For all the full answer, please refer to the pages attached at the end.

HOMEWORK 5 – CS210

Print this PDF and write your solutions in the space given. Your submission should consist of a scan of this PDF with addional scanned pages that you think are necessary to better explain your answer. Remember to fill out your name and student ID. You may write your answer directly into the PDF fields.

Name KAMALIKA PODDAR **SID** [redacted]

Question	1	2	3	4	5	6	7	8
Points	10	10	10	10	10	10	20	20

1. For each of the following functions, what do the first- and second- order optimality conditions say about whether 0 is a minimizer on $\mathbb{R}$?

   (a) $f(x) = -x^4 + 5x^2 - 4$

   Ans 0 is the minimizer for first and second order optimality conditions. Pg-1

   (b) $f(x) = e^{-x^2}$

   Ans 0 can't be the minimizer as it doesn't satisfies both $f'(x) = 0$ & $f''(x) \geq 0$ condition. Pg-1

   (c) $f(x) = \frac{1}{\cos(x)}$, $-\pi/2 < x < \pi/2$

   0 is not the minimizer as it satisfies second order but doesn't satisfies first order optimality conditions. Pg-2

2. For each of the following statements, indicate whether the statement is true or false.

   T/F If a function is unimodal on a closed interval, then it has exactly one minimum on the interval.

   T/F In minimizing a unimodal function of one variable by golden section search, the point discarded at each iteration is always the point having the largest function value.

   T/F For minimizing a real-valued function of several variables, the steepest descent method is usually more rapidly convergent than Newtooon's method.

3. Determine the critical points of each of the following functions and characterize each as a minimum, maximum, or inflection point. Also determine whether each function has a global minimum or maximum on $\mathbb{R}$.

(a) $f(x) = x^5 - 10x^4 + 31x^3 - 30x^2$

- Critical points $= 0, 1.07659, 2.5468, 4.37662$.
- Maximum, minimum, Maximum, minimum resp.
- Inflection points are $0.436624, 1.85062, 3.71276$.
- No global max or min.

Pg-3

(b) $f(x) = -x^5 - 3x^4 + 5x^3 + 15x^2 - 4x - 12$

- Critical points $= -2.62899, -1.50219, 0.126215, 1.60497$.
- Minimum, Maximum, Minimum, Maximum resp.
- Inflection points $= -2.17263, -0.665224, 1.03786$.
- No global max or min.

Pg-4

(c) $f(x) = \sin(x^2), 0 \leq x \leq \pi$.

- Critical points $0, \sqrt{\pi/2}, \sqrt{\frac{3\pi}{2}}, \sqrt{\frac{5\pi}{2}}$

Pg-5

- Min, Max, min, Max resp.
- Inflection points $= 0.808252, 1.81447, 2.5223, 3.0$[illegible]
- Function has global Minimum.

4. Determine the critical points of each of the following functions and characterize each as a minimum, maximum, or saddle point. Also determine whether each function has a global minimum or maximum on $\mathbb{R}^2$.

(a) $f(x, y) = x^2 - 16xy - 8x + y^2 + 16$

- Critical points $- \left(-\frac{4}{63}, \frac{-32}{63}\right)$ is a saddle point
- No global max or min.

Pg-6

(b) $f(x, y) = x^2 - 4y^2 - 8xy(x - y - 1)$

- Critical point $- (0,0)$ $(1.29, 0.23)$
- Both are saddle points.
- No global Min or Max

Pg-7

5. Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(\mathbf{x}) = x^4 + y^4 + xy(x + y - 1).$$

(a) At what point does f attain a minimum?

At $(-1,-1)$ & $(1/4, 1/4)$ $f(x)$ attain minimum.

Pg - 8

(b) Perform one iteration of Newton's method for minimizing f using as starting point $\mathbf{x}_0 = (0, 5)^T$.

$$x_1 = \begin{bmatrix} 0.57387 \\ 3.3487 \end{bmatrix}$$

Pg - 8

(c) Explain whether this is a good or bad step and in what sense.

$\|f'(x_1)\| = 147.5940$.

The step is bad, as it does not has a derivative close to zero.

Pg 9

6. Let A be a square matrix n by n.

(a) Is $f(x) = ||Ax - b||_2^2$ a convex function? Justify your answer.

We find second derivate and prove it to be positive semi-definite matrix.
Proof is attached.

Pg-10

(b) When does $g(x) = x^T Ax + b^T x + c$, for a real constant c, is convex?

When it is ≥ 0. Explaination is attached.

Pg-10

(c) Assuming $g(x) = x^T Ax + b^T x + c$ is convex, how many iterations of Newton's method are necessary to minimize $g(x)$?

$O(\log \log (1/\epsilon))$ iterations are needed to minimize $g(x)$ using proximal Newton Method.

Pg-11

7. Suppose we are given three points (x_1, y_1), (x_2, y_2), (x_3, y_3) with distinct x values.

(a) Show that the vertex of the parabola $y = ax^2 + bx + c$ through these points is given by:

$$x = x_2 - \frac{(x_2 - x_1)^2(y_2 - y_3) - (x_2 - x_3)^2(y_2 - y_1)}{2(x_2 - x_1)(y_2 - y_3) - (x_2 - x_3)(y_2 - y_1)}$$

Derivation is proved at the end.

