Name: Aneesh Panchal
SR No: 06-18-01-10-12-24-1-25223
Email ID: aneeshp@iisc.ac.in
Date: September 13, 2024
Assignment No: Assignment 1
Course Code: E0230
Course Name: Computational Methods of Optimization
Term: AUG 2024

Solution 1

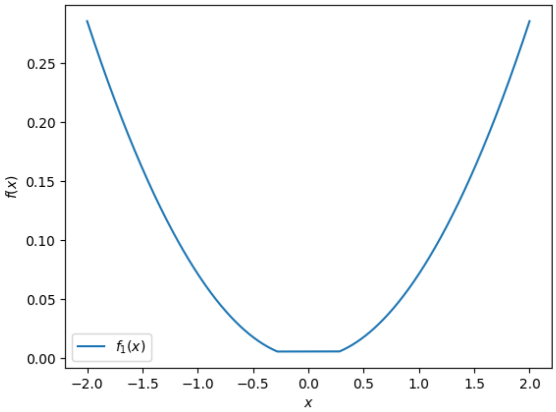
1. (a) Convexity can be tested using forward difference scheme. Negative values implies negative slope and Positive values implies positive slope. So, there must be negative values followed by either zero or, positive values. If slope values are increasing then function is convex else not.

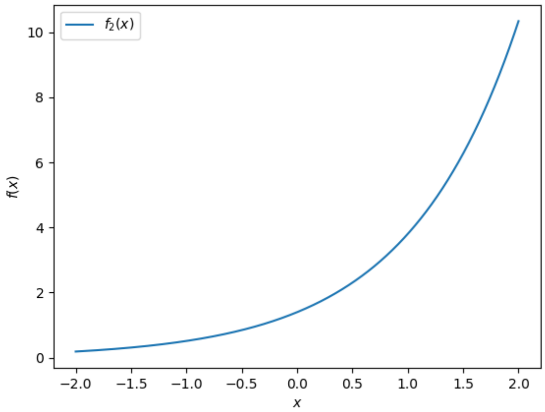
$$\text{Forward Difference Scheme, } slope = \frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i}$$

(b) f_1 is convex and f_2 is also convex function over domain $[-2, 2]$ with assumption of 10000 points in the interval. f_2 is strictly convex, however f_1 is not strictly convex.
Strict convexity can also be tested using forward difference scheme. If function is convex and slope is 0 for more than 1 point then function is not strictly convex, else function is strictly convex. $x*$ can be evaluated when either slope is 0 or, slope changes from negative to positive.
$x*$ and $f(x*)$ for the functions f_1 and f_2 are as follows,

$$f_1(x) = 0.005600, \quad \text{for } x \in [-0.222222, 0.222222], \quad \text{that is Multiple Minima.}$$

$$f_2(x) = 0.1894694, \quad \text{for } x = -2.0, \quad \text{that is Single Minima.}$$





2. (a) Yes, the function $f3$ is Coercive.
As it is given in the question that the oracle represents a quartic polynomial which means polynomial of degree 4. Hence, it can be easily interpolate using 5 points from oracle. Lets assume the polynomial turns out to be,

$$p(x) = a_0x^4 + a_1x^3 + a_2x^2 + a_3x + a_4$$

This polynomial is coercive if $a_0 > 0$ because a function is said to be coercive when $\lim_{x\to\infty} p(x) = \infty$.

If the condition quartic polynomial is not given then we can proceed with scaling and checking the values upto a specific range i.e. checking $x = 0, 10^1, 10^2, 10^4, 10^8, \ldots$ and the oracle will be coercive if $f(0) < f(10^1) < f(10^2) < \ldots$.

(b) As we have interpolated the polynomial in the previous part then, we can proceed to simply solving the polynomial and as we know that all the polynomial of degree n are n times continuously differentiable.

$$Oracle : 0.14285714x^4 - 0.1x^3 - 0.71428571x^2 + 0.21428571x + 0.19285714$$

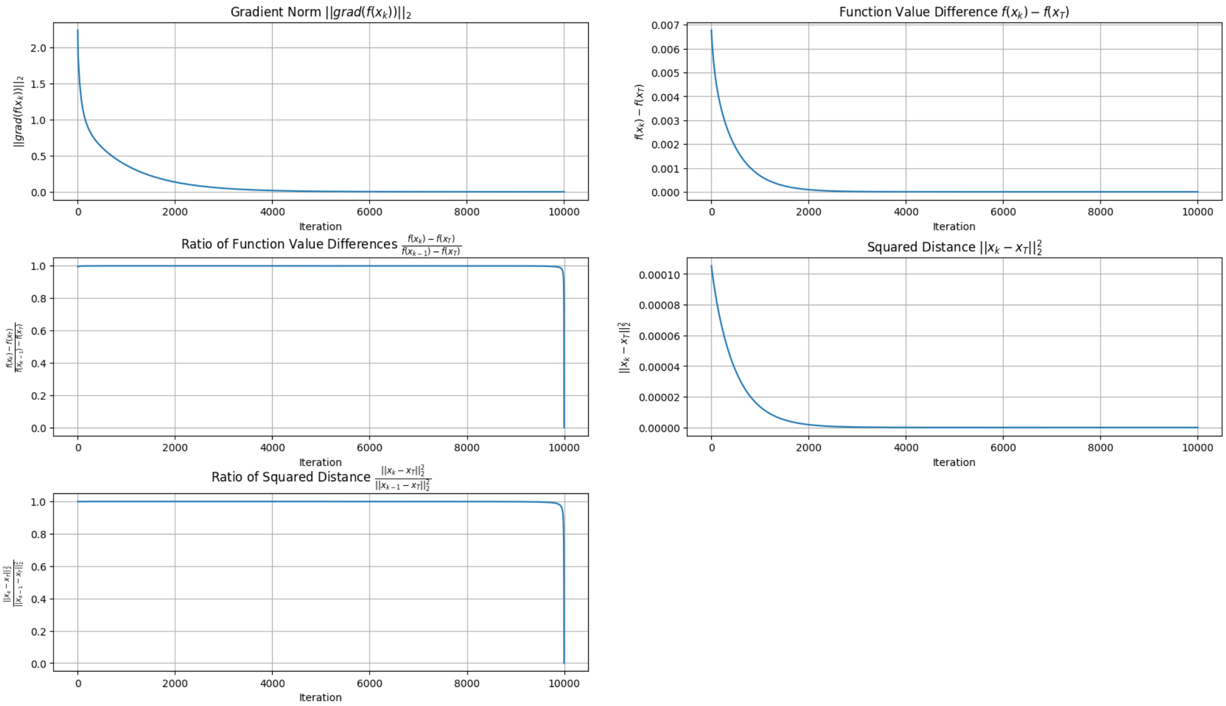
Hence, the results are as follows,
Roots: $[-2.0026369586263377, -0.40414530901623713, 0.6902388437031207, 2.416543423939612]$
Minima: $[-1.4206150925389083, 1.7988731352792633]$
Local Maxima: $[0.1467419572597644]$

