ASSIGNMENT-1 REPORT

Optimization Methods

Hrishikesh Nakka

Roll Number: 2023201012

1 Trid Function

1.1 Mathematical Formulation

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

1.2 Jacobian

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

1.3 Hessian

$$\nabla^2 f(x) = \begin{bmatrix} 2 & -1 & \cdots & -1 \\ -1 & 2 & \cdots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & \cdots & 2 \end{bmatrix}_{d\times d}$$

1.4 Minima

To simplify solving for minima, We take $\mathbf{d = 2}$, since the test cases take 2-dimensional space.

This means

$$\nabla f(\mathbf{x}) = \begin{bmatrix} 2(x_1 - 1) - x_2 \\ 2(x_2 - 1) - x_1 \end{bmatrix}$$

and

$$\nabla^2 f(\mathbf{x}) = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

Since the Hessian is positive definite, we can say that the function is convex and it minimizes at $\mathbf{x}$ which gives $\mathbf{f(x) = 0}$. Solving for

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

here when $\mathbf{d = 2}$, we get

$$\mathbf{x} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

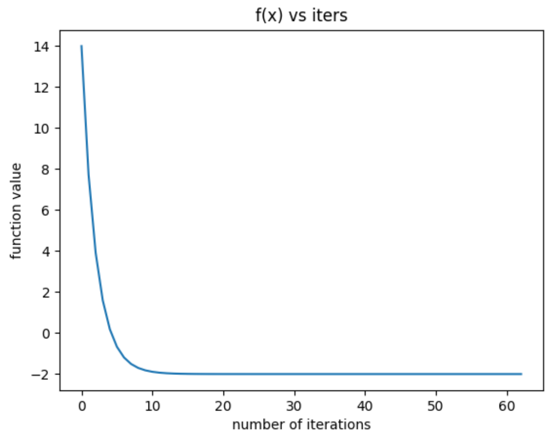

Figure 1: trid function [-2. -2.] Backtracking vals

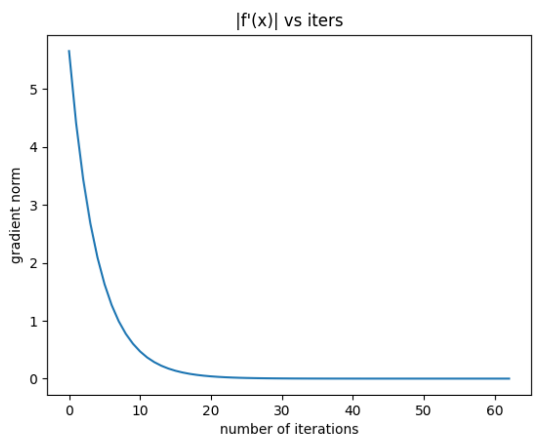

Figure 2: trid function [-2. -2.] Backtracking gradient

1.5 Algorithms that failed to converge:

All the given algorithms from all the given initial points were able to converge for this function.

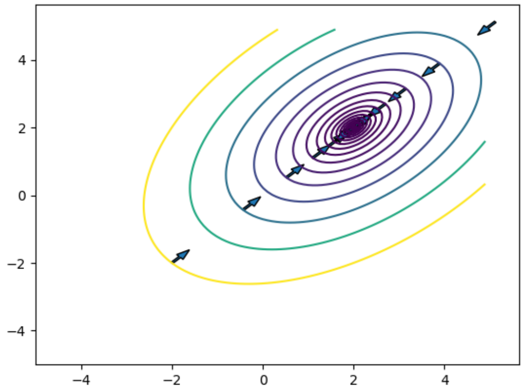

Figure 3: trid function [-2. -2.] Backtracking contour

1.6 Graphs

2 Three Hump Camel Function

2.1 Mathematical Formulation

$$f(x,y) = 2x^2 - 1.05x^4 + \frac{x^6}{6} + xy + y^2$$

2.2 Jacobian

$$\nabla f(x,y) = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} 4x - 4.2x^3 + x^5 + y \\ x + 2y \end{bmatrix}$$

2.3 Hessian

$$\nabla^2 f(x,y) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 4 - 12.6x^2 + 5x^4 & 1 \\ 1 & 2 \end{bmatrix}$$

2.4 Minima

Setting gradient to zero we get

$4 - 4.2x^3 + x^5 + y = 0$

$x + 2y = 0$

by solving this two equations we get solutions as

(1.0705,-0.53525), (-1.0705,0.53525), (1.74755235 -0.87377617), (-1.74755235 0.87377617) and (0,0).