Page-(2)

(b) Use this formula to propose an iterative technique for minimizing a function of one variable without using any of its derivatives.

Quadratic Interpolation by Powell can be used to minimize the function without using any of its derivatives

Page (14)

(c) What happens when the three points are collinear? Does this suggest a failure mode of successive parabolic interpolation?

Explained in page number (16). Yes it suggest a failure mode of successive parabolic interpolation.

(d) Does the formula in item (a) distinguish between maxima and minima of parabolas? Does this suggest a second failure mode?

No it does not, ~~the~~ This show ~~the~~ second failure. Page (17)

8. **Programming assigment**. Write a program to find a minimum of Resenbrock's function,

$$f(x, y) = 100(y - x^2)^2 + (1 - x)^2$$

using each of the following method:

(a) Steepest descent

(b) Newton

(c) Damped Newton (Newton's method with a line search).

You should try each of the methods from each of the three starting points:

(a) $(-1, 1)$

(b) $(0, 1)$

(c) $(2, 1)$

Plot the path taken in the plane by the approximate solutions for each method from each starting point. In particular, your submission should consist of three plots, each one for each method, and the three paths (one for each starting point) in each plot.

Attached in the Pdf.

1.a $f(x) = -x^4 + 5x^2 - 4$

$f'(x) = -4x^3 + 10x = 0$

$x = 0, \pm\sqrt{\frac{5}{2}}$

$f''(x) = -12x^2 + 10$

We need to know $f''(x) \geq 0$ so for that substitute 0 in $f''(x)$, we get.

$f''(0) = 10$

∴ We know 0 is the minimizer for both first order & second order, as it satisfies both

$f'(x) = 0$ & $f''(x) \geq 0$.

1.b $f(x) = e^{-x^2}$

$f'(x) = e^{-x^2}(-2x)$

~~if~~ $x = 0$

∴ $f''(x) = 4x^2 e^{-x^2} - 2e^{-x^2}$

Substituting 0 in $f''(x)$, we get

$f''(0) = 4 \times 0 \times 1 - 2 \times 1$

$= -2$

$f''(x) \not\geq 0$. therefore 0 is not the minimizer of ~~first and~~ second order optimality condition.

0 is a minimizer for first order optimality condition.

∴ 0 can't be the minimize as it doesn't satisfies both the condition.

1.e $f(x) = \frac{1}{\cos x}$ $\quad -\pi/2 \le x \le \pi/2$

$f'(x) = \frac{\sin(x)}{\cos^2(x)}$ $\quad \therefore x = -0.819975$

~~$\therefore x = \pi n$~~

~~when $n = -\frac{\pi}{2}$ then n is~~

~~$n\pi = -\frac{\pi}{2}$~~

~~$n = -\frac{1}{2}$~~ ~~$-\frac{1}{2} < n < \frac{1}{2}$~~

~~$n \cdot \frac{\pi}{2}$~~ ~~$n\pi = \frac{\pi}{2}$~~

~~$n = \frac{1}{2}$~~

~~$\therefore$ if $n = 0$ then $x = 0$.~~

Putting x value in $f''(x)$

$$f''(x) = \frac{2\sin^2(x)}{\cos^3(x)} + \frac{1}{\cos x}$$

$$= \cancel{\frac{2\times 0}{4}} + \frac{1}{1}$$

$$f''(x) = 4.83248 \ge 0$$

0 is a not a minimizer as it doesn't satisfies first order, ~~and~~ but satisfies second order optimality conditions.

2. True :- this is by definition of unimodal.
   False :- They keep the point with largest function value
   False :- Newton's better

3.a $f(x) = x^5 - 10x^4 + 31x^3 - 30x^2$

$f'(x) = 5x^4 - 40x^3 + 93x^2 - 60x$

Equating $f'(x)$ to zero to find all the roots,

The roots are

$x = 0, 1.07659, 2.5468, 4.37662$

The above are the critical points of the function.

$f''(x) = 20x^3 - 120x^2 + 186x - 60$

$\therefore x = 0.436624, 1.85062, 3.71276$

Substituting critical points in $f''(x)$, we get

$f''(0) = -60$ (Max)

$f''(1.07659) = 26.11624$ (Min)

$f''(2.5468) = -34.2584$ (Max)

$f''(4.37662) = 132.1409$ (Min)

After plotting the graph, we get that this doesn't have a global minimum or maximum.

To find inflection points, taking random numbers in the range scale:- as 0.25, 1, 2, 4, & substituting in $f''(x)$, we get

$f''(0.25) = -20.6875$ } Concavity changes

$f''(1) = 26$ } change

$f''(2) = -8$ } change

$f''(4) = 44$

Inflection points are $\{0.436624, 1.85062, 3.71276\}$

3.b $f(x) = -x^5 - 3x^4 + 5x^3 + 15x^2 - 4x - 12$ (4)

$f'(x) = -5x^4 - 12x^3 + 15x^2 + 30x - 4$

Critical points are $= -2.62899, -1.50219, 0.126215, 1.60497$.

