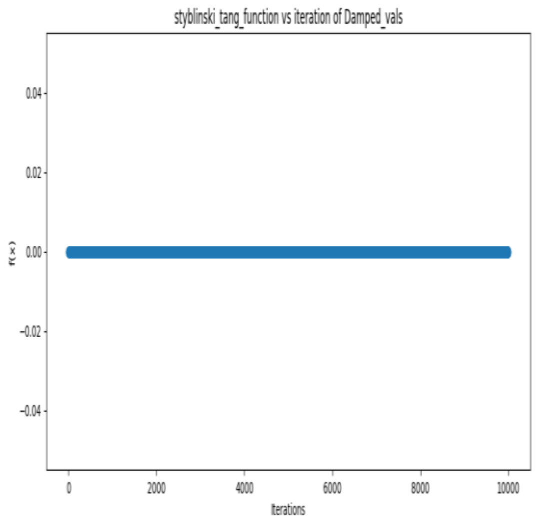
Plot f(x) vs iterations Styblinski-Tang Function

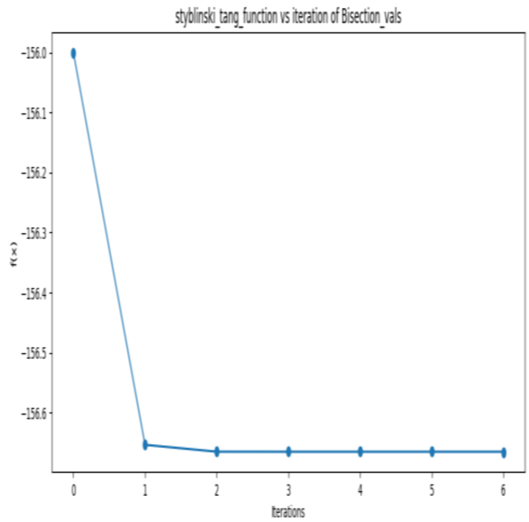




initial point [3,-3,3,-3]

initial point [3,3,3,3]

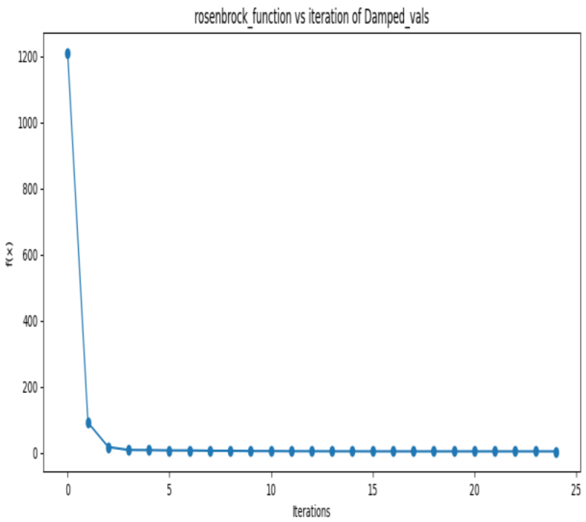


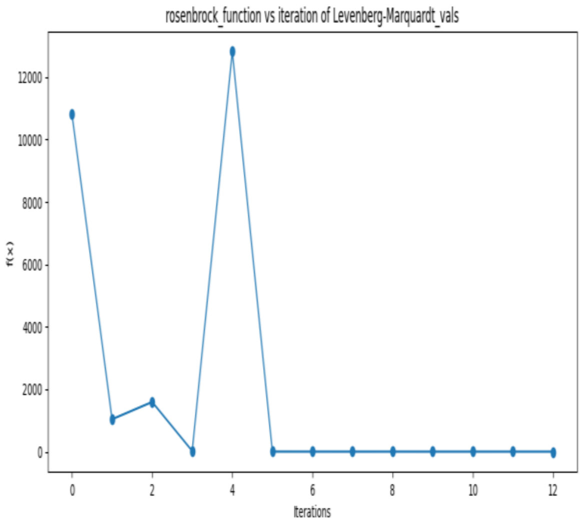


initial point [0,0,0,0]
damped condition the f(x) is not Converge

initial point [-3,-3,-3,-3]

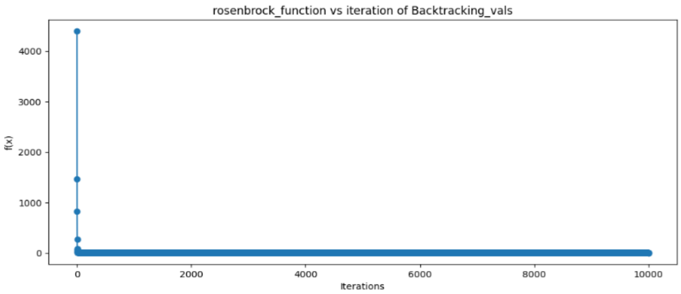
Plot f(x) vs iterations Rosenbrock Function