Substituting (0, 0) in hessian we get

$$\nabla^2 f(x,y) = \begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix}$$

which is positive definite hence it is a minima

Substituting (-1.74755235, 0.87377617) in hessian we get

$$\nabla^2 f(x,y) = \begin{bmatrix} 12.153 & 1 \\ 1 & 2 \end{bmatrix}$$

which is also positive definite hence it is a minima

Substituting (1.74755235, -0.87377617) in hessian we get

$$\nabla^2 f(x,y) = \begin{bmatrix} 12.153 & 1 \\ 1 & 2 \end{bmatrix}$$

which is also positive definite hence it is a minima

Substituting (1.0705, -0.53525) in hessian we get

$$\nabla^2 f(x,y) = \begin{bmatrix} -3.879 & 1 \\ 1 & 2 \end{bmatrix}$$

which is not a positive definite hence it is not a minima

Substituting (-1.0705, 0.53525) in hessian we get

$$\nabla^2 f(x, y) = \begin{bmatrix} -3.879 & 1 \\ 1 & 2 \end{bmatrix}$$

which is not a positive definite hence it is not a minima

out of these five points, the function has the minimum value at point (0,0) so it is considered as global minimum and other two are considered as local minima.

2.5 Algorithms that failed to converge:

- **Steepest Descent with backtracking** was not able to converge to global minima with initial point as [-2,1], [2,-1].But were able to converge to some local minima.
- **Steepest Descent with bisection** was not able to converge to global minima with initial point as [-2,1], [2,-1], [-2,-1], [2,1].But were able to converge to some local minima.
- **Pure Newton algorithm** was not able to converge to global minima with initial point as [-2,1], [2,-1], [-2,-1], [2,1].But were able to converge to some local minima.
- **Damped Newton algorithm** was not able to converge to global minima with initial point as [-2,1], [2,-1], [-2,-1], [2,1].But were able to converge to some local minima.
- **Combined Newton algorithm** was not able to converge to global minima with initial point as [-2,1], [2,-1], [-2,-1], [2,1].But were able to converge to some local minima.
- **Levenberg-Marquardt Newton algorithm** was not able to converge to global minima with initial point as [-2,-1], [2,1].But were able to converge to some local minima.

