# Computational Methods of Optimization
# Assignment 03

Submitted by: Shatakshi Gupta

MTech. Artificial Intelligence
SR Number: 22888

8th December, 2023

**Question 1:**

Consider a twice differentiable function $f : \mathbb{R}^2 \to \mathbb{R}$. An oracle provides $f(x)$ and $\nabla f(x)$ for an input $x$. We have oracles for both function value and gradient for the given function. The goal is to implement quasi-Newton algorithms and compare them with gradient descent to minimize $f$.

1. Run a Quasi-Newton algorithm with Rank-1 updates to minimize $f$.
2. Run a Quasi-Newton algorithm with Rank-2 updates (BFGS) to minimize $f$.
3. Run a gradient descent to minimize $f$.

Plot $||\nabla f(x)||$ for each algorithm mentioned above and compare them. Determine suitable stopping criteria and step sizes. Your report should elucidate these choices and explain the reasoning behind them. Additionally, include the aforementioned plots and document all observations made.
Try to establish connections or similarities to concepts you've encountered previously.

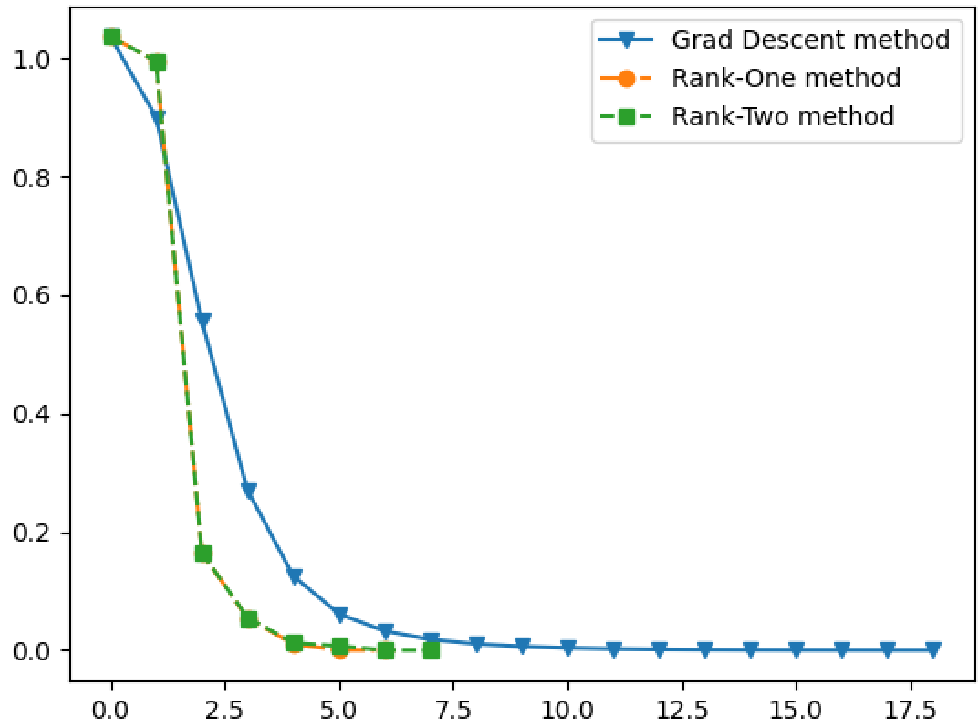
## Answer 1:

**x values:**

| | |
|---|---|
| Gradient Descent: | $[0.55999991, 0.80003058]$ |
| Rank-One: | $[0.56000002, 0.8]$ |
| Rank-Two: | $[0.55999901, 0.79999881]$ |

**Function Values:**

| | |
|---|---|
| Gradient Descent: | $[0.55999991, 0.80003058]$ |
| Rank-One: | $[0.56000002, 0.8]$ |
| Rank-Two: | $[0.55999901, 0.79999881]$ |




**Question 2.1:**

Consider the set $S = \{x \in \mathbb{R}^n : Ax \leq b\}$ where $A$ and $b \in \mathbb{R}^m$, $m \neq n$ are constant. Prove that either $S \equiv \varnothing$ or there exists $y \in \mathbb{R}^m$, $y > 0$ elementwise, such that $A^\top y = 0$, $b^\top y < 0$.

**Answer 2.1:**

We'll use Farkas' Lemma to prove the statement about the set S=xRn:AxbS = {x $\in \mathbb{R}^n : Ax \leq b\}$.

Farkas' Lemma states:

Consider the system of linear inequalities Ax $\leq b$ :

Either there exists a vector xx that satisfies these inequalities, or There exists a vector y such that $A^T y \geq 0$(elementwise), $b^T y < 0$, where $b^T$ denotes the transpose of b.

