# MTH603/AM8211
# Fall 2023

## Assignment Cover Page

Assignment Number: 1

Student Last Name: Walaa

Student First Name: Mariam

Student ID: 500-744-640

*If you used any resources other than what is listed on the course outline, or if you collaborated with any of your colleagues, list below these resources or the names of your colleagues along with their id's:*

You are expected to know what constitutes academic integrity, to avoid committing academic offences, and to take responsibility for your actions. Provide your signature below to indicate that you are aware of all policies of the course (described in the course outline), including copyright.

**Signature:** Mariam Walaa

# Assignment 1

Mariam Walaa

September 20, 2023

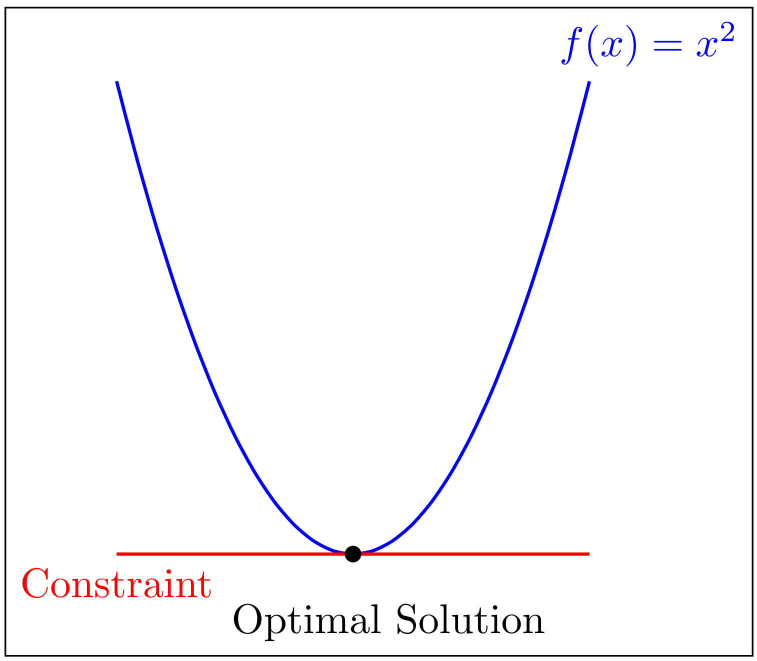


Figure 1: Non-linear Optimization Plot

## Question 1(A)

A general nonlinear program can be formulated as either a maximization or minimization problem, where each constraint can take one of the following forms:

- $g(x) \leq 0$
- $g(x) \geq 0$
- $g(x) = 0$

Here, $g : \mathbb{R}^n \to \mathbb{R}$.
Prove that any nonlinear program can be equivalently expressed in the form:

$$\text{minimize } f(x) \quad \text{subject to } g_i(x) \leq 0 \quad \text{for } i = 1, 2, \ldots, m$$

**Solution:**

We want to show that any nonlinear program can be written in the form

$$\text{minimize } f(x) \text{ subject to } g_i(x) \leq 0 \text{ for } i = 1, 2, \ldots, m$$

Consider an arbitrary NLP with the following form:

$$\text{minimize } f(x)$$

subject to:

$$g_i(x) \leq 0 \text{ for } i = 1, \ldots, q \text{ (inequality constraints)}$$
$$h_j(x) = 0 \text{ for } j = 1, \ldots, p \text{ (equality constraints)}$$
$$r_k(x) \geq 0 \text{ for } k = 1, \ldots, s \text{ (inequality constraints)}$$

where $q + p + s = m$.
The constraints $g_i, h_j, r_k$ are all possible constraints an NLP can consist of.
We can write each inequality constraint $r_k(x) \geq 0$ as $r'_k(x) := -r_k(x) \leq 0$.
Similarly, we can write inequality constraints $h_j(x) = 0$ as $h'_j(x) := -r_k(x) \leq 0$ and $h''_j(x) := r_k(x) \leq 0$.
Now, we have transformed the original NLP into the following form:

$$\text{minimize } f(x)$$

subject to:

$$g_i(x) \leq 0 \text{ for } k = 1, \ldots, q$$
$$h'_j(x) \leq 0 \text{ for } j = 1, \ldots, p$$
$$h''_j(x) \leq 0 \text{ for } j = 1, \ldots, p$$
$$r'_k(x) \leq 0 \text{ for } k = 1, \ldots, s$$

We have now shown that any NLP can be written in the specified form, where the objective function is to minimize $f(x)$ subject to a set of inequality constraints $g_i(x) \leq 0$.