Name: Aneesh Panchal
SR No: 06-18-01-10-12-24-1-25223
Email ID: aneeshp@iisc.ac.in
Date: September 13, 2024
Assignment No: Assignment 1
Course Code: E0230
Course Name: Computational Methods of Optimization
Term: AUG 2024

Solution 2

(a). **ConstantGradientDescent:**
Given and assumed parameters are as follows,
$x_0 = [0, 0, 0, 0, 0], \quad \alpha = 10^{-5}, \quad Iterations, T = 10000$

The final results are as follows,
$x* : [-3.96463545e-05, -5.00000000e-04, -1.00000000e-03, -2.00000000e-03, -9.99954827e-03]$
$f(x*) : -0.006769823167055695$
$Iterations : 10000$



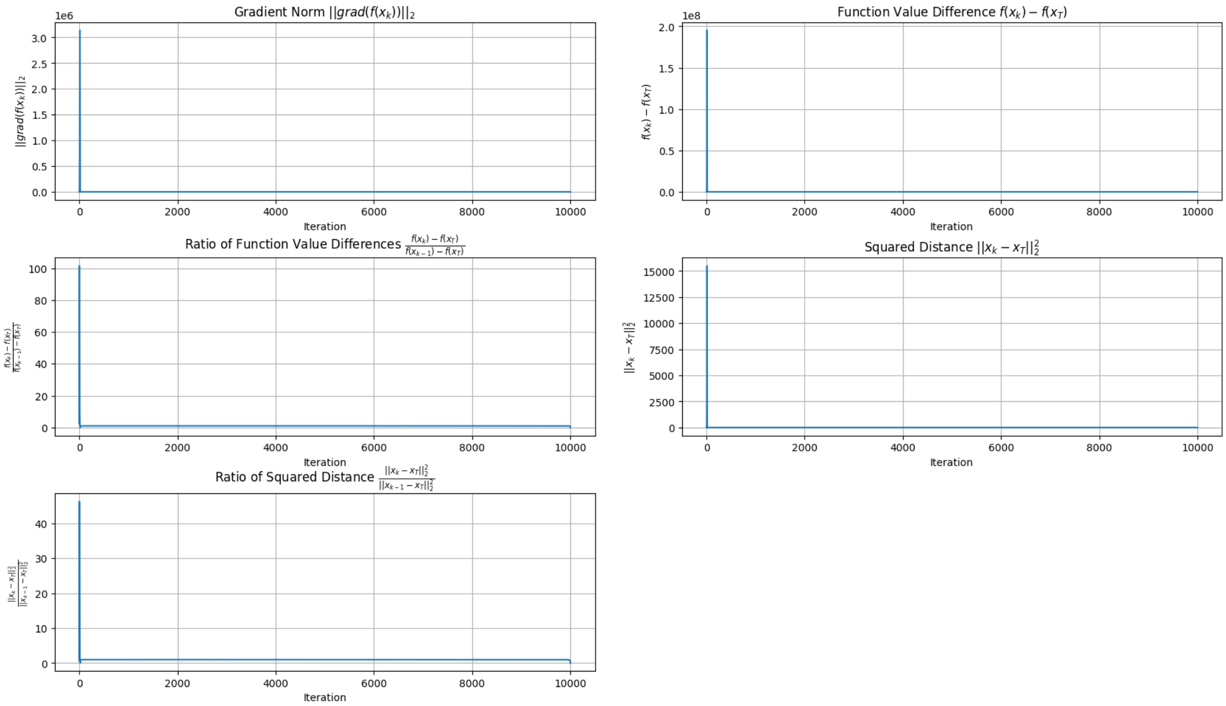
(b). **DiminishingGradientDescent:**
$x_0 = [0, 0, 0, 0, 0], \quad \alpha_0 = 10^{-3}, \quad Iterations, T = 10000$

The final results are as follows,
$x* : [0.00133996, -0.0005, -0.001, -0.00133017, -0.00163586]$
$f(x*) : 0.02084391990743637$
$Iterations : 10000$

$x* : [-3.96463545e-05, -5.00000000e-04, -1.00000000e-03, -1.98871635e-03, -6.27461458e-03]$
$f(x*) : -0.006075866519042196$
$Iterations : 10000$

Comparison of **ConstantGradientDescent** and **DiminishingGradientDescent:**
No, DiminishingGradientDescent values are not matching with ConstantGradientDescent. This may be due to the reason that in Diminishing case, initially the values of α are not small enough to converge to a local minima region and when it becomes small it becomes so small that the step taken becomes very small and hence remaining gap becomes large which can't be covered.
Remark: Similar behaviour can also be seen in Ques 3(4) and Ques 3(5).