$f''(x) = -20x^3 - 36x + 30x + 30$.

$x = -2.17263, -0.665224, 1.03786$.

$\therefore f''(\cancel{x}-2.62899) = 65.7231$ (Min)

$f''(-1.50219) = -28.5063$ (Max)

$f''(0.126215) = 33.1727$ (Min)

$f''(1.60497) = -97.2701$ (Max)

There is no global minimum or maximum in $f(x)$.

To find inflection points we take

$-3, -1, 1, 2$

$$\left.\begin{aligned} \therefore f''(-3) &= 156 \\ f''(-1) &= -16 \end{aligned}\right\} \text{change}$$

$$\left.\begin{aligned} f''(-1) &= -16 \\ f''(1) &= 4 \end{aligned}\right\} \text{change}$$

$$\left.\begin{aligned} f''(1) &= 4 \\ f''(2) &= -214 \end{aligned}\right\} \text{change}$$

$\therefore$ the inflection points are $(-2.17263, -0.665224, 1.03786)$

3.c $f(x) = \sin(x^2), \quad 0 \leq x \leq \pi$

$f'(x) = 2x \cos x^2$

$x = 0, \sqrt{\frac{\pi}{2}}, \cancel{\sqrt{2\pi}}, \sqrt{\frac{3\pi}{2}}, \sqrt{\frac{5}{2}}$

$f''(x) = 2\cos(x^2) - 4x^2 \sin(x^2)$

$x = 0.808252, 1.81447, 2.52223, 3.07855$

$f''(0) = 2$ (Min)

$f''\left(\sqrt{\frac{\pi}{2}}\right) = -2\pi$ (Max)

$f''\left(\sqrt{\frac{3\pi}{2}}\right) = 6\pi$ (Min)

$f''\left(\sqrt{\frac{5\pi}{2}}\right) = -10\pi$ (Max)

The function $f(x)$ has a global minimum after we plot it, we get the obtained graph.

To find inflection points:-

take the points as ~~0, 1, 2, 3, 4~~ 0.5, 1.5, 2.1, 2.9, 3.25 we get,

$f''(0.5) = 1.69042$

$f''(1.5) = -8.25901$

$f''(2.1) = 16.2444$

$f''(2.9) = -29.6282$

$f''(3.25) = 37.5101$

The inflection points are — 0.80852, 1.81447, 2.52223, 3.07855

4.a $f(x,y) = x^2 - 16xy - 8x + y^2 + 16$

$f_x = 2x - 16y - 8$

$f_y = -16x + 2y$

$2x - 16y - 8 = 0$

$-16x + 2y = 0$

Find Critical Points

$y = 8x$

$x = -\frac{4}{63}, \quad y = -\frac{32}{63}$

Critical Point is $\left(-\frac{4}{63}, -\frac{32}{63}\right)$

$f\left(-\frac{4}{63}, -\frac{32}{63}\right) = 16.2539$

$f_{xx} = 2$

$f_{yy} = 2$

$f_{xy} = -16$

$f_{xx} f_{yy} - (f_{xy})^2 = 2 \times 2 - (16)^2$

$= 4 - 256$

$= -252 < 0$

$\therefore \left(-\frac{4}{63}, -\frac{32}{63}, 16.2539\right)$ is a saddle point

There is ~~not~~ no global minimum or maximum for the function $f(x,y)$.

4.6 $f(x,y) = x^2 - 4y^2 - 8xy(x-y-1)$ (8)

$f(x,y) = x^2 - 4y^2 - 8x^2y + 8xy^2 + 8xy$

$f_x = 2x - 16xy + 8y^2 + 8y$

$f_y = -8y - 8x^2 + 16xy + 8x$

$2x - 16xy = 8y^2 + 8y = 0$

$-8y - 8x^2 + 16xy + 8x = 0$

$x = \frac{4y(y-1)}{(1-8y)}$

$y = \frac{(x^2 - x)}{(2x-1)}$

Critical points are $(0,0)$ & $(1.29, 0.23)$

$\therefore f(1.29, 0.23) = 1.310084$

$f_{xx} = 2 - 16y$

$f_{yy} = -8 + 16x$

$f_{xy} = -16x + 16y + 8$

$f_{xx} f_{yy} - (f_{xy})^2$

$= (2-16y)(-8+16x) - (-16x + 16y + 8)^2$

for $f(0,0) = (2-0)(-8+0) - (-0+0+8)^2$

$= -80$

$f(1.29, 0.23) = -101.5168$

Therefore at critical point $(0,0)$ & $(1.29, 0.23)$ the surface has saddle points and the function has no global min & global max.