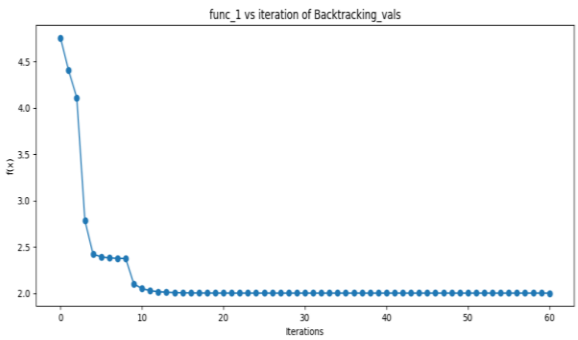
initial point [-2,2,2,2]

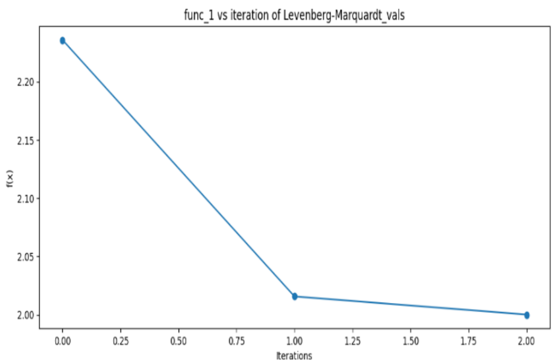
initial point [3,3,3,3]



At initial point [2,2,2,2]
backtracking condition the function is not Converge

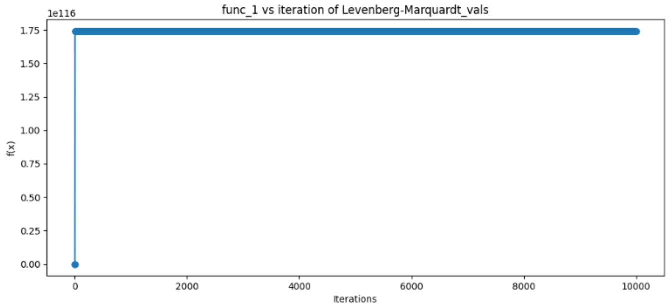
Plot f(x) vs iterations Root of Square Function(func_1)





initial point [-0.5, 0.5]

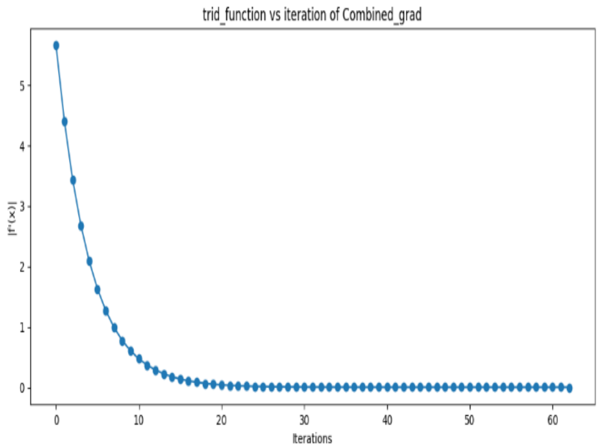
initial point [-3.5, 0.5]



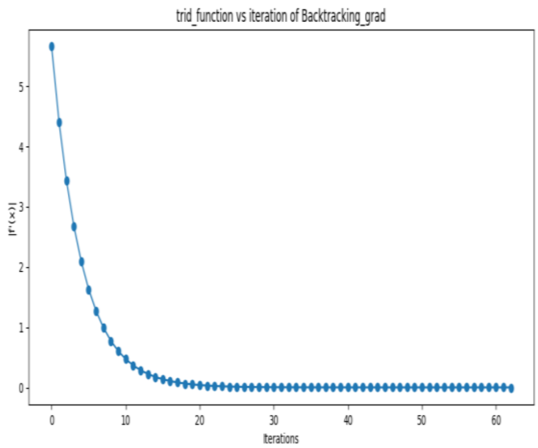
At initial point [-3, 3] Levenberg-Marquardt condition the function is not Converge

|f′(x)| vs iterations

Plot |f′(x)| vs iterations of Trid Function.