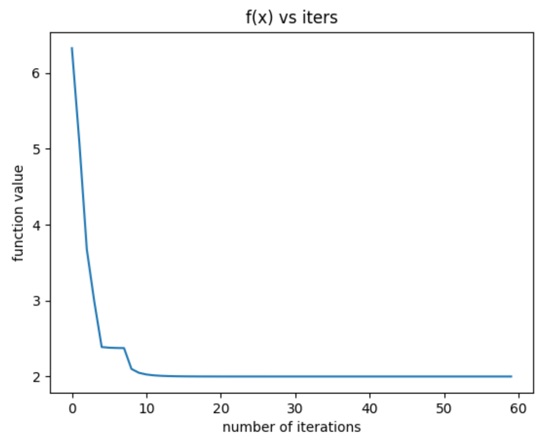

Figure 4: three hump camel function [-2. 1.] Backtracking vals

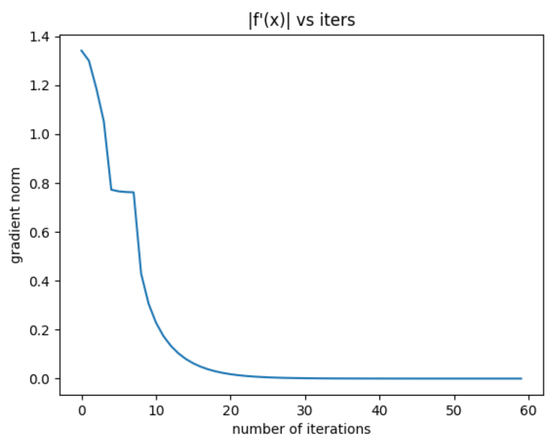

Figure 5: three hump camel function [-2. 1.] Backtracking grad

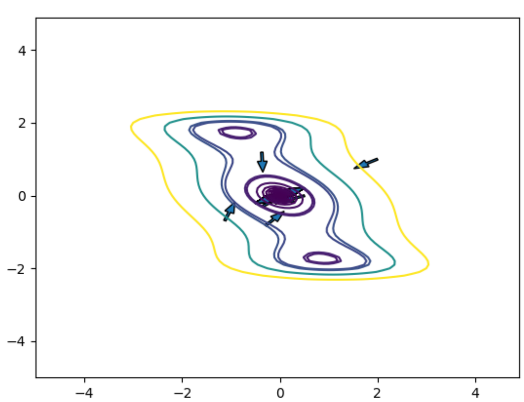

Figure 6: three hump camel function[-2. 1.] Backtracking cont

2.6 Graphs

3 Styblinski-Tang Function

3.1 Mathematical Formulation

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

3.2 Jacobian

$$\nabla f(\mathbf{x}) = \begin{bmatrix} 2x_1^3 - 32x_1 + 5 \\ 2x_2^3 - 32x_2 + 5 \\ \vdots \\ 2x_d^3 - 32x_d + 5 \end{bmatrix}$$

3.3 Hessian

$$\nabla^2 f(x) = \begin{bmatrix} 6x_1^2 - 32 & 0 & \cdots & 0 \\ 0 & 6x_2^2 - 32 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 6x_d^2 - 32 \end{bmatrix}$$

3.4 Minima

By equating gradient equal to zero

$$2x_1^3 - 32x_1 + 5 = 0, 2x_2^3 - 32x_2 + 5 = 0, \cdots, 2x_d^3 - 32x_d + 5 = 0$$

we get the solutions as

(0.1567, 0.1567, ... , 0.1567), (2.7468, 2.7468, ..., 2.7468), (-2.9035, -2.9035, ...,-2.9035)

Substituting (0.1567,...) in Hessian we get

$$\nabla^2 f(x) = \begin{bmatrix} -31.852 & 0 & \cdots & 0 \\ 0 & -31.852 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -31.852 \end{bmatrix}$$

Hessian to be not positive definite and so it is not a local minima

Substituting (2.7468, 2.7468,...) we get

$$\nabla^2 f(x) = \begin{bmatrix} 13.0456 & 0 & \cdots & 0 \\ 0 & 13.0456 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 13.0456 \end{bmatrix}$$

Hessian is positive definite matrix and so it is one of the minima

Substituting (-2.9035, -2.9035,...) we get

$$\nabla^2 f(x) = \begin{bmatrix} 18.58 & 0 & \cdots & 0 \\ 0 & 18.58 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 18.58 \end{bmatrix}$$

Hessian is positive definite matrix and so it is one of the minima

Out of these two minima, substituted in f(x) we get minimum value at (-2.9035,...) as -78.33

Hence (-2.9035, -2.9035, ...) is considered as global minima.

3.5 Algorithms that failed to converge:

- **Steepest Descent with backtracking** was not able to converge to global minima with initial point as [3, -3, 3, -3].But were able to converge to some local minima.
- **Steepest Descent with bisection** was not able to converge to global minima with initial point as [3 , 3, 3, 3], [3, -3, 3, -3].But were able to converge to some local minima.
- **Pure Newton algorithm** was not able to converge to global minima with initial point as [0, 0, 0, 0], [3 , 3, 3, 3], [3, -3, 3, -3].But were able to converge to some local minima.
- **Damped Newton algorithm** was not able to converge to global minima with initial point as [0, 0, 0, 0], [3 , 3, 3, 3], [3, -3, 3, -3].But were able to converge to some local minima.
- **Combined Newton algorithm** was not able to converge to global minima with initial point as [3 , 3, 3, 3], [3, -3, 3, -3].But were able to converge to some local minima.
- **Levenberg-Marquardt Newton algorithm** was not able to converge to global minima with initial point as [3 , 3, 3, 3], [3, -3, 3, -3].But were able to converge to some local minima.