5.a $f(x) = x^4 + y^4 + xy(x+y-1)$ ⑧

$f(x) = x^4 + y^4 + x^2y + xy^2 - xy$

$f_x = 4x^3 + 2xy + y^2 - y$

$f_y = 4y^3 + x^2 + 2xy - x$

$4x^3 + 2xy + y^2 - y = 0$

$4y^3 + x^2 + 2xy - x = 0$

$x = -1, 0, 1/4$

$y = -1, 0, 1/4$

Therefore the critical points are $(-1,-1), (0,0), (1/4, 1/4)$

$D = f_{xx} f_{yy} - f_{xy}^2$

$f_{xx} = 12x^2 + 2y \qquad f_{xy} = 2x + 2y - 1$

$f_{yy} = 12y^2 + 2x$

$\therefore (12x^2 + 2y)(12y^2 + 2x) - (2x + 2y - 1)^2$

$f(-1,-1) = 75$ — (Min)

$f(0,0) = -1$ — (Max)

$f(1/4, 1/4) = 25/16$ (Min)

Thus the minimum is at ~~(-1~~ $(-1,-1)$ & $(1/4, 1/4)$

5.b The Hessian Matrix is

$$H_f(x) = \begin{bmatrix} 12x^2 + 2y & 2x + 2y - 1 \\ 2x + 2y - 1 & 12y^2 + 2x \end{bmatrix}$$

We need $H(x_k) s_k = -\nabla f(x_k)$

$$\begin{bmatrix} 12x^2+2y & 2x+2y-1 \\ 2x+2y-1 & 12y^2+2x \end{bmatrix} S_0 = \begin{bmatrix} -20 \\ -500 \end{bmatrix}$$ (9)

$\therefore$ we get after solving as

$$S_0 = \left(\frac{-500}{973}, \frac{-4820}{2919}\right)$$

$$x_1 = x_0 + S_0 = \begin{bmatrix} 0 \\ 5 \end{bmatrix} + \begin{bmatrix} \frac{-500}{973} \\ \frac{-4820}{2919} \end{bmatrix} = \begin{bmatrix} -0.51387 \\ 3.3487 \end{bmatrix}$$

<u>5.c</u> $f'(x_0) = f'(0,5)$

$$\|f'(0,5)\| = \|(20, 500)\| = 500.3998$$

$$\|f'(x_1)\| = \|f'(-0.51387, 3.3487)\|$$

$$= \|3.88072, 147.543\|$$

$$= 147.5940$$

The step is bad, as it does not has a derivative close to zero.

(10)

6.a To prove $f(x) = \|Ax - b\|_2^2$ a convex function

~~First derivative is $\nabla f(x) = 2A^T(Ax-b)$,~~

~~& 2nd derivative $= 2A^TA$~~

$\because$ we have $\|Ax-b\|^2 = \|b\|^2 - 2b^TAx + \|Ax\|^2$

Differentiating, we get,

$$\frac{d f(x)}{dx} = -2b^TA + 2(Ax)^TA$$

$$= -2b^TA + 2x^TA^TA$$

$$= 2A^T(Ax^T - b)$$

$$\therefore \frac{d^2 f(x)}{dx^2} = 2A^TA,$$ which is a positive semi-definite matrix.

Thus we can say $f(x)$ is a convex function.

6.b $g(x) = x^TAx + b^Tx + c$,

A function is convex, if $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$ for all $\lambda \in [0,1]$

It suffices to show for a quadratic function $f(x) = x^TQx$,

P.T.O →

Therefore using the definition of a convex function:

$$(\lambda x+(1-\lambda)y)^T Q(\lambda x+(1-\lambda)y) \leq \lambda x^T Q x+(1-\lambda)y^T Q y$$

Equality holds for $\lambda = 0$ or 1. Therefore considering $\lambda \in (0,1)$.

The left hand side simplifies to:

$$\lambda^2 x^T Q x + (1-\lambda)^2 y^T Q y + \lambda(1-\lambda) x^T Q y + \lambda(1-\lambda) y^T Q x$$

$$\leq \lambda x^T Q x + (1-\lambda) y^T Q y$$

Rearranging the terms and simplifying we obtain:

$$\lambda(1-\lambda) x^T Q x + \lambda(1-\lambda) y^T Q y - \lambda(1-\lambda) x^T Q y - \lambda(1-\lambda) y^T Q x \geq 0$$

$$\Rightarrow x^T Q x + y^T Q y - x^T Q y - y^T Q x \geq 0$$

$$\Rightarrow (x-y)^T Q (x-y) \geq 0$$

which is true for positive semi-definite $Q \geq 0$.

<u>6.e</u> We should perform $x^{(k+1)} = x^{(k)} - F(x^{(k)})^{-1} g^{(k)}$ iteration to minimize the $g(x)$ by Newton's method.

(Using Proximal Newton method, we require $O(\log\log(1/\epsilon))$ iterations to minimize $g(x)$)

7.a (18)

Points $(x_1, y_1), (x_2, y_2), (x_3, y_3)$

$$y = ax^2 + bx + c$$

Substitute (x_1, y_1) we get

$$\rightarrow y_1 = ax_1^2 + bx_1 + C$$

$$(x_2, y_2) \Rightarrow y_2 = ax_2^2 + bx_2 + c$$

$$(x_3, y_3) \Rightarrow y_3 = ax_3^2 + bx_3 + C$$

Solving for a, b, c, we get.