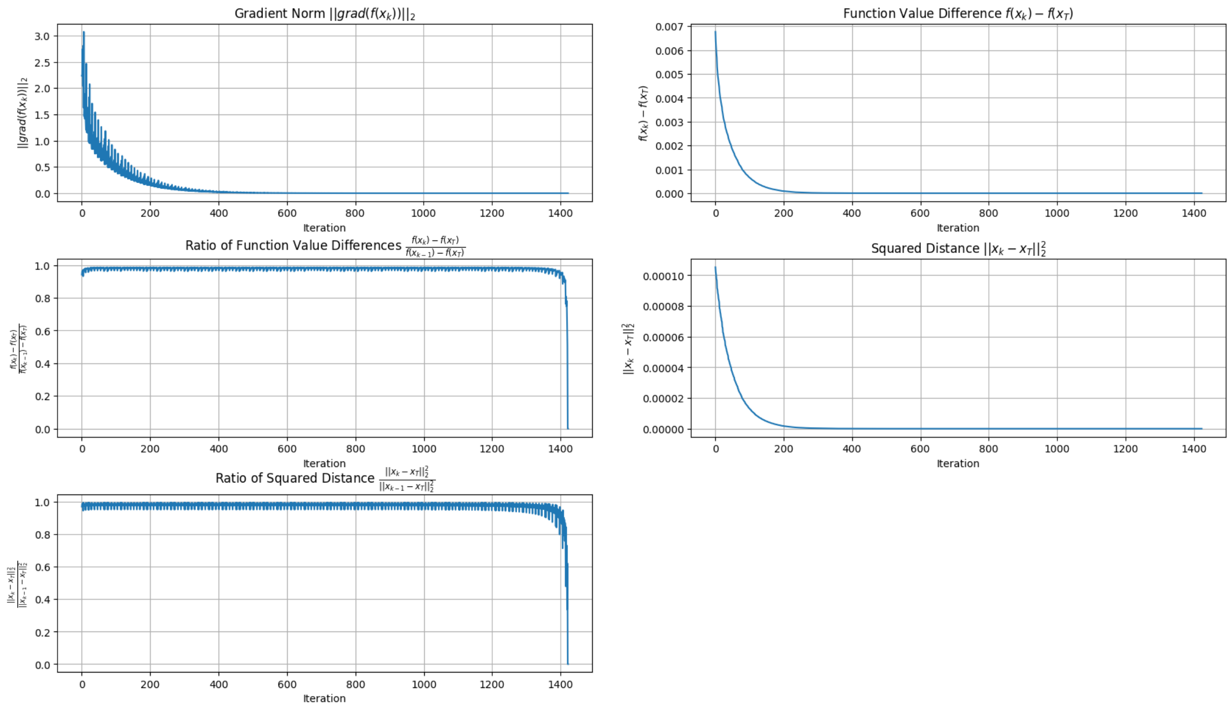
Name: Aneesh Panchal
SR No: 06-18-01-10-12-24-1-25223
Email ID: aneeshp@iisc.ac.in
Date: September 13, 2024
Assignment No: Assignment 1
Course Code: E0230
Course Name: Computational Methods of Optimization
Term: AUG 2024



(c). **InExactLineSearch:**
Given and assumed parameters are as follows,
$x_0 = [0, 0, 0, 0, 0], \quad c_1 = 0.5, \quad c_2 = 0.5, \quad \gamma = 0.5, \quad MaxIterations, T = 10000$

NOTE: Additionally we assumed a condition on α updation that if α becomes 0 (due to machine precision) then break the iterations as it will again give the same values of x and $f(x)$. In our case, $\alpha \to \alpha_0 2^{-1075}$ (machine precision limit).

The final results are as follows,
$x* : [-3.96463288e-05, -5.00000000e-04, -1.00000000e-03, -2.00000000e-03, -9.99999513e-03]$
$f(x*) : -0.0067698231772576585$
$Iterations : 1423$



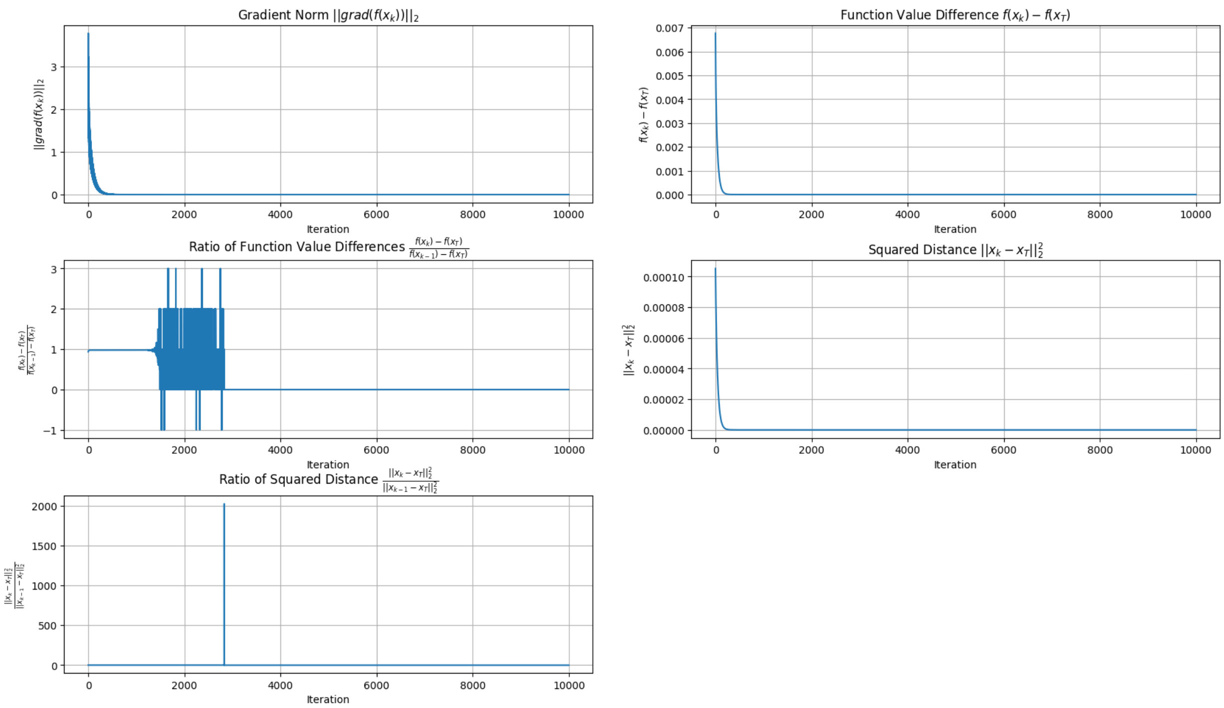
Analysis of InExactLineSearch:
Values of c_1, c_2 and γ are very crucial for accurate results.

c_1	c_2	γ	$f(x*)$	$Iterations$
0.5	0.5	0.5	-0.0067698231772576585	1423
0.75	0.25	0.5	0.0	1
0.25	0.75	0.5	-0.0067698231772588095	1699
0.5	0.5	0.1	-0.000355885	2
0.75	0.25	0.1	0.0	1
0.25	0.75	0.1	-0.00061542986977165	3

