Computational Methods of Optimization
Assignment 02

Submitted by: Shatakshi Gupta

MTech. Artificial Intelligence
SR Number: 2288

24th October, 2023

Question 1:

Trisection Method

1. Compute r_1 and s_1 in terms of x_1 and y_1.
2. In how many iterations does the process end? Prove it and report your answer in terms of x_1, y_1, and ϵ.
3. You are given oracle access to a function $g(\alpha)$. (See below.) Minimize it using the trisection method. Choose the initial interval as $[x_1, y_1] = [-1, 1]$ and the tolerance $\varepsilon = 10^{-4}$. Report the number of iterations, the estimate of the minimizer α^*, and the estimate of the minimum value $g(\alpha^*)$ (both up to 9 decimal places).

Answer 2:

Compute r_1 and s_1 in terms of x_1 and y_1:

The interval $[x_1, y_1]$ is divided into three equal intervals, so the points r_1 and s_1 are given by:

$$r_1 = x_1 + \frac{y_1 - x_1}{3}$$

$$s_1 = x_1 + \frac{2(y_1 - x_1)}{3}$$

How many iterations does the process end?

At each step, we reduce the interval size by a factor of $\frac{2}{3}$. After k iterations, the interval size will be:

$$\frac{1}{3^k}(y_1 - x_1)$$

We want this interval size to be at most ε, so we set up the inequality:

$$\frac{1}{3^k}(y_1 - x_1) \leq \varepsilon$$

Solving for k, we get:

$$k \geq \log_3\left(\frac{y_1 - x_1}{\varepsilon}\right)$$

Since k must be an integer, we take the ceiling function:

$$k = \lceil \log_3 \left(\frac{y_1 - x_1}{\varepsilon} \right) \rceil$$

So, the process ends in k iterations, where $k = \lceil \log_3 \left(\frac{y_1 - x_1}{\varepsilon} \right) \rceil$.

Results from the code are:-
Number of iterations: 10
Estimate of the minimizer α^*: 0.045199749
Estimate of the minimum value $g(\alpha^*)$: 21.712852962

Question 2:

Consider minimizing the function $f(x) = \sqrt{1 + 2x^2}$ with the starting point denoted by x_0.

1. Show that $f(x)$ is convex.
2. Say, we choose $x_0 = 1$. Show that Newton's method diverges.
3. For what values of x_0 does Newton's method converge?
4. Now, apply Damped Newton's method from $x_0 = 1$ where, in each iteration, the step size is found using backtracking with parameters $\beta = 0.9$ and $c = 0.1$. Take the stopping condition as $|f'(x)| \leq 10^{-4}$. Report the final point (up to 9 decimal places) and the number of iterations, i.e., the final value of k.

Answer 2:

1. Show that $f(x) = \sqrt{1 + 2x^2}$ is convex.

A function is convex if its second derivative is non-negative. Let's compute the second derivative:

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

$$f''(x) = \frac{d^2}{dx^2} \sqrt{1 + 2x^2}$$

Using the chain rule and product rule, we can find the second derivative:

$$f''(x) = \left(4x^2 + 1\right)^{-\frac{3}{2}} \cdot 8x^2$$

Since $\left(4x^2 + 1\right)^{-\frac{3}{2}}$ and x^2 are both non-negative, $f''(x)$ is non-negative for all x. Therefore, $f(x)$ is convex.

A function is convex if its second derivative (Hessian matrix) is non-negative for all values of x. To demonstrate the convexity of $f(x)$, we'll compute its second derivative and verify that it's non-negative.

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

First, find the first and second derivatives of $f(x)$:

$$f'(x) = \frac{d}{dx} \left(\sqrt{1 + 2x^2} \right) = \frac{1}{2} \left(1 + 2x^2\right)^{-\frac{1}{2}} \cdot 4x = \frac{2x}{\sqrt{1 + 2x^2}}$$

Now, find the second derivative:

$$f''(x) = \frac{d}{dx}\left(\frac{2x}{\sqrt{1+2x^2}}\right) = 2\cdot\frac{d}{dx}\left(\frac{x}{\sqrt{1+2x^2}}\right) = 2\cdot\frac{\sqrt{1+2x^2} - x \cdot \frac{d}{dx}\sqrt{1+2x^2})}{1+2x^2} = 2\cdot\frac{1}{(1+2x^2)\sqrt{1+2x^2}}$$