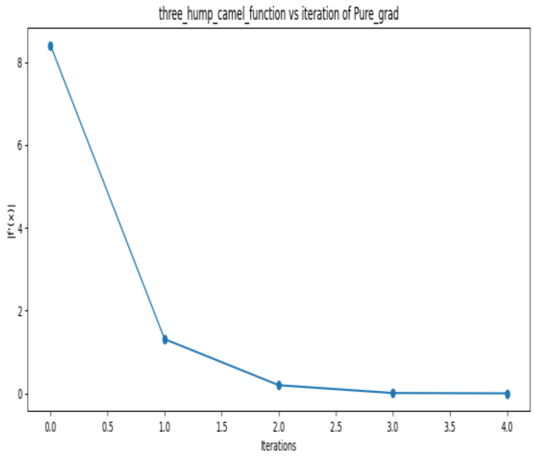


initial point [-2,-2]

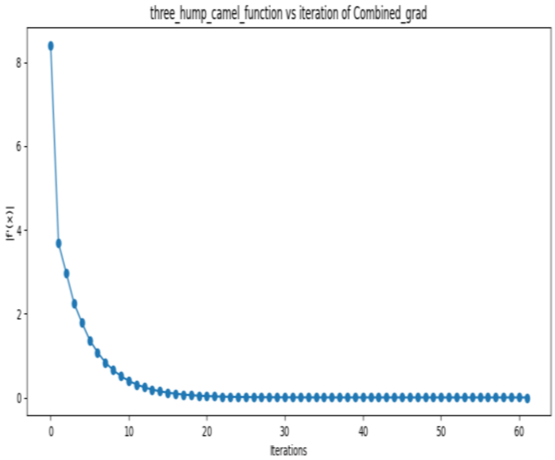


initial point [-2,-2]

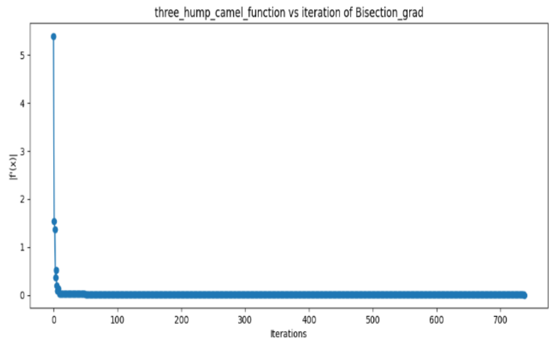
Plot |f′(x)| vs iterations of Three hump camel function



initial point [2,1]

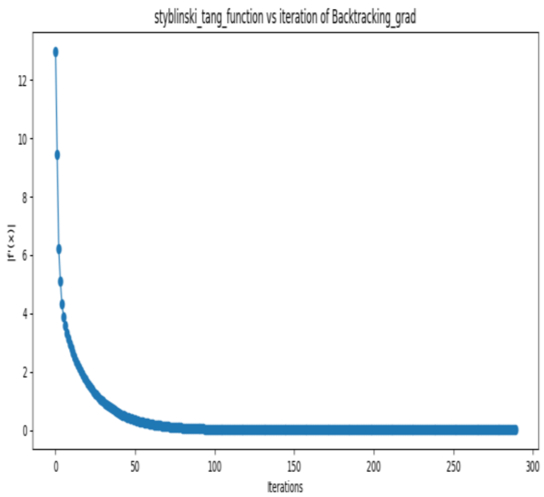


initial point [-2,-1]

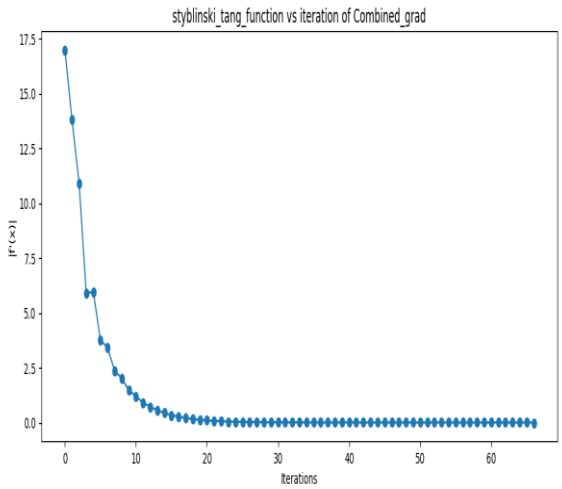


initial point [2,-1]