(d). **ExactLineSearch:**
Given initial value of x is as follows, $x_0 = [0, 0, 0, 0, 0]$

How to find $p_k^T A p_k$:
For finding $p_k^T A p_k$, a simple analysis is done which is as follows,
As $f(x) = \frac{1}{2}x^T Ax + b^T x$ and hence, $\nabla f(0) = b$
As $p_k = -\nabla f(x_k)$ and hence we can use, $f(p_k) = \frac{1}{2}p_k^T A p_k + \nabla f(0)^T p_k$
And hence, $p_k^T A p_k = 2\left(f(p_k) - \nabla f(0)^T p_k\right)$

The final results are as follows,
$x* : [-3.96463545e-05, -5.00000000e-04, -1.00000000e-03, -2.00000000e-03, -1.00000000e-02]$
$f(x*) : -0.00676982317725885$
$Iterations : 10000$



Name:	Aneesh Panchal	**Assignment No:**	Assignment 1
SR No:	06-18-01-10-12-24-1-25223	**Course Code:**	E0230
Email ID:	aneeshp@iisc.ac.in	**Course Name:**	Computational Methods of Optimization
Date:	September 13, 2024	**Term:**	AUG 2024

Solution 3

1. We have given function, $f(x, y) = e^{xy}$

$$f_x(x, y) = ye^{xy}, \quad f_y(x, y) = xe^{xy}, \quad f_{xx}(x, y) = y^2e^{xy}, \quad f_{yy}(x, y) = x^2e^{xy}, \quad f_{xy} = e^{xy}(1 + xy)$$

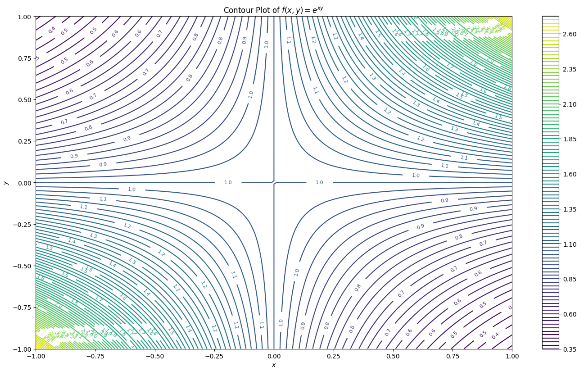
For stationary points, $f_x(x, y) = 0$ and $f_y(x, y) = 0$ i.e. $x = 0$ and $y = 0$,
For $(x, y) = (0, 0)$, $\quad f_{xx}(0, 0)f_{xx}(0, 0) - (f_{xy}(0, 0))^2 = -1 < 0$ which means is saddle point.
Hence, the only stationary point is $(\mathbf{x}, \mathbf{y}) = (\mathbf{0}, \mathbf{0})$ which is **Saddle Point**.

2. As we know generally gradient descent is of the form, $x_{i+1} = x_i - \alpha\nabla(f(x_i))$. Hence,

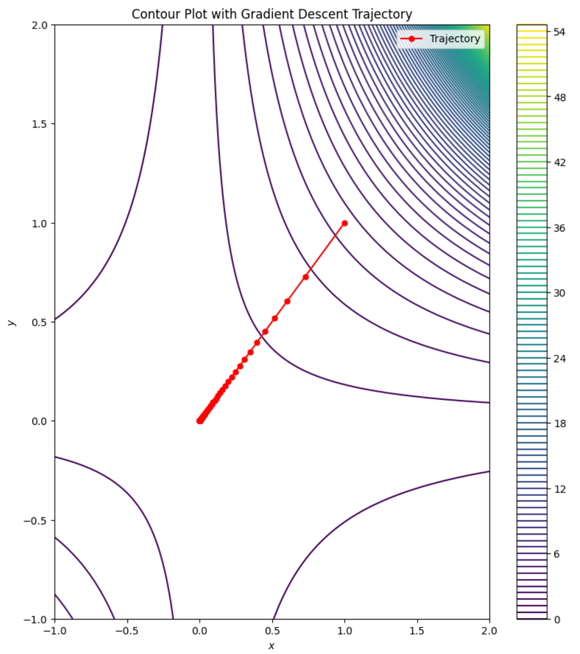
$$[x_{i+1}, y_{i+1}]^T = [x_i, y_i]^T - \alpha[f_x(x, y), f_y(x, y)]^T \quad \equiv \quad [x_{i+1}, y_{i+1}]^T = [x_i - (\alpha e^{x_i y_i})\, y_i, y_i - (\alpha e^{x_i y_i})\, x_i]^T$$

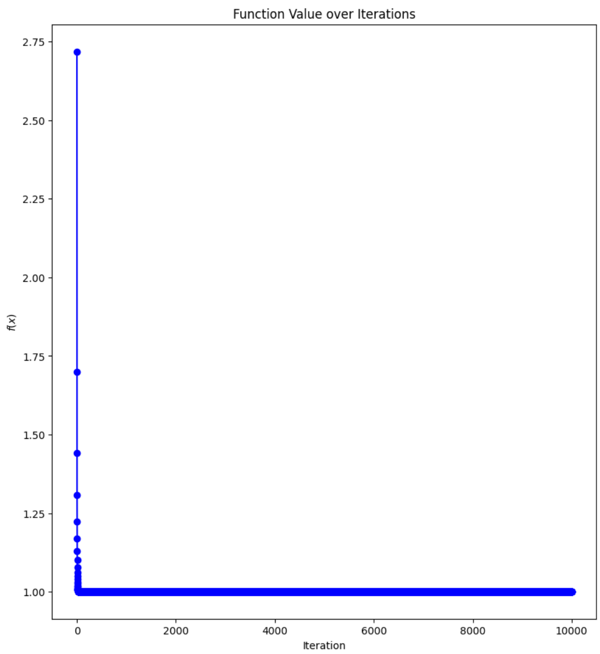
So when $x_i = y_i$, it is equivalent to,
$[x_{i+1}, y_{i+1}]^T = x_i \left(1 - \alpha e^{x_i^2}\right) [1, 1]^T$
Hence, when the initial point is on line $x = y$ then the next iterations will also stays on line $x = y$.