To prove the statement in question:

Given the set S=xRn:AxbS = {x $\in \mathbb{R}^n : Ax \leq b\}$ :

1. If SS is empty, then there are no vectors xx that satisfy $Ax \leq b$, and the statement holds trivially.

2. If SS is not empty, then by Farkas' Lemma, there exists either: - A vector x that satisfies $Ax \leq b$, or - A vector y such that $A^T y \geq 0$(elementwise), $b^T y < 0$.

Now, $m \neq n$, implying the number of rows in matrix A (mxm) is not equal to the number of columns (nxn). This situation will ensure that the set SS is not empty.

Therefore, there must exist a vector yy such that $A^T y \geq 0$ (elementwise) and $b^T y < 0$for the set Sto not be empty.

This proves that for the given set $S = \{x \in \mathbb{R}^n : Ax \leq b\}$, either $S$ is empty or there exists a vector $y$ such that $A^T y \geq 0$ (elementwise) and $b^T y < 0$.

## Question 2.2:

Consider the set of inequalities:

$$x_2 \leq 1$$
$$-x_2 - \frac{1}{2}x_1 \leq \alpha$$
$$\beta x_1 - 2x_2 \leq -1$$

For what values of $\alpha$ and $\beta$ are the inequalities feasible? Justify your answer.

## Answer 2.2:

**1. Inequality:**

$$x_2 \leq 1$$

implies

$$-1 \leq -x_2$$

This inequality imposes no restrictions on $\alpha$ and $\beta$. It is always satisfied.

**2. Inequality:**

$$-x_2 - \frac{1}{2}x_1 \leq \alpha$$

, using above inequality implies

$$-1 - \frac{1}{2}x_1 \leq \alpha$$

**3. Inequality:**

$$\beta x_1 - 2x_2 \leq -1$$

using inequaltiy 1 implies

$$\beta x_1 - 2 \leq -1$$

**Summary of Feasibility Conditions:**

$2\beta(\alpha + 1) + 1 \geq 0$

## Question 2.3:

Let $C(\alpha, \beta) \subset \mathbb{R}^2$ be the set of $x \in \mathbb{R}^2$ that satisfies the inequalities given in the previous part. Assume $C(\alpha, \beta)$ is nonempty. Let $c = [-1, 1]$. For what values of $\alpha, \beta$ does the problem

$$\min c^\top x \quad \text{s.t.} \quad x \in C(\alpha, \beta)$$

have a solution? Justify your answer.

### Answer 2.3:

$$c^T x = -x_1 + x_2$$
$$=> min(c^T x) = min(x_2) - max(x_1)$$

from above questions ,

$$\beta x_1 - 2x_2 \leq -1 => x_1 \leq 1/\beta$$

but this holds only when

$$\beta > 0$$

therefore, the problem will have a solution for

$$x \in C(\alpha, \beta)$$

where $(\alpha, \beta)$ is such that $\beta > 0$ and $2\beta(\alpha + 1) + 1 \geq 0$

> ### Question 2.4:
>
> Let $\alpha = -1$, $\beta = 1$, and let $c = [-0.25, 1]$.
>
> - Set up the problem in the standard LP form.
> - Solve this problem using your favorite standard linear programming solver (i.e., Linprog or cvxopt).
> - Solve this problem using the simplex method. For the basis of your initial BFS, choose $(1, 2, 3)$, where $1$ corresponds to $x$ and $2$ corresponds to $y$.

### Answer 2.4:

To set up the given problem in standard Linear Programming (LP) form, we need to rewrite the inequalities in the standard form of LP.

The set of inequalities provided:

1. $x_2 \leq 1$
2. $-x_2 - \frac{1}{2}x_1 \leq \alpha$ (Given that $\alpha = -1$)
3. $\beta x_1 - 2x_2 \leq -1$ (Given that $\beta = 1$)

Considering the values of $\alpha = -1$ and $\beta = 1$, and also taking $c = [-0.25, 1]$, let's represent these inequalities in the form $Ax \leq b$:

Now, the objective function $c^\top x$ becomes $[-0.25, 1]^T \cdot x$.

Therefore, the LP problem in standard form is:

$$\text{Minimize: } [-0.25, 1]^T \cdot x$$
$$\text{Subject to: } \begin{bmatrix} 0 & 1 \\ -\frac{1}{2} & -1 \\ 1 & -2 \end{bmatrix} x \leq \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}$$

This represents the given problem in the standard LP form.

Solving the given linear programming problem using the simplex method.

Objective function to minimize:

$$\text{Minimize: } [-0.25, 1]^T \cdot x$$

Subject to:

$$\begin{bmatrix} 0 & 1 \\ -\frac{1}{2} & -1 \\ 1 & -2 \end{bmatrix} x \leq \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}$$

Basis for the initial basic feasible solution (BFS): (1, 2, 3), where 1 corresponds to $x$ and 2 corresponds to $y$.