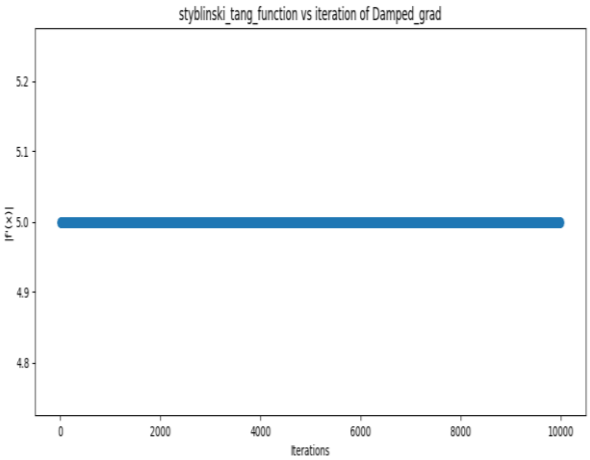
Plot |f′(x)| vs iterations Styblinski-Tang Function

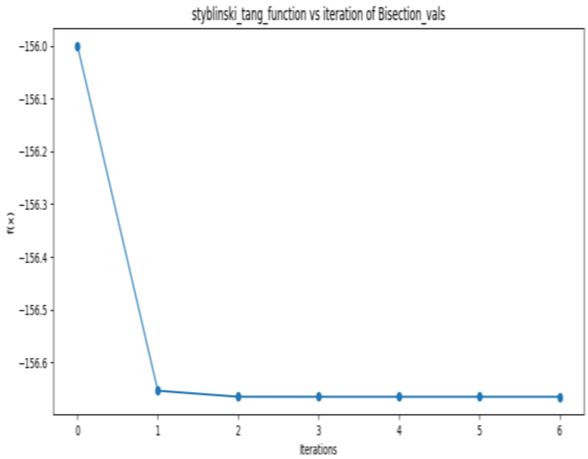


initial point [3,-3,3,-3]



initial point [3,3,3,3]

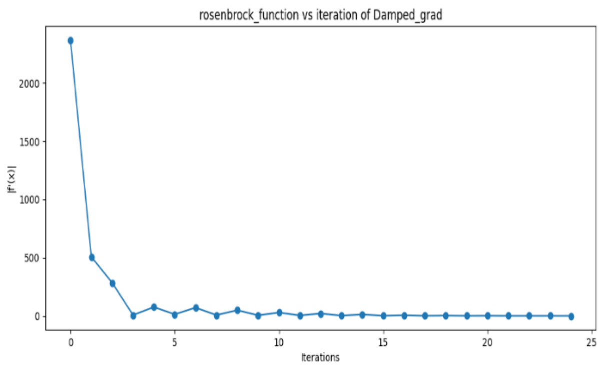




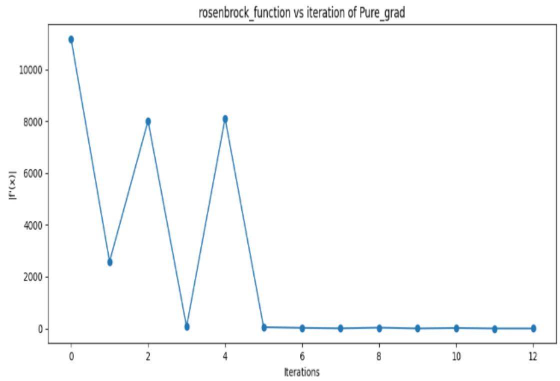
initial point [0,0,0,0] at
damped condition the f(x) is not Converge

initial point [-3,-3,-3,-3]

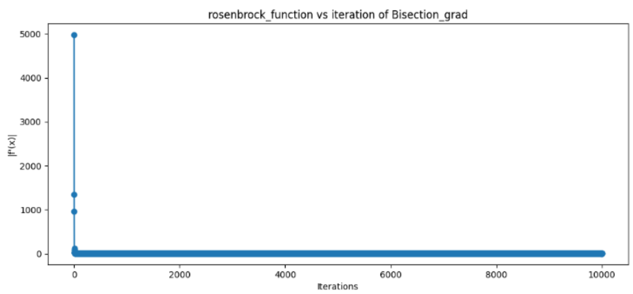
Plot |f′(x)| vs iterations Rosenbrock Function



initial point [-2,2,2,2]