3. Contour Plot for equation $f(x, y) = e^{xy}$ with $-1 \le x \le 1$ and $-1 \le y \le 1$.



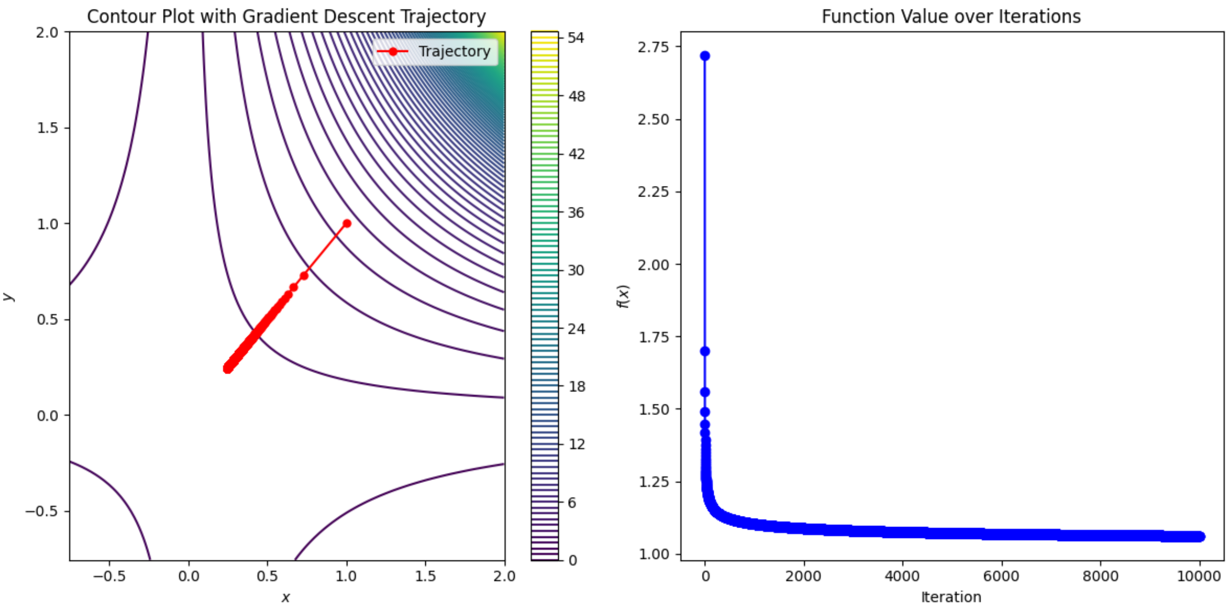
4. Parameters to be consider for the **Gradient Descent with fixed step size** are as follows,
$(x_0, y_0) = (1, 1), \quad \alpha = 0.1, \quad iterations = 10000$



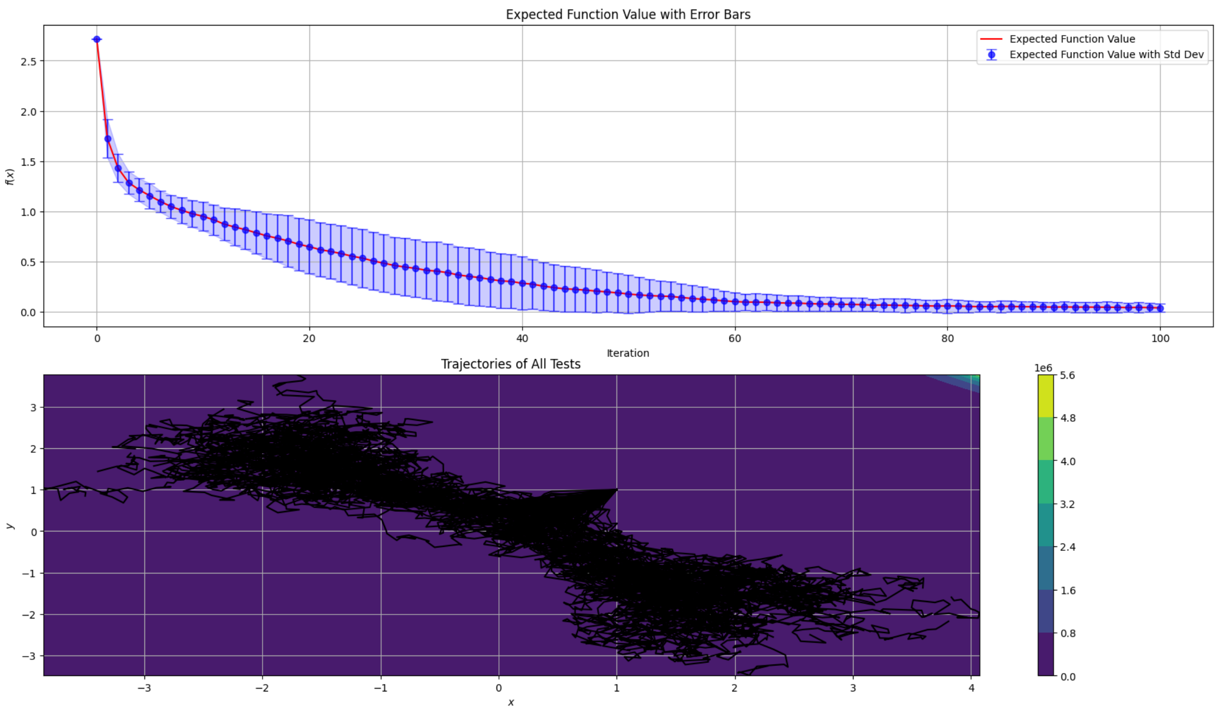


Name: Aneesh Panchal
SR No: 06-18-01-10-12-24-1-25223
Email ID: aneeshp@iisc.ac.in
Date: September 13, 2024
Assignment No: Assignment 1
Course Code: E0230
Course Name: Computational Methods of Optimization
Term: AUG 2024