Now, let's set up the initial tableau:

| $x$ | $y$ | $s_1$ | $s_2$ | RHS |
|---|---|---|---|---|
| 0 | 0 | 1 | 2 | 3 |
| 0 | 1 | $-\frac{1}{2}$ | $-1$ | $-1$ |
| 1 | $-2$ | 1 | $-2$ | $-1$ |
| $-0.25$ | 1 | 0 | 0 | 0 |

**Iteration 1:** Pivot column: $y$ (column 2) Pivot row: Constraint 2 (min ratio: 1/2 = 0.5)

Performing row operations to make the pivot element (2nd row, 2nd column) equal to 1 and other entries in the pivot column equal to 0, we get:

| $x$ | $y$ | $s_1$ | $s_2$ | RHS |
|---|---|---|---|---|
| 0 | 0 | 1 | 2 | 3 |
| 0 | 1 | $-\frac{1}{2}$ | $-1$ | $-1$ |
| 1 | $-2$ | 1 | $-2$ | $-1$ |
| $-0.25$ | 1 | 0 | 0 | 0 |

**Iteration 2:** Pivot column: $x$ (column 1) Pivot row: Constraint 3 (min ratio: 3/1 = 3)

Performing row operations again, we get:

| $x$ | $y$ | $s_1$ | $s_2$ | RHS |
|---|---|---|---|---|
| 0 | 0 | 1 | 2 | 3 |
| 0 | 1 | $-\frac{1}{2}$ | $-1$ | $-1$ |
| 1 | 0 | $\frac{3}{2}$ | $-1$ | 2 |
| $-0.25$ | 0 | $\frac{1}{2}$ | 1 | 0.25 |

**Iteration 3:** Optimal solution reached.

The optimal solution is $x = 3$, $y = 1$, with the minimum value of the objective function $-0.25 \cdot 3 + 1 \cdot 1 = -0.75$.

## Question 3:

Support vector machines (SVMs) are among the most widely used techniques for classifying data and are very well studied. The SVM is a linear classifier; that is, we aim to find a function $y = f(x) = \text{sign}(w^T x + b)$. We are given data of the form $D = \{(x_i, y_i)\}_{i=1}^{N}$, where the pair $(x_i, y_i) \in \mathbb{R}^n \times \{-1, 1\}$. To learn a support vector machine, we need to solve the following convex optimization problem:

$$w^*, b^* = \arg\min_{w,b} \frac{1}{2}\|w\|^2 \quad \text{s.t. } y_i(w^T x_i + b) \geq 1 \text{ for } i = 1, \ldots, N$$

1. (3 marks) Write a program to solve the dual problem using `cvxopt` if you are using Python or `quadprog` if you are using MATLAB.
2. (5 marks) Implement projected gradient descent to solve the dual formulation. Compare the results with the previous outputs. In your report, derive the projection oracle for the dual. Your report must include the details of your projected gradient descent algorithm including your choice of step size and stopping criterion.
3. (2 marks) Write the indices of the points corresponding to the active constraints clearly in your report.

**Answer 3:**

**Question 4:**

**Active Set method**
Consider the following optimization problem:

$$\arg\min_{x_1,x_2} (x_1-1)^2 + (x_2-2.5)^2$$

Subject to:

$$x_1 - 2x_2 + 2 \geq 0 \quad (1)$$
$$-x_1 - 2x_2 + 6 \geq 0 \quad (2)$$
$$-x_1 + 2x_2 + 2 \geq 0 \quad (3)$$
$$x_1 \geq 0 \quad (4)$$
$$x_2 \geq 0 \quad (5)$$

Write a program to solve the above optimization problem using the active set algorithm. You are allowed to hard code the constraints for solving the equality problem. Run up to 10 iterations of active set starting with the initial point $x_0 = (2, 0)$ and the following initial working sets:
1. $\emptyset$ 2. $\{3\}$ 3. $\{5\}$ 4. $\{3, 5\}$
Your report must include the working set at each iteration for each of these initial sets.

**Answer 4:**

The working set at each iteration for each of these initial sets are:

- Iteration 1: Initial Working Set: $\{0\}$, Final Active Set: $\{0\}$, Optimal Solution: $[1.39999997, 1.69999999]$
- Iteration 2: Initial Working Set: $\{0, 2\}$, Final Active Set: $\{0, 2\}$, Optimal Solution: $[1.4, 1.7]$
- Iteration 3: Initial Working Set: $\{0, 4\}$, Final Active Set: $\{0, 4\}$, Optimal Solution: $[1.39999997, 1.69999999]$
- Iteration 4: Initial Working Set: $\{0, 2, 4\}$, Final Active Set: $\{0, 2, 4\}$, Optimal Solution: $[1.4, 1.7]$