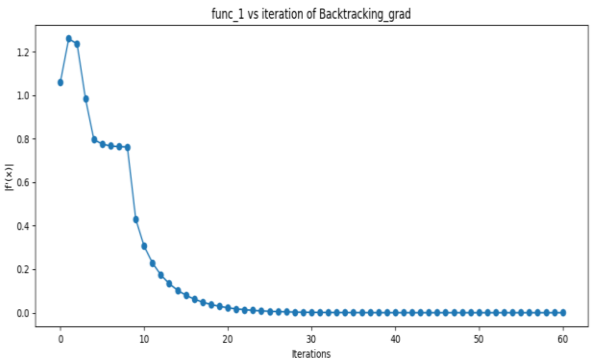


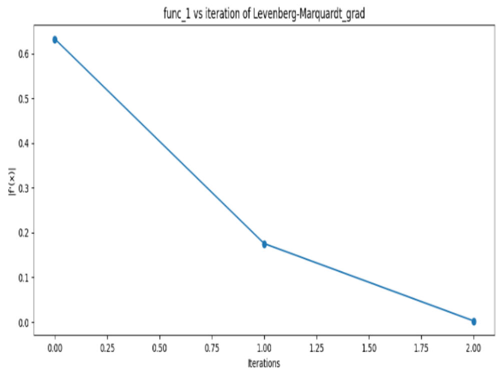
initial point [3,3,3,3]



initial point [2,2,2,2]
at backtracking the function is not Converge

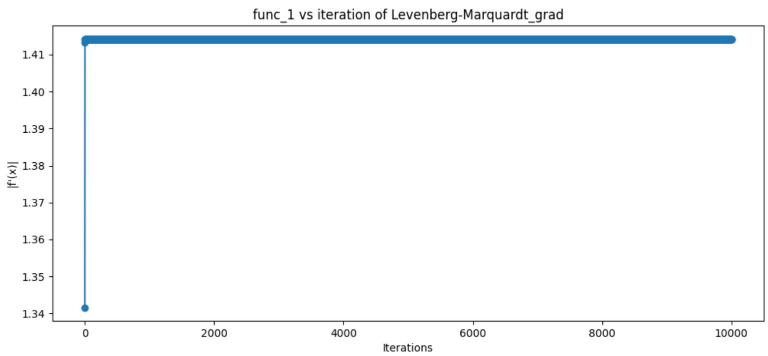
Plot |f′(x)| vs iterations Root of Square Function(func_1)





initial point [-3.5, 0.5]

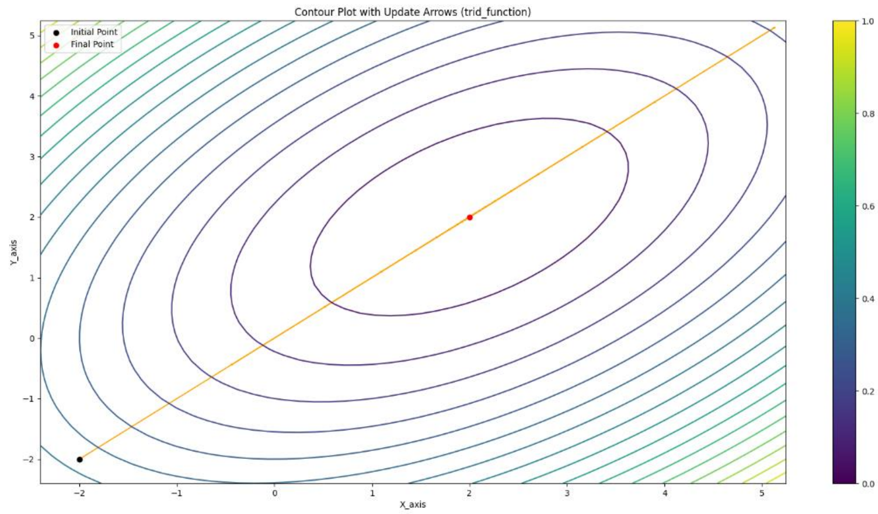
initial point [-0.5, 0.5]



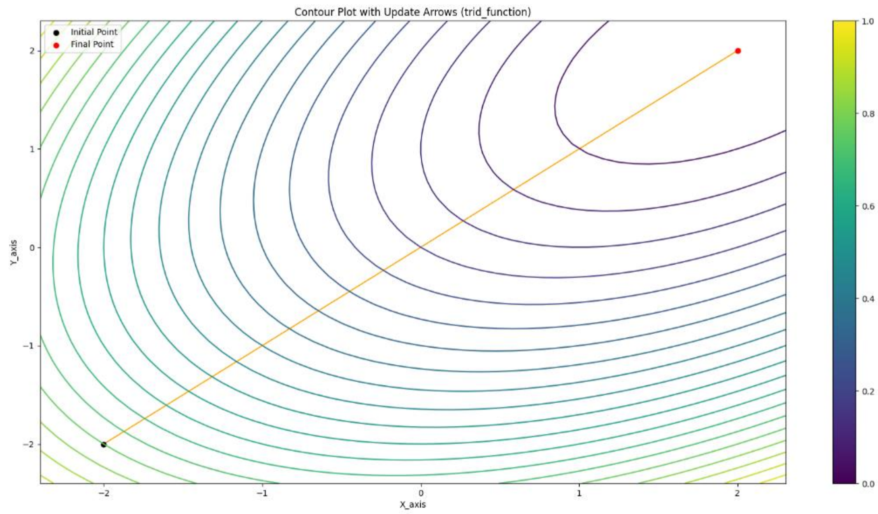
initial point [-3, 3] Levenberg-Marquardt condition the function is not Converge