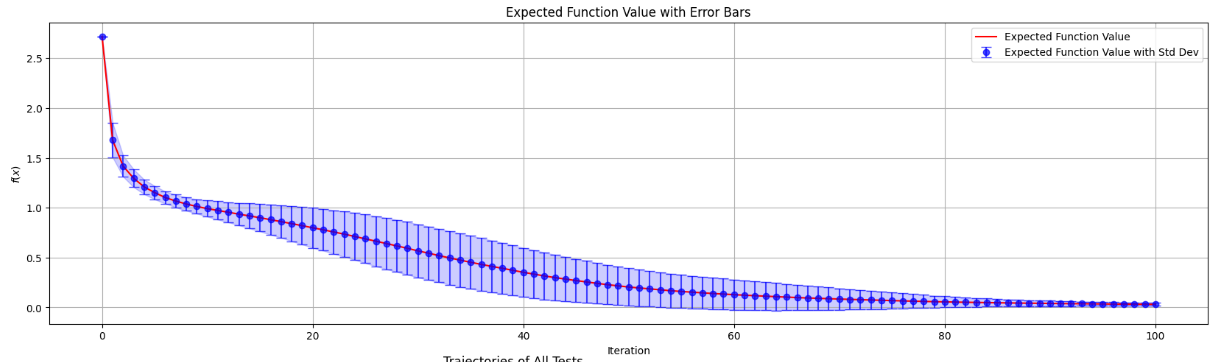
5. Parameters to be consider for the **Gradient Descent with diminishing step size** are as follows,
$(x_0, y_0) = (1, 1), \quad \alpha_0 = 0.1, \quad iterations = 10000$

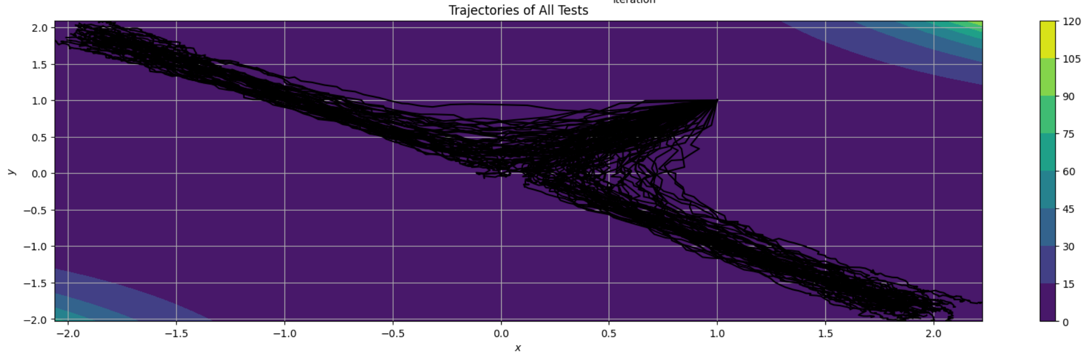


6. Parameters to be consider for the **Gradient Descent with fixed step size and fixed noise variance** are as follows,
$(x_0, y_0) = (1, 1), \quad \alpha = 0.1, \quad iterations = 100, \quad Number of Experiments = 100, \quad \sigma^2 = 1$

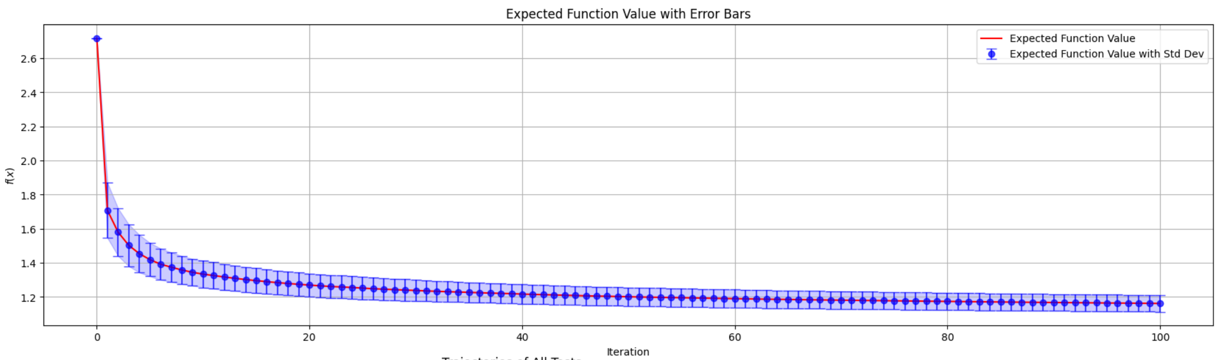


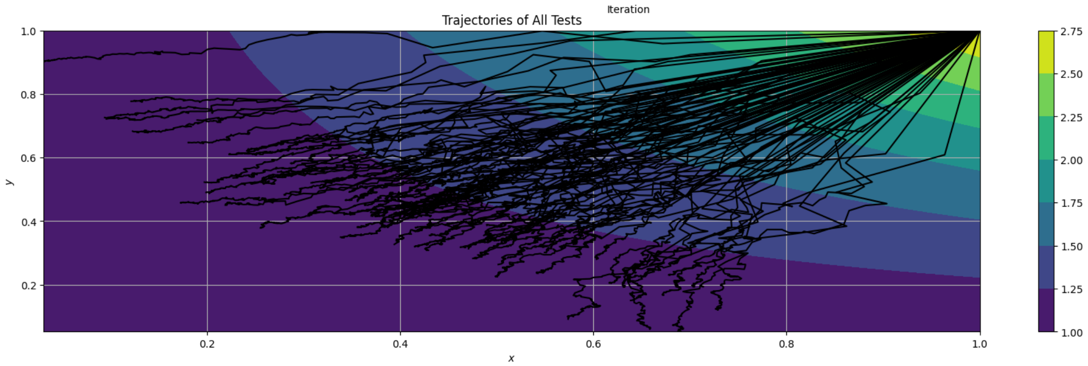
7. Parameters to be consider for the **Gradient Descent with fixed step size and diminishing noise variance** are as follows, $(x_0, y_0) = (1, 1), \quad \alpha = 0.1, \quad iterations = 100, \quad Number of Experiments = 100, \quad \sigma_0^2 = 1$