## Question 1(B)

Suppose we have a nonlinear program with an objective function $f(x)$. What happens to the optimizers and the optimal solution if we change the objective function to $603 \cdot f(x) - 2018$?

**Solution:**

We want to analyze what happens to the optimizers and optimal solution if we change the objective function to $603 \cdot f(x) - 2018$.

Original nonlinear program:

$$\text{minimize } f(x) \quad \text{subject to constraints}$$

Modified nonlinear program:

$$\text{minimize } 603 \cdot f(x) - 2018 \quad \text{subject to the same constraints}$$

When we change the objective function of a nonlinear program from $f(x)$ to $603 \cdot f(x) - 2018$, we are scaling and shifting the original objective function. This transformation does not change the optimizers or the optimal solution in terms of the original function's critical points and their corresponding values.

- Scaling the objective function by a constant factor (603 in this case) does not affect the location of the optimal solution or the values of the optimizers. It simply scales the optimal value $f(x*)$ for some minimizer $x*$. If the optimal value is $f(x*) = 10$, then the new value will be $603 * 10$ after scaling.
- Shifting the objective function by a constant value (2018 in this case) decreases the value of the objective function at its optimal value by 2018, after scaling it by 603. For example, if $x*$ is the minimizer and $f(x*) = 10$ is the optimal value, then the new optimal value will be $603 * 10 - 2018$.

In summary, changing the objective function to $603 \cdot f(x) - 2018$ does not alter the critical points, optimizers, or the location of the optimal solution in the decision variable space. It only modifies the scale and value of the objective function at its optimum.

## Question 2(A)

Prove that the intersection of two convex sets $A$ and $B$ in $\mathbb{R}^n$, denoted as $A \cap B$, is itself convex.

**Proof:** To prove that the intersection of two convex sets $A$ and $B$ in $\mathbb{R}^n$, denoted as $A \cap B$, is convex, we'll use the definition of convex sets.

**Definition:** A set $C$ is convex if, for any two points $x$ and $y$ in $C$, the line segment connecting $x$ and $y$ is entirely contained within $C$. Mathematically, this can be expressed as:

$$\text{For all } x, y \in C \text{ and for all } t \text{ in the interval } [0,1], \quad tx + (1-t)y \in C.$$

Now, let's prove that $C := A \cap B$ is convex:

1. Take any two points, $x$ and $y$, in $A \cap B$.
   Since $x$ and $y$ are in both $A$ and $B$: $x, y \in A$ and $x, y \in B$.
2. Now, let's consider the line segment connecting $x$ and $y$, denoted as $L$:
   $$L = \{tx + (1-t)y \mid t \in [0,1]\}.$$
3. We want to show that every point in $L$ is also in $A \cap B$.
   - Consider any $t$ in the interval $[0,1]$.
   - By the convexity of set $A$, we know that
     $$tx + (1-t)y \in A.$$
   - Similarly, by the convexity of set $B$, we know that
     $$tx + (1-t)y \in B.$$
4. Since $tx + (1-t)y$ is both in $A$ and in $B$, it is in their intersection:
   $$tx + (1-t)y \in A \cap B.$$

As wanted, we have shown that for any $t$ in $[0,1]$, $tx + (1-t)y$ is in $A \cap B$.
Therefore, $A \cap B$ is convex, as it satisfies the definition of convexity.

## Question 2(B)

Suppose we have a function $f(x) = c \cdot x$, where $c$ and $x$ are both vectors. Prove that $f$ is a convex function.

**Solution:**

To prove that the function $f(x) = c \cdot x$, where $c$ and $x$ are both vectors, is convex, we can use the definition of convexity.