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

Trid Function

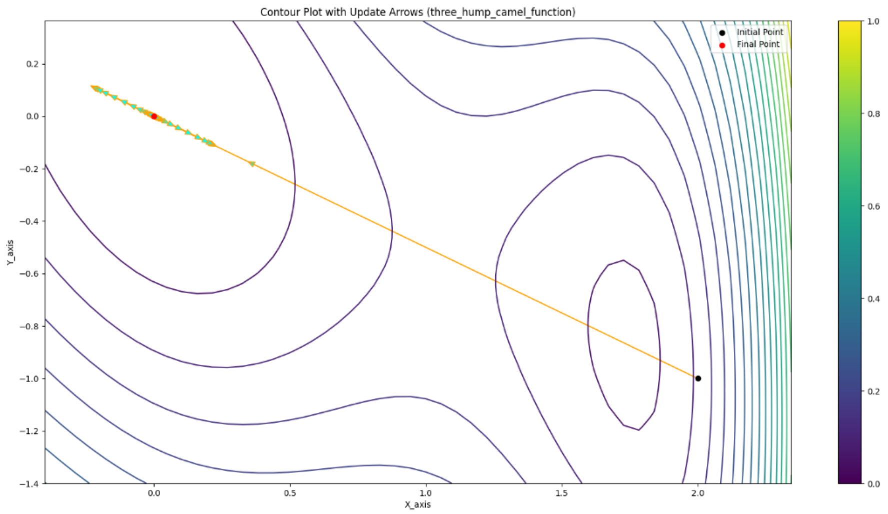


Initial points [-2,-2] at combined condition, final point [2,2]

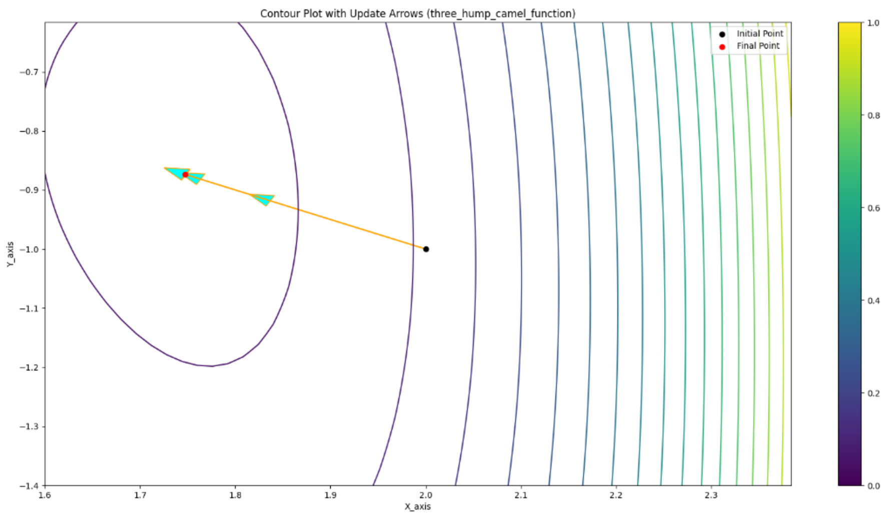


Initial points [-2,-2] at bisection condition, final point [2,2]