$$a = \frac{(x3(y2-y1) + x^2(y1-y3) + x1(y3-y2))}{(x1-x_2)(x1-x_3)(x_2-x_3)}$$

$$b = \frac{(x_3^2(y_1-y_2) + x_2^2(y_3-y_1) + x_1^2(y_2-y_3))}{(x_1-x_2)(x_1-x_3)(x_2-x_3)}$$

$$c = \frac{x_2 x_3(x_2-x_3) y_1 + x_3 x_1(x_3-x_1) y_2 + x_1 x_2(x_1-x_2) y_3}{(x_1-x_2)(x_1-x_3)(x_2-x_3)}$$

Vertex is a point ~~f~~ where first derivative is zero.

$$x = -\frac{b}{2a} \quad \& \quad y = \frac{c-b^2}{4a}$$

$$\therefore x = \cancel{x_3^2(y_1-y} \; \frac{x_3^2(y_2-y_1) + x_2^2(y_1-y_3) + x_1^2(y_3-y_2)}{2(x_3(y_2-y_1) + x_2(y_1-y_3) + x_1(y_3-y_2)}$$

(13)

To prove we have

$$x = \frac{x_2 - (x_2-x_1)^2(y_2-y_3) - (x_2-x_3)^2(y_2-y_1)}{2(x_2-x_1)(y_2-y_3) - (x_2-x_3)(y_2-y_1)}$$

In a

$$= 2x_2(x_2-x_1)(y_2-y_3) - x_2(x_2-x_3)(y_2-y_1) - (x_2-x_1)^2(y_2-y_3) -$$

$$\therefore x = \frac{x_3^2y_2 - x_3^2y_1 + x_2^2y_1 - x_2^2y_3 + x_1^2y_3 - x_1^2y_2}{2x_3y_2 - 2x_3y_1 + 2x_2y_1 - 2x_2y_3 + 2x_1y_3 - 2x_1y_2}$$

$$= \frac{[x_2^2y_2 - 2x_2x_1y_2 + x_1^2y_2 - x_2^2y_3 - 2x_2x_1y_3 + x_1^2y_3] - x_2^2y_2 + 2x_2x_1y_2 - 2x_1^2y_2 + 2x_2x_1y_3 + x_3^2y_2 - x_3^2y_1 + x_2^2y_1}{2x_3y_2 - 2x_3y_1 + 2x_2y_1 - 2x_2y_2 + 2x_2y_2 - 2x_2y_3 + 2x_1y_3 - 2x_1y_2}$$

$$= \frac{(x_2-x_1)^2(y_2-y_3) - 2x_2y_2 \ldots - 2x_2^2y_2 - 2x_2x_1y_2 + 2x_2^2y_1 - 2x_2x_3y_1}{2(2x_2(y_2-y_3) - 2x_1(y_2-y_3)) - 2(x_2y_2 - x_3y_2 - x_2y_1 + x_3y_1)}$$

$(x_2-x_1)^2(y_2-y_3)$

$x_2^2 - 2x_2x_1 + x_1^2$

$x_2^2y_2 - 2x_2x_1y_2 + x_1^2y_2$

$-x_2^2y_3 - 2x_2x_1y_3 + x_1^2y_3$

$-2x_2y_3 + 2x_1y_3$

$-2x_1y_2$

$2x_2y_2$

$$= \frac{(x_2-x_1)^2(y_2-y_3) - (x_2-x_3)^2(y_2-y_1) + 2x_3^2y_2 - 2x_2x_3y_2 - 2x_3^2y_1 - 2x_2x_3y_1 + 2x_2x_1y_2 - 2x_1^2y_2 + 2x_2x_1y_3}{2(x_2-x_1)(y_2-y_3) - 2(y_2-y_1)(x_2-x_3)}$$

$-2x_2y_2$

$x_2y_2 - x_3y_2 - x_2y_1 + x_3y_1$

$(x_2-x_3)(y_2-y_1)$

$(x_2^2 + x_3^2 - 2x_2x_3)(y_2-y_1)$

$x_2^2y_2 - x_3^2y_2 + 2x_2x_3y_1 - 2x_2x_3y_2 + x_3^2y_1 - 2x_2x_3y_1$

$$\text{num} = 2x_3^2 y_2 - 2x_3^2 y_1 - 2x_1^2 y_2 - 2x_2 x_3 y_2 - 2x_2 x_3 y_1 + 2x_2 x_1 y_2 + 2x_2 x_1 y_3$$

$$2x \; Qy^2 (\cancel{x_3^2 x} - x_1^2$$

= Proved

$$x = x_2 - \frac{(x_2 - x_1)^2 (y_2 - y_3) - (x_2 - x_3)^2 (y_2 - y_1)}{2(x_2 - x_1)(y_2 - y_3) - (x_2 - x_3)(y_2 - y_1)}$$

7.b Quadratic Interpolation was first proposed by Powell and it uses values of the function f to be minimized at three points to fit a parabola,

$$y = ax^2 + bx + c \text{ through those points.}$$

The method starts with an initial point say $x = 0$, with a function value $p_0 = f(x_0)$ & at step size β. two more function evaluations are performed as described in the following steps to determine the points for the polynomial fit. In general, however, we start with a situation where we have already bracketed the minimum between $x_1 = x_l$ and $x_2 = x_u$ by using the bracketing method described earlier. In that case we will only need an intermediate point x_0 in the interval (x_l, x_u)