3.6 Graphs

4 Rosen Brock Function

4.1 Mathematical Formulation

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

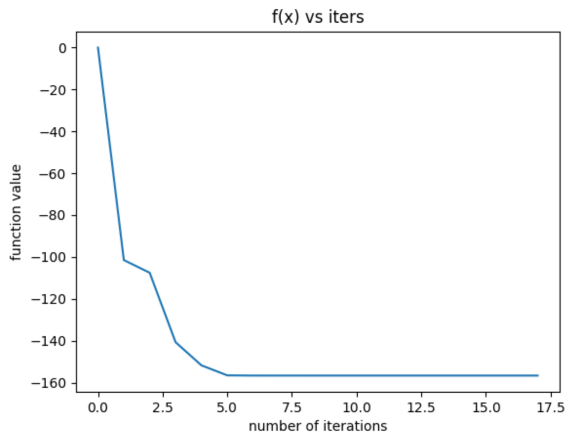

Figure 7: styblinski tang function [0. 0. 0. 0.] Bisection vals

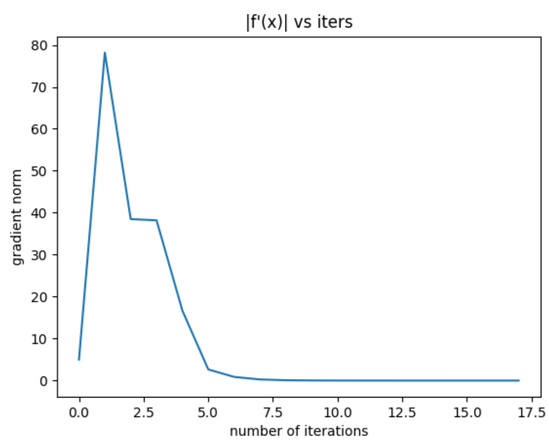

Figure 8: styblinski tang function [0. 0. 0. 0.] Bisection grad

4.2 Jacobian

$$\nabla(f)x = \begin{bmatrix} 400(x_1^2 - x_2)x_1 + 2(x_1 - 1) \\ 400(x_2^2 - x_3)x_2 + 2(x_2 - 1) + 200(x_2 - x_1^2) \\ 400(x_3^2 - x_4)x_3 + 2(x_3 - 1) + 200(x_3 - x_2^2) \\ \vdots \\ 400(x_{d-1}^2 - x_d)x_{d-1} + 2(x_{d-1} - 1) + 200(x_{d-1} - x_{d-2}^2) \\ 200(x_d - x_{d-1}^2) \end{bmatrix}$$

4.3 Hessian

$$\nabla^2 f(x) = \begin{bmatrix} 1200x_1^2 - 400x_2 + 2 & -400x_1 & \cdots & \cdots & 0 \\ -400x_1 & 1200x_2^2 - 400x_3 + 202 & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & -400x_{d-1} & 200 \end{bmatrix}$$

4.4 Algorithms that failed to converge

- **Pure Newton algorithm** was not able to converge to global minima with initial point as [-2,2,2,2] but were able to converge to local minima.
- **Damped Newton algorithm** was not able to converge to global minima with initial point as [-2,2,2,2] but were able to converge to local minima.
- **Combined Newton algorithm** was not able to converge to global minima with initial point as [-2,2,2,2] but were able to converge to local minima.
- **Levenberg-Marquardt Newton algorithm** was not able to converge to global minima with initial point as [-2,2,2,2] but were able to converge to local minima.