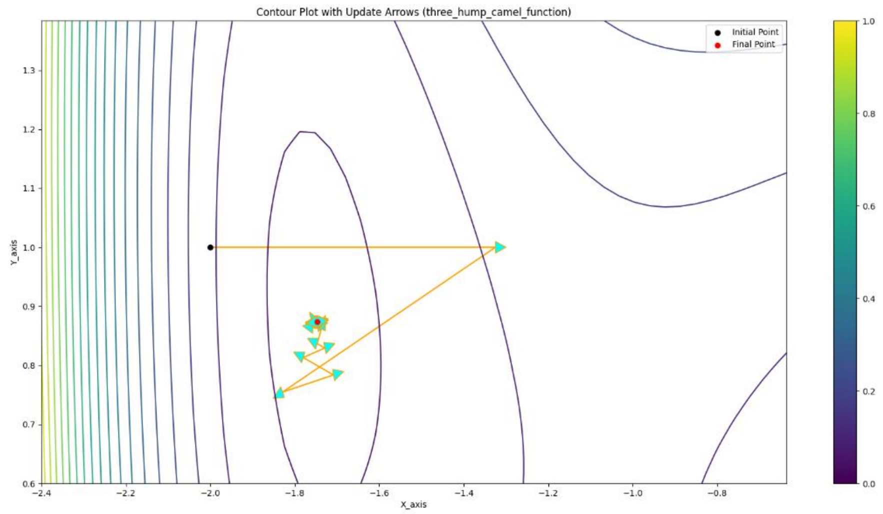
Three Hump Camel



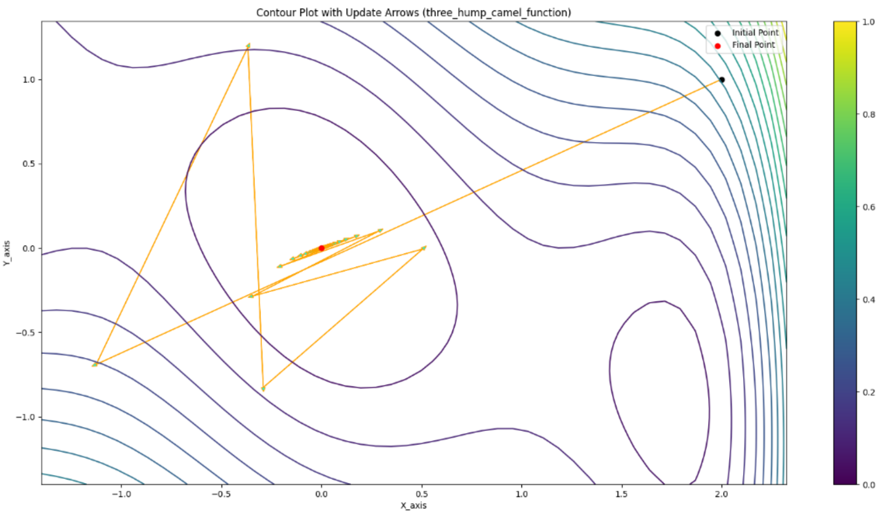
Initial points [2,-1] at combined condition, final point [0,0]



Initial points [2,-1] at damped condition, final point [1.748,-0.874]

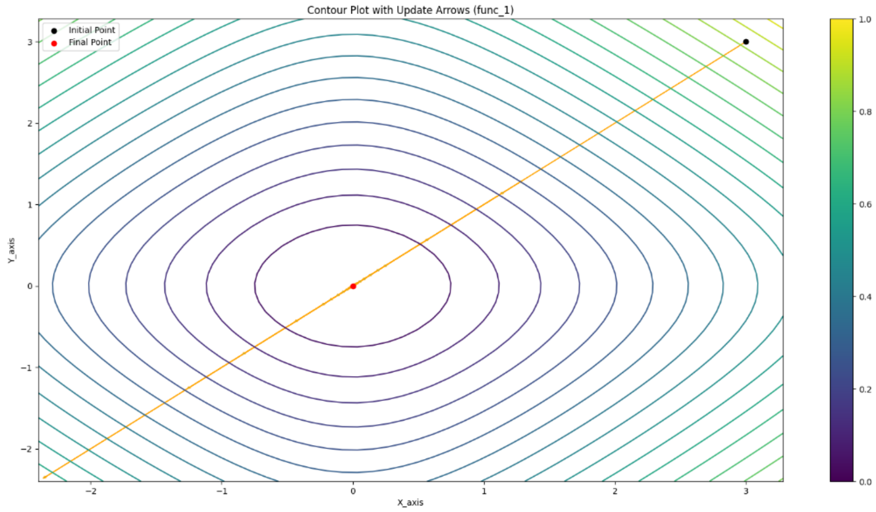


Initial points [-2,1] at Bisection condition, final point [0.874, -1.748]