(13)

1. Evaluate $p_1 = p(\beta) = f(x_0 + \beta s)$
2. If $p_1 < p_0$, then evaluate $p_2 = p(2\beta) = f(x_0 + 2\beta s)$. Otherwise evaluate $p_2 = p(-\beta) = f(x_0 - \beta s)$. The constants a, b and c in equation can now be uniquely expressed in terms of the function values p_0, p_1 and p_2 as:

$$a = p_0$$

$$b = \frac{4p_1 - 3p_0 - p_2}{2\beta} \;\&\; c = \frac{p_2 + p_0 - 2p_1}{2\beta^2}, \text{ if } p_2 = f(x_0 + 2\beta s)$$

3. The value of $\alpha = \alpha^*$ at which $p(\alpha)$ is extremized for the current cycle is then given by

$$\alpha^* = -\frac{b}{2c}$$

4. α^* corresponds to a minimum of p if $c > 0$, and the prediction ~~based on~~ is repeated using $(x_0 + \alpha^* s)$ as the initial point for the next cycle with p_0 of $(x_0 + \alpha^* s)$ until the desired accuracy is obtained.
5. If the point $\alpha = \alpha^*$ corresponds to a maximum of p rather than a minimum, or if it corresponds to a minimum of p which is at a distance greater than a prescribed maximum β_{max}, then the maximum allowed step is taken in the direction of decreasing f and the point furthest away from this new point is discarded in order to repeat the process.

7.c ~~Since, it is stipulated that the quadratic function implies that no pair of the three~~

The condition that our parabola be realizable as a graph ~~graph~~ quadratic function implies that no pair of the three points lie on the same vertical line. Moreover, since it is stipulated that a quadratic function be nontrivial, we cannot have three co-linear points. But suppose the parabola contains three points which lie on a common line. Then this line has an equation of the form $y = dx + e$ for some $d, e \in \mathbb{R}$. Since the line is coincident with the parabola ~~d~~ at three points, we expect the difference of the quadratic and the line to have three distinct roots. This is impossible, since the resulting difference

$$ax^2 + (b-d)x + c - e = 0$$

is a quadratic provided $a \neq 0$, when it has at most two distinct solutions.

This suggest a failure mode of successive parabolic interpolation.

17

7d A vertex has co-ordinates as

$$\cancel{P = \left(-\frac{b}{2a}, c - \frac{b^2}{4a} \right)}$$

$$= P = \left(-\frac{b}{2a}, f\left(\frac{-b}{2a}\right) \right)$$

The second co-ordinate of the vertex will detect the maximum or minimum value of $f(x)$.

Therefore the formula does distinguish between maxi & minimum, therefore it is a second failure mode.

8 Answer is attached ~~as a~~ with the code.

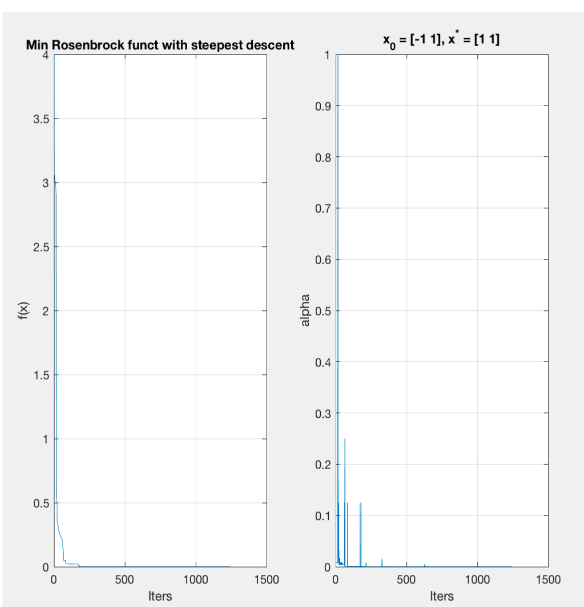
CODE:

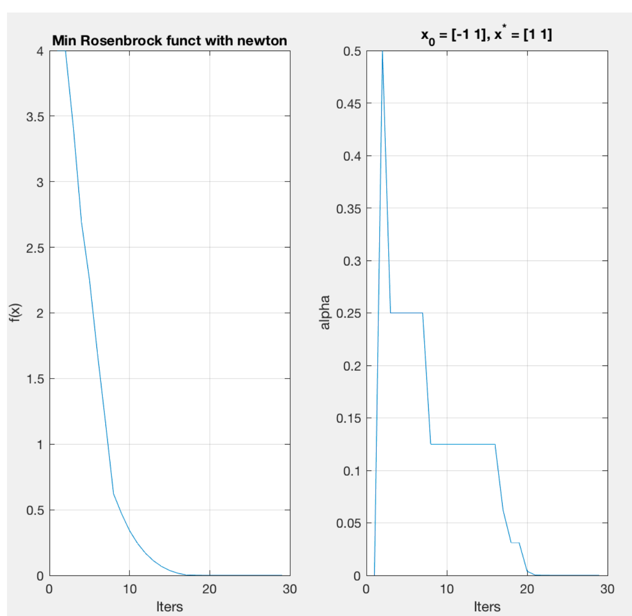
```
function resenbrock()

alpha_max = 1;
rho = 0.5;
c = 0.5;

x_0 = [-1; 1];
[x_p, f_obj, alpha, iter] =  backtracking_min(x_0, 'steepest descent', ...
                                               alpha_max, rho, c);
plot_results(x_p, f_obj, alpha, iter, 'steepest descent', x_0);
[x_p, f_obj, alpha, iter] =  backtracking_min(x_0, 'newton', alpha_max, rho, c);
plot_results(x_p, f_obj, alpha, iter, 'newton', x_0);
grid on

x_0 = [0; 1];
[x_p, f_obj, alpha, iter] =  backtracking_min(x_0, 'steepest descent', ...
                                               alpha_max, rho, c);
plot_results(x_p, f_obj, alpha, iter, 'steepest descent', x_0);
[x_p, f_obj, alpha, iter] =  backtracking_min(x_0, 'newton', alpha_max, rho, c);
plot_results(x_p, f_obj, alpha, iter, 'newton', x_0);
grid on

x_0 = [2;1];
[x_p, f_obj, alpha, iter] =  backtracking_min(x_0, 'steepest descent', ...
                                               alpha_max, rho, c);
plot_results(x_p, f_obj, alpha, iter, 'steepest descent', x_0);
[x_p, f_obj, alpha, iter] =  backtracking_min(x_0, 'newton', alpha_max, rho, c);
plot_results(x_p, f_obj, alpha, iter, 'newton', x_0);
grid on


end

function [x, f_obj, alpha, iter] = backtracking_min(x_0, method, ...
                                                       alpha_max, rho, c)
iter = 1;
tol = 1e-6;
max_iter = 1e6;
f_obj = zeros(max_iter, 1);
alpha = alpha_max * ones(max_iter, 1);
alpha(1) = 0;

x = x_0;
f_obj(iter) = rb_function(x);
while norm(rb_gradient(x)) > tol && iter < max_iter
  iter = iter + 1;
  p = step_dir(x, method);
  [alpha(iter), f_obj(iter)] = step_length(x, p, alpha(iter), rho, c);
  x = x + alpha(iter) * p;
end
f_obj = f_obj(1:iter);
alpha = alpha(1:iter);
end

function p = step_dir(x_k, method)
% Return a unit direction of search
if strcmp(method, 'newton')
  p = - rb_hessian(x_k)^-1 * rb_gradient(x_k);
else
  p = - rb_gradient(x_k);
end
p = p / norm(p);
end

function [alpha, f_x_k] = step_length(x_k, p_k, alpha_max, rho, c)
% Return the step length based on first Wolfe condition
alpha = alpha_max;
f_x_k = rb_function(x_k);
while (rb_function(x_k + alpha * p_k) > ...
      f_x_k + c * alpha * rb_gradient(x_k)' * p_k)
  alpha = rho * alpha;
end
end
```

```
% Visualize iteration
function plot_results(x_p, f_obj, alpha, iter, method, x_0)
figure;
subplot(1, 2, 1);
plot(1:iter, f_obj);
title(['Min Rosenbrock funct with ', method])
ylabel('f(x)');
xlabel('Iters');
grid on
subplot(1, 2, 2);
plot(1:iter, alpha)
title(['x_0 = ', mat2str(x_0', 3), ', x^* = ', mat2str(x_p', 3)]);
ylabel('alpha');
xlabel('Iters');
grid on
end

% Function to minimize, its gradient and hessian

function f = rb_function(x)
f = 100*(x(2) - x(1)^2)^2 + (1 - x(1))^2;
end

function gf = rb_gradient(x)
gf = [2 * x(1) - 400 * x(1) * (- x(1)^2 + x(2)) - 2;
      200 * (x(2) - x(1)^2)];
end

function hf = rb_hessian(x)
hf = [2 + 1200 * x(1)^2 - 400 * x(2), -400*x(1);
      -400 * x(1), 200];
end
```

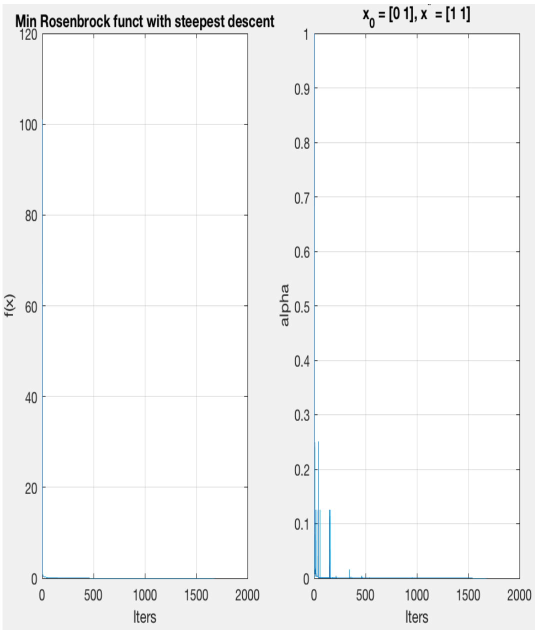
Graphs:

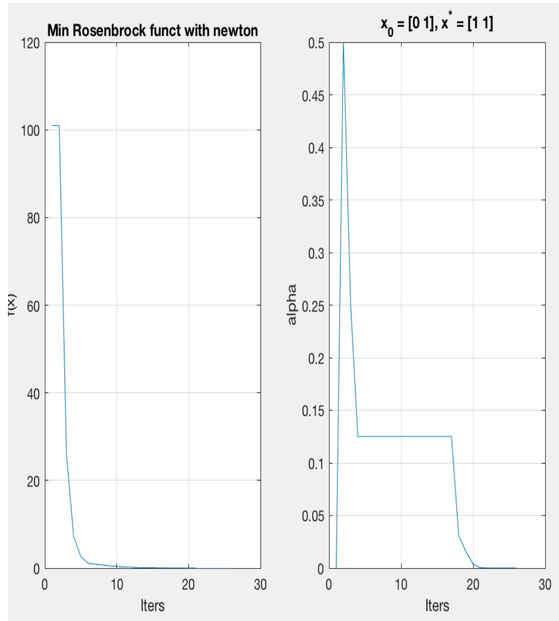
- Starting point [-1,1]



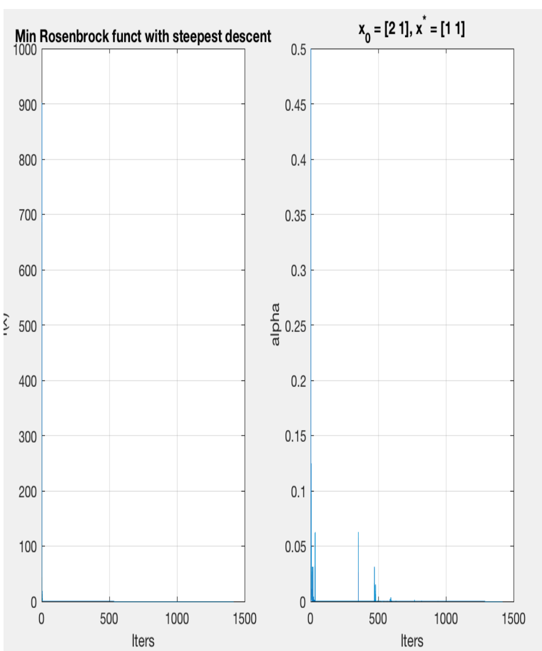


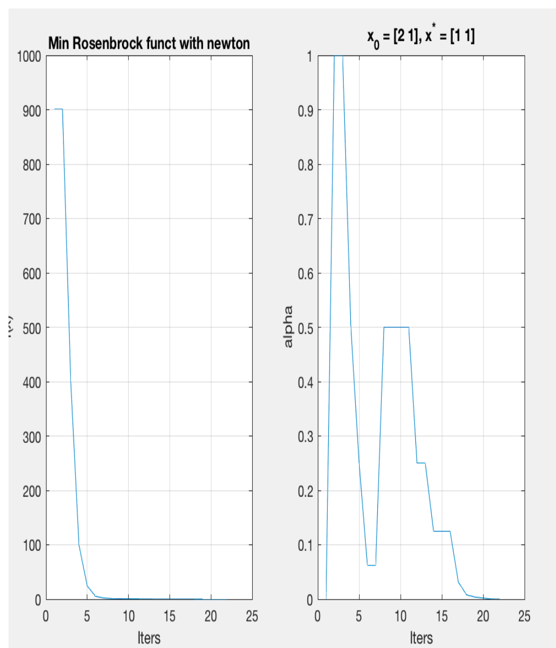
- Starting point- [0,1]





- Starting point= [2,1]





References

- https://www.wyzant.com/resources/answers/58435/find_and_classify_the_critical_points_of_f_x_x_3_1_x_4_as_local_maxima_and_minima
- http://www.wolframalpha.com/calculators/equation-solver/
- https://www.coursehero.com/file/11051399/sample-test1-solution
- http://www.math.pitt.edu/~trenchea/math1070/MATH1070_5_Rootfinding.pdf
- http://web.stanford.edu/class/cs205/homework/hw5_solutions.pdf
- http://web.mae.ufl.edu/nkim/eas6939/UnconstrainedOpt.pdf
- https://see.stanford.edu/materials/lsocoee364a/03convexfunctions.pdf
- https://math.stackexchange.com/questions/483339/proof-of-convexity-of-linear-least-squares
- https://math.stackexchange.com/questions/526657/show-convexity-of-the-quadratic-function
- http://www.stat.cmu.edu/~ryantibs/convexopt-S15/lectures/14-newton.pdf
- https://www.cs.ccu.edu.tw/~wtchu/courses/2014s_OPT/Lectures/Chapter%209%20Newton%27s%20Method.pdf
- http://www.stat.cmu.edu/~ryantibs/convexopt-F15/lectures/17-prox-newton.pdf
- https://github.com/escorciav/amcs211/blob/master/hw3/hw3_3.m