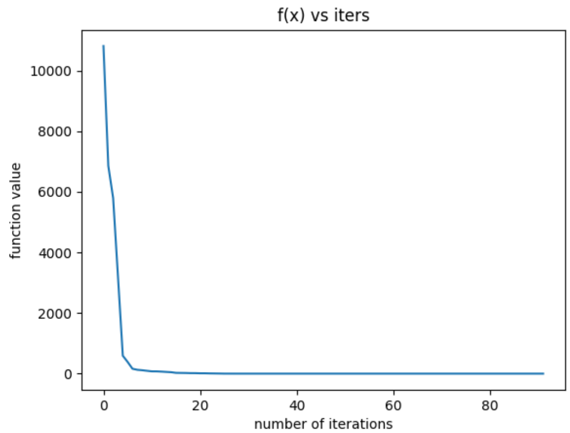

Figure 9: rosenbrock function [3. 3. 3. 3.] Combined vals

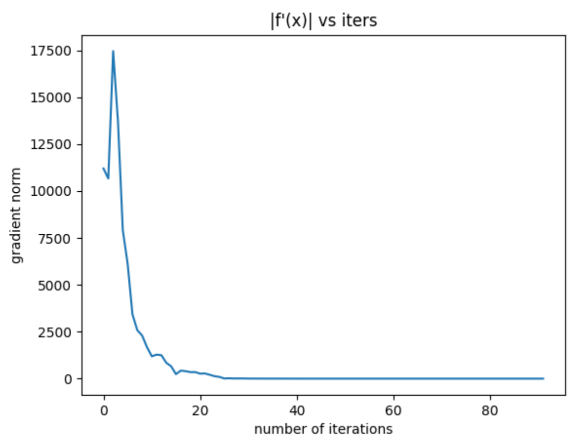

Figure 10: rosenbrock function [3. 3. 3. 3.] Combined grad

4.5 Graphs

5 Root of a square Function

5.1 Mathematical Formulation

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

5.2 Jacobian

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \frac{x_1}{\sqrt{1+x_1^2}} \\ \frac{x_2}{\sqrt{1+x_2^2}} \end{bmatrix}$$

5.3 Hessian

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

5.4 Minima

Equations from setting each component of the gradient to zero:

$$\begin{cases} \frac{x_1}{\sqrt{1+x_1^2}} = 0 \\ \frac{x_2}{\sqrt{1+x_2^2}} = 0 \end{cases}$$

Solving these equations, we find $x_1 = 0$ and $x_2 = 0$.

Therefore, the only critical point is $(0, 0)$. At the critical point $(0, 0)$, the Hessian matrix becomes:

$$\mathbf{H} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Since the Hessian matrix is positive definite (all eigenvalues are positive), the critical point $(0, 0)$ is a local minimum.

Therefore, the function $f(\mathbf{x}) = \sqrt{1 + x_1^2} + \sqrt{1 + x_2^2}$ has a minimum value at $\mathbf{x} = (0, 0)$.

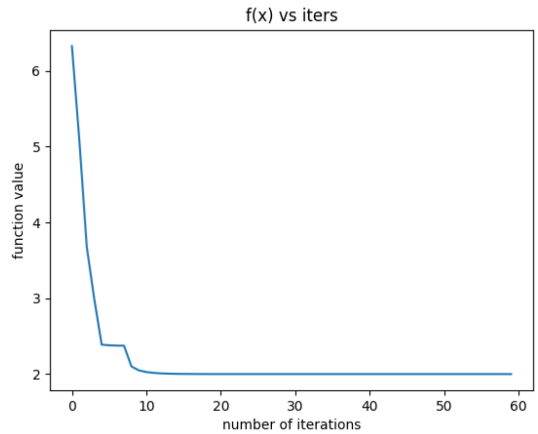

Figure 11: func 1[3. 3.] Backtracking vals

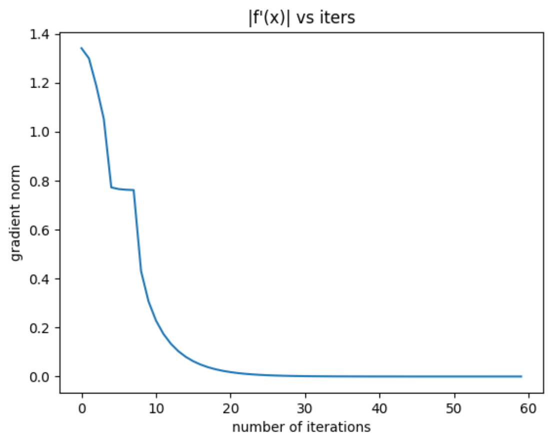

Figure 12: func 1[3. 3.] Backtracking grad

5.5 Algorithms that failed to converge

- **Pure Newton algorithm** was not able to converge to minima with initial points as [3,3] and [-3.5,0.5].
- **Levenberg-Marquardt Newton algorithm** was not able to converge to minima with initial points as [3,3] and [-3.5,0.5].

5.6 Graphs

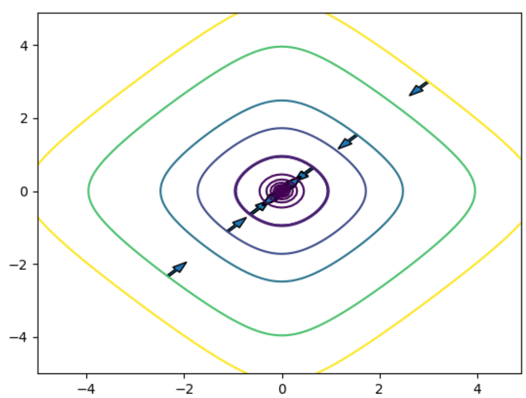

Figure 13: func 1[3. 3.] Backtracking cont