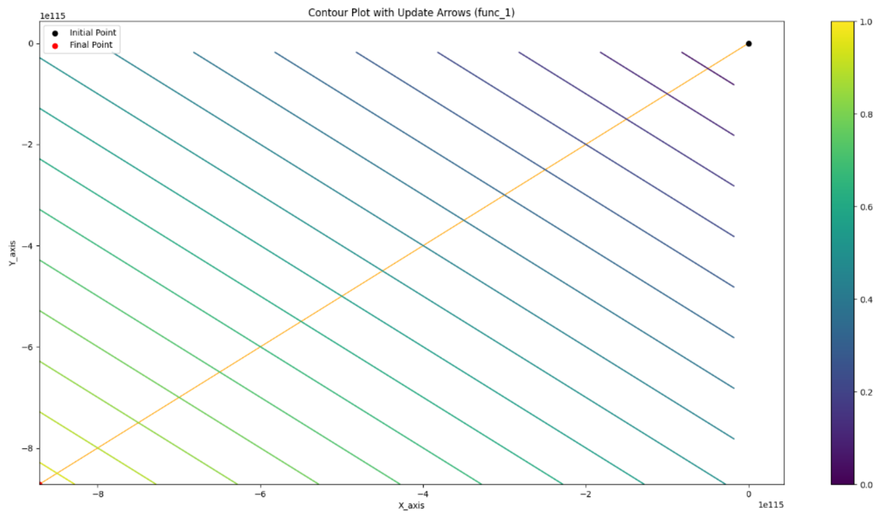


Initial points [2,1] at Backtracking condition, final point [0, 0]

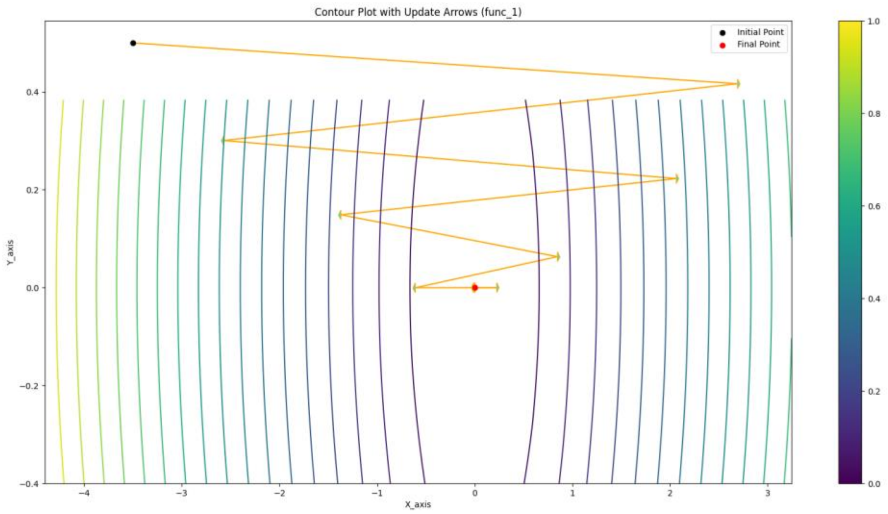
Root of Square Function(func_1)



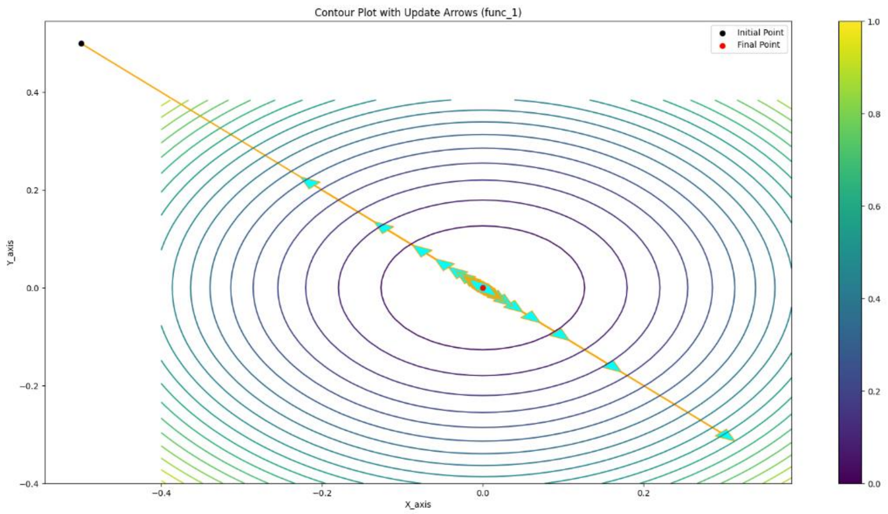
Initial points [3,3] at combined condition, final point [0, 0]



Initial points [-3.5,0.5] at Levenberg_Marquart condition



Initial points [-3.5,0.5] at Levenberg_Marquart condition , final point [0,0]



Initial points [-0.5,0.5] at Backtracking condition , final point [0,0]