$f(x)$ is convex as for all x :

$$f''(x) = 2 \cdot \frac{1}{(1+2x^2)\sqrt{1+2x^2}} \geq 0$$

2. Show that Newton's method diverges when $x_0 = 1$.

Starting with $x_0 = 1$, the Newton's update for $f(x) = 1 + 2x^2$ is:

$$x_{k+1} = x_k - \frac{f''(x_k)}{f'(x_k)} = x_k - x_k(1 + 2x_k^2) = -2x_k^3$$

So, with $x_0 = 1$:

$$x_1 = -2, \quad x_2 = -16, \quad x_3 = -4096, \ldots$$

As we can see, the sequence diverges rapidly, confirming that Newton's method diverges when starting from $x_0 = 1$.

3. For what values of x_0 does Newton's method converge?

For $|x_0| < 1$, the method converges very rapidly to the solution $x^* = 0$.

4. Apply Damped Newton's method with backtracking.

Now, let's apply Damped Newton's method using backtracking with the given parameters $\beta = 0.9$ and $c = 0.1$. We'll also set a stopping condition of $|\nabla f(x)| \leq 10^{-4}$. Reporting the the final point (up to 9 decimal places) and the number of iterations, i.e., the final value of k :-

Final point: 1.3917×10^{-5}
Number of iterations: 17

Question 3:

Nesterov's Accelerated Gradient Descent

In this question, we will compare gradient descent and accelerated gradient descent. You are given access to a oracle that, on input x, returns a tuple of the form $(\text{float}, \text{numpy.array})$. The first entry of this tuple is the value $f(x)$ and the second entry is the gradient $\nabla f(x)$ represented as a numpy array. (See below.)

1. Perform gradient descent from the starting point $x_0 = 0$ with a constant step size of $\alpha = 0.00004$. Use the stopping condition as $\|\nabla f(x)\| \leq 10^{-4}$. Report the number of iterations, the final point, and the function value at the final point (both up to 9 decimal places).
2. Perform accelerated gradient descent from the starting point $x_0 = 0$ with parameters $\alpha = 0.00004$ and $\theta = 0.142$. Use the stopping condition as $\|\nabla f(x)\| \leq 10^{-4}$. Report the number of iterations, the final point, and the function value at the final point (both up to 9 decimal places).

Answer 3:

Gradient Descent:

- Iterations: 32
- Final Point: [0.00721296 0.00721296 0.00721296 0.00721296 0.00721296]
- Final Function Value: 93.137342860

Accelerated Gradient Descent:

- Iterations: 25
- Final Point: [0.00721296 0.00721296 0.00721296 0.00721296 0.00721296]
- Final Function Value: 93.137342860

Question 4:

Conjugate Directions Method

For this question, you are required to query the given oracle to get a positive definite matrix $Q_{5\times5}$. (See the instructions below.) Let $e_0, e_1, \ldots, e_4$ be the standard basis vectors where e_i has its $(i+1)$th entry as one and the rest of the entries as zeros. The following iterative method can be used to find $u_0, u_1, \ldots, u_4$ that are Q-conjugate.

$$u_0 = e_0, \quad (4.1)$$

For i from 1 to 4:

$$u_i = e_i - \sum_{j=0}^{i-1} \frac{e_i^T Q u_j}{u_j^T Q u_j} u_j. \quad (4.2)$$

Answer the following questions:

1. Use the procedure in eq. (4.1) and eq. (4.2) to find the Q-conjugate vectors $u_0, u_1, \ldots, u_4$. Report the vectors. Keep your precision of every coordinate up to 9 decimal places.
2. Let $b = [s_1, s_2, s_3, s_4, s_5]^T$ where $s_1 s_2 s_3 s_4 s_5$ is the suffix of length 5 of your SR Number. Use the conjugate directions method to minimize $\frac{1}{2}x^T Q x - b^T x$ using the u_i's as the conjugate directions. Use $x_0 = 0$ as your starting point.
   (a) Report the final minimum x^* and $f(x^*)$ (both up to 9 decimal places).
   (b) How many iterations does your method take?

Answer 4:

Conjugate Directions Method results:

(a) Minimum $x^* = [-1.41319228, -0.31404273, 0.52893591, 0.79340386, 0.63472309]$, $f(x^*) = -6.101016387$

(b) Number of iterations: 5