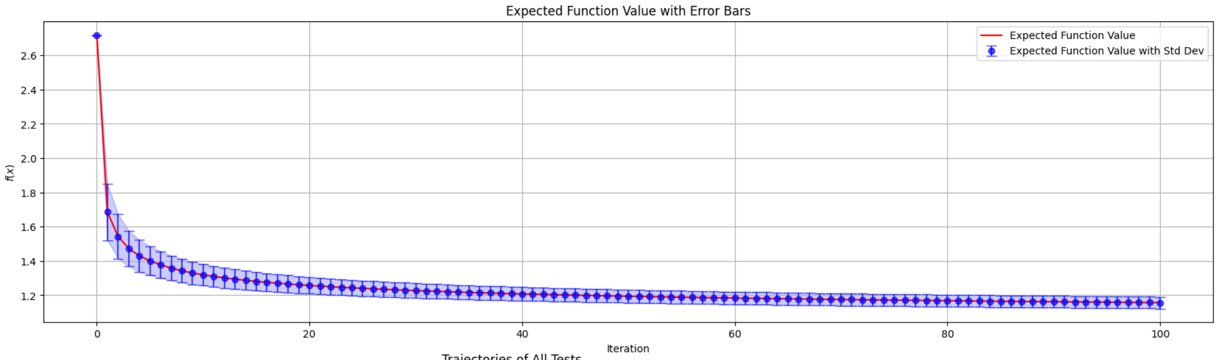
8. Parameters to be consider for the **Gradient Descent with diminishing step size and fixed noise variance** are as follows, $(x_0, y_0) = (1, 1), \quad \alpha_0 = 0.1, \quad iterations = 100, \quad Number of Experiments = 100, \quad \sigma^2 = 1$

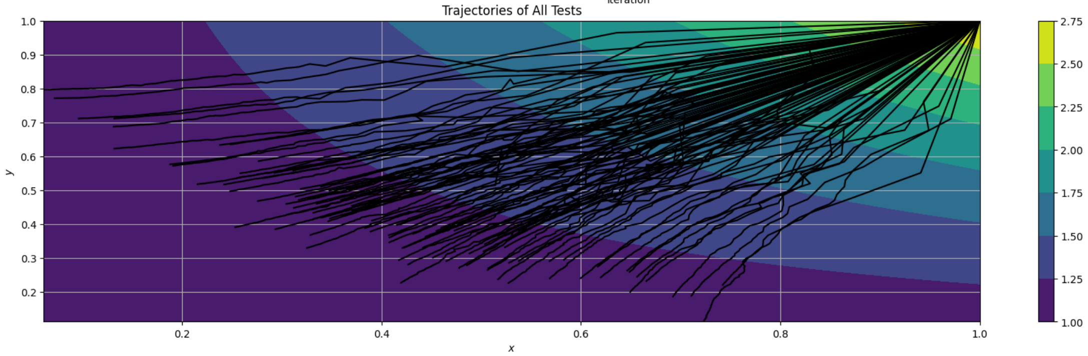




Name:	Aneesh Panchal	**Assignment No:**	Assignment 1
SR No:	06-18-01-10-12-24-1-25223	**Course Code:**	E0230
Email ID:	aneeshp@iisc.ac.in	**Course Name:**	Computational Methods of Optimization
Date:	September 13, 2024	**Term:**	AUG 2024

9. Parameters to be consider for the **Gradient Descent with diminishing step size and diminishing noise variance** are as follows,
$(x_0, y_0) = (1, 1), \quad \alpha_0 = 0.1, \quad iterations = 100, \quad NumberofExperiments = 100, \quad \sigma_0^2 = 1$





10. We have given relationship, $\quad \theta^{(t+1)} = \theta^{(t)} - \alpha_t \left(\nabla f\left(\theta^{(t)}\right) + \zeta^{(t)}\right) \quad$ where, $\zeta^{(t)} \sim \mathcal{N}\left(0, \sigma_t^2 I\right)$
Let us assume that $f \in C^1$, then we get,

$$f\left(\theta^{(t+1)}\right) = f\left(\theta^{(t)} - \alpha_t\left(\nabla f\left(\theta^{(t)}\right) + \zeta^{(t)}\right)\right) = f\left(\theta^{(t)}\right) - \left(\alpha_t\left(\nabla f\left(\theta^{(t)}\right) + \zeta^{(t)}\right)\right)^T \nabla f\left(\theta^{(t)}\right) \tag{1}$$

As we know Law of Total Expectation is given as, If X and Y are random variable sampled from same space then we have,

$$\mathbb{E}[X] = \mathbb{E}\left[\mathbb{E}[X|Y]\right] \tag{2}$$

Hence, we can rewrite equation (1) as,

$$\mathbb{E}\left[f\left(\theta^{(t+1)}\right)\right] = \mathbb{E}\left[\mathbb{E}\left[f\left(\theta^{(t+1)}\right)\middle|\theta^{(t)}\right]\right] \tag{3}$$

Now as we have given $\theta^{(t)}$, hence we can assume that $\theta^{(t)}$, $f\left(\theta^{(t)}\right)$ and $\nabla f\left(\theta^{(t)}\right)$ as constant. Hence,

$$\mathbb{E}\left[f\left(\theta^{(t+1)}\right)\right] = \mathbb{E}\left[\mathbb{E}\left[f\left(\theta^{(t)}\right) - \left(\alpha_t\left(\nabla f\left(\theta^{(t)}\right) + \zeta^{(t)}\right)\right)^T \nabla f\left(\theta^{(t)}\right)\right]\right]$$

$$\mathbb{E}\left[f\left(\theta^{(t+1)}\right)\right] = \mathbb{E}\left[f\left(\theta^{(t)}\right) - \alpha_t\left(\nabla f\left(\theta^{(t)}\right)\right)^T \nabla f\left(\theta^{(t)}\right) - \alpha_t \mathbb{E}\left[\zeta^{(t)}\right]^T \nabla f\left(\theta^{(t)}\right)\right] \tag{4}$$

As we know $\zeta^{(t)} \sim \mathcal{N}\left(0, \sigma_t^2 I\right)$. Hence, $\mathbb{E}\left[\zeta^{(t)}\right] = 0$,

$$\mathbb{E}\left[f\left(\theta^{(t+1)}\right)\right] = \mathbb{E}\left[f\left(\theta^{(t)}\right) - \alpha_t\left(\nabla f\left(\theta^{(t)}\right)\right)^T \nabla f\left(\theta^{(t)}\right)\right]$$