**Definition:** A function is convex if, for all $x_1$ and $x_2$ in its domain and for all $t$ in the interval $[0, 1]$, the following inequality holds:

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2)$$

In this case, we have:

$$f(x) = c \cdot x$$

Now, let's prove that $f(x)$ is convex:

(a) Take any two vectors $x_1$ and $x_2$ in the domain of $f(x)$.

(b) Consider any $t$ in the interval $[0, 1]$.

(c) We want to prove that:

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2)$$

Substitute the expression for $f(x)$ into the inequality:

$$c \cdot (tx_1 + (1-t)x_2) \leq tc \cdot x_1 + (1-t)c \cdot x_2$$

(d) Now, we can simplify both sides of the inequality:

$$tc \cdot x_1 + (1-t)c \cdot x_2 \leq tc \cdot x_1 + (1-t)c \cdot x_2$$

The right-hand side of the inequality is equal to itself.

Since the inequality holds, we have shown that the function $f(x) = c \cdot x$ is convex, as it satisfies the definition of convexity.

## Question 3(A)

Suppose we have an unconstrained nonlinear program:

$$\text{minimize } \|Ax - b\|$$

Show that it is equivalent to the following linear program:

$$\text{minimize } \sum_{i=1}^{n} |t_i|$$

subject to

$$-t_i \leq t_i \leq |t_i| \text{ for } i = 1, 2, \ldots, n$$

### Solution

Let $t_i = |(Ax - b)_i|$ for $i = 1, ..., m$. Consider the unconstrained nonlinear program:

$$\text{minimize } \|Ax - b\|$$

Suppose $\bar{x}$ is the solution that minimizes this objective. Then, by definition, it must satisfy:

$$\|A\bar{x} - b\| \leq \|Ax - b\| \text{ for all } x$$

Now, let's look at the linear program:

$$\text{minimize } \sum_{i=1}^{m} t_i$$

subject to

$$-t_i \leq (Ax - b)_i \leq t_i \text{ for } i = 1, 2, \ldots, m$$

The term $-t_i \leq (Ax - b)_i \leq t_i$ enforces that $t_i$ should be non-negative and that it has an absolute value of $t_i$. Therefore, for each component $i$, the objective $\sum_{i=1}^{m} t_i$ represents the absolute value of the difference between $Ax$ and $b$ for that component.

First, we want to show that if $\bar{x}$ minimizes the LP, it also minimizes the NLP.
If $\bar{x}$ minimizes $\|Ax - b\|$, it means that for each component $i$, $t_i$ is minimized in magnitude, and therefore, $\sum_{i=1}^{n} |t_i|$ is also minimized because it represents the sum of these magnitudes.

Conversely, we want to show that if $\bar{x}$ minimizes the NLP, it also minimizes the LP.
If $\bar{x}$ minimizes $\sum_{i=1}^{m} t_i$, it implies that for each component $i$, $t_i$ is minimized in magnitude, which is equivalent to minimizing $\|Ax - b\|$.

Therefore, the two problems are indeed equivalent, and any solution that minimizes one objective function also minimizes the other.

## Question 3(B)

Suppose we have an unconstrained linear program:

$$\text{minimize } \sum_{i=1}^{m} t_i$$

subject to

$$-t \leq Ax - b \leq t \text{ for } i = 1, 2, \ldots, m$$

Convert this linear program to the form:

$$\min \; c \cdot y$$

subject to

$$Dy \leq e$$

Define the vectors $c$, $y$, $D$, and $e$ accordingly.

### Solution

To convert the given linear program into the form min $c \cdot y$ subject to $Dy \leq e$, we redefine the variables and matrices as follows:

- $c$ is an $m$-dimensional vector with all elements equal to 1.
- $y$ is an $m$-dimensional vector representing the variables $t_i$.
- $D$ is a $2m \times n$ matrix where the first $m$ rows correspond to $Ax - b$ and the last $m$ rows correspond to $b - Ax$.
- $e$ is an $2m$-dimensional vector representing the variables $t_i$.

With these definitions, we have the objective function $c \cdot y = \sum_{i}^{m} t_i$, as wanted.
Now, the original linear program can be expressed as:

$$\text{minimize } c \cdot y$$

subject to

$$Dy \leq e$$

This new linear program is in the desired form min $c \cdot y$ subject to $Dy \leq e$ and is equivalent to the original problem.