Finally, the required relation is as follows,

$$\mathbb{E}\left[f\left(\theta^{(t+1)}\right)\right] = \mathbb{E}\left[f\left(\theta^{(t)}\right)\right] - \alpha_t \mathbb{E}\left[\left(\nabla f\left(\theta^{(t)}\right)\right)^T \nabla f\left(\theta^{(t)}\right)\right] \tag{5}$$

Solution 4

1. As given function is polynomial of degree 4 and hence, is 4 times continuously differentiable,

$$f(x) = x(x-1)(x-3)(x+2) = x^4 - 2x^3 - 5x^2 + 6x$$

$$f'(x) = 4x^3 - 6x^2 - 10x + 6 \implies f'(x) = 2(2x-1)(x^2 - x - 3)$$
$$f''(x) = 12x^2 - 12x - 10 \implies f''(x) = 2(6x^2 - 6x + 5)$$

Extremum points are (for which $f'(x) = 0$),

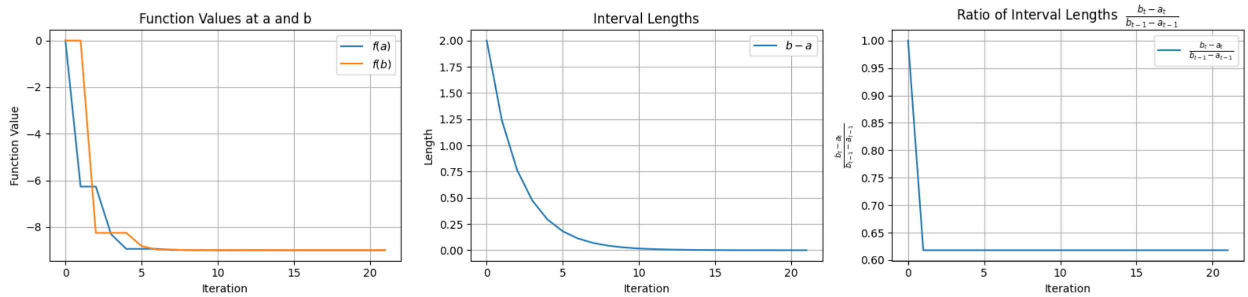
$$x = \frac{1-\sqrt{13}}{2}\ (x_1,\ \text{say}),\ \frac{1}{2}\ (x_2,\ \text{say}),\ \frac{1+\sqrt{13}}{2}\ (x_3,\ \text{say})$$

Out of which $f''(x_1) > 0, f''(x_2) < 0$ and $f''(x_3) > 0$.
Hence, x_1 and x_3 are minima and x_2 is local maxima.

2. (a) **Golden Section Search:**
Parameters to be consider for the **Golden Section Search** are as follows,
$a = 1, \quad b = 3, \quad thresold = 10^{-4}$

Hence the resulting values are as follows,
Number of Iterations required: **21**
Optimal x value: **2.3027522726894833**
Function Value: $-$**8.999999992903064**

Iteration	a	b	$\lvert b-a\rvert$
0	1	3	2
1	1.7639320225002102	3	1.2360679774997898
2	1.7639320225002102	2.5278640450004204	0.7639320225002102
3	2.055728090000841	2.5278640450004204	0.4721359549995796
4	2.23606797749979	2.5278640450004204	0.2917960675006306
5	2.23606797749979	2.416407864998738	0.18033988749894814
6	2.23606797749979	2.347524157501472	0.111456180001682
7	2.278640450004206	2.347524157501472	0.06888370749726569
8	2.278640450004206	2.3212129225086224	0.042572472504416314
9	2.2949016875157726	2.3212129225086224	0.02631123499284982
10	2.2949016875157726	2.3111629250273396	0.01626123751156694
11	2.2949016875157726	2.3049516849970555	0.01004999748128288
12	2.298740444966772	2.3049516849970555	0.006211240030283616
13	2.3011129275460562	2.3049516849970555	0.003838757450999264
14	2.3011129275460562	2.3034854101253406	0.002372482579284352
15	2.3020191352536257	2.3034854101253406	0.0014662748717149121
16	2.302579202417771	2.3034854101253406	0.0009062077075694397
17	2.302579202417771	2.303139269581916	0.0005600671641450283
18	2.302579202417771	2.3029253429611956	0.00034614054342441136
19	2.302711416340475	2.3029253429611956	0.00021392662072061697
20	2.302711416340475	2.3028436302631787	0.0001322139227037944
21	2.302711416340475	2.3027931290384918	8.171269801682257e-05



(b) **Fibonacci Search Method:**
Parameters to be consider for the **Fibonacci Search Method** are as follows,
$a = 1, \quad b = 3, \quad thresold = 10^{-4}, \quad T = 21$

Hence the resulting values are as follows,
Number of Iterations required: **21**
Optimal x value: **2.3027950310559007**
Function Value: $-$**8.999999995110633**

Iteration	a	b	$\lvert b - a \rvert$
0	1	3	2
1	1.7639320225002102	3	1.2360679779158041
2	1.7639320225002102	2.5278640450004204	0.7639320220841959
3	2.055728090000841	2.5278640450004204	0.47213595583160783
4	2.23606797749979	2.5278640450004204	0.2917960662525876
5	2.23606797749979	2.416407864998738	0.1803398895790198
6	2.23606797749979	2.347524157501472	0.11145617667356733
7	2.278640450004206	2.347524157501472	0.06888371290545159
8	2.278640450004206	2.3212129225086224	0.04257246376811574
9	2.2949016875157726	2.3212129225086224	0.026311249137335846
10	2.2949016875157726	2.3111629250273396	0.016261214630779897
11	2.2949016875157726	2.3049516849970555	0.010050034506556393
12	2.298740444966772	2.3049516849970555	0.006211180124223503
13	2.3011129275460562	2.3049516849970555	0.0038388543823328902
14	2.3011129275460562	2.3034854101253406	0.002372325741891057
15	2.3020191352536257	2.3034854101253406	0.0014665286404418332
16	2.302579202417771	2.3034854101253406	0.0009057971014487798
17	2.302579202417771	2.303139269581916	0.0005607315389921652
18	2.302579202417771	2.3029253429611956	0.0003450655624561705
19	2.302711416340475	2.3029253429611956	0.00021566597653510655
20	2.302711416340475	2.3028436302631787	0.00012939958592106393
21	2.302711416340475	2.3027931290384918	8.626639061404262e-05