SSY281
Assignment 4
Optimization basics and QP problem

Manish Suvarna

February 19, 2020

Optimization Basics

1 Question 1a

Since the interest of this question is at satisfying a particular local minima condition in the neighbourhood of the feasible points x_1 and x_2, the given problem can be treated as a convex optimization problem, accordingly the point can now be framed as shown in Equation 1

$$z = \theta x_1 + (1 - \theta)x_2 \quad \theta \in \mathbb{R} \tag{1}$$

Hence the condition at which $z = \frac{x_1+x_2}{2}$ is achieved from Equation 1 when $\theta = 0.5$

2 Question 1b

On referring to the definition of convex function, one can treat z as a convex function of convex combination of x_1 and x_2 which is less than or equal to the convex combination of the function at x_1 and x_2 as shown in Equation 2. On observing this equation, $f(z) \leq f(x_1)$ if $f(x_1) \geq f(x_2)$ and vice-versa. Hence z will never be worse than X_1 and x_2. But if $f(x_1) = f(x_2)$, then $f(z)$ can be equal to or better than x_1 and x_2

$$f(x_1\lambda + (1 - \lambda)x_2) \leq f(x_1)\lambda + (1 - \lambda)f(x_2) \tag{2}$$

3 Question 1c

Yes.

4 Question 2a

Halfspaces are defined as sets in the form of $\{x|a^T x \leq b\}$ which are convex in nature. The given set of slab is defined as $\{x|\alpha \leq a^T x \leq \beta\}$. This set can be represented as the intersection of two hyperspaces by rewriting its constraint function as shown in Equation 3

$$\begin{aligned} S_1 &= \{x| - a^T x \leq -\alpha\} \\ S_2 &= \{x|a^T x \leq \beta\} \end{aligned} \tag{3}$$

Since, intersection of two convex sets is said to be convex, the given given set of slab is therefore convex in nature.

5 Question 2b

The given set M can be rewritten as M= $\{x| \ \ ||x-y|| - f(y) \leq 0 \forall \ \ y \in S\}$. Let g(x) = $||x-y|| - f(y)$, for M to be convex, for any given two points x_1 and x_2, $z = \theta x_1 + (1-\theta)x_2$ must also belong to M, (g(z) ≤ 0).

$$g(z) = g(\theta x_1 + (1-\theta)x_2) = ||\theta x_1 + (1-\theta)x_2 - y|| - f(y) \tag{4}$$

Since the points x_1 and $x_2 \in$ M, $g(x_1), g(x_2) \leq 0$,

$$\theta g(x_1) + (1-\theta)g(x_2) \leq 0 \tag{5}$$

Equation 5 can be worked upon to get the form of g(z) as follows,

$$\begin{aligned} \theta g(x_1) + (1-\theta)g(x_2) = \theta(||x_1 - y|| - f(y)) + (1-\theta)(||x_2 - y|| - f(y)) \\ = ||\theta x_1 - \theta y|| - \theta f(y) + ||x_2 - y|| - f(y) - ||\theta x_2 - \theta y|| + \theta f(y) \\ = ||\theta x_1 + (1-\theta)x_2 - y|| - f(y) \end{aligned} \tag{6}$$

On comparing Equation 4, 5 and 6, g(z)$\leq$ 0, hence M is convex.

6 Question 2c

Let g(x) = $||x - x_0||_2 - ||x - y||_2 \quad \forall \, y \in S$

$$\begin{aligned} g(x) \leq 0 \\ \implies (x - x_0)^T(x - x_0) \leq (x - y)^T(x - y) \\ \implies (y - x_0)^T x + (y - x_0)x^T \leq y^T y - x_0^T x_0 \\ \implies 2(y - x_0)^T x \leq y^T y - x_0^T x_0 \end{aligned} \tag{7}$$

The final derived expression in Equation 7 is dual to that of the form of union of two half spaces. Since union of convex sets cannot be convex always, the give set is not convex.

7 Question 3a

By definition, 1-norm of matrix is basically the sum of absolute values of the columns of the given matrix. Hence the given 1-norm objective function can be written as shown in Equation 8 which is now representing a set of linear equations, which can be recasted as a linear programming problem as shown in 9

$$min|Ax_1|_1 \implies min \sum |a_i| x \tag{8}$$

$$\begin{aligned} min \quad 1^T t \\ subject \quad to, \quad Ax - t \leq 0 \\ -Ax - t \leq 0 \\ Fx \leq g \\ x \in \mathbb{R}^n, \quad t \in \mathbb{R}^m \end{aligned} \tag{9}$$

8 Question 3b

Whereas the ∞ norm objective $||Ax||_\infty$ is the maximum values of $a_i x$ where a_i^T is rows of the matrix A. Therefore the objective function under ∞ norm now becomes as shown in 10

$$|Ax|_\infty = max(a_i x) \tag{10}$$

The equivalent linear programming problem of this objective function is as shown in 11

$$\begin{aligned} min \quad t \\ subject \quad to, \quad a_i^T x - t \leq 0 \\ -a_i^T x - t \leq 0 \\ Fx \leq g \\ x \in \mathbb{R}^n, \quad t \in \mathbb{R}^m \end{aligned} \tag{11}$$

9 Question 4

Based on the answer for the previous question, function N1.m and Ninf.m MATLAB files were filled accordingly by using the command linprog. Since the involved decision variable is meant to be a regular point now, the constrained inequalities are entered as the equality constraint in the respective functions.

10 Question 5a

The solution obtained for the given Quadratic programming using quadprog command aree as shown in Equation 12 with an optimal cost function of 4.64.

$$[x_1 \ \ x_2 \ \ u_0 \ \ u_1]^T = \begin{bmatrix} 2.5 \\ 0.625 \\ 1.5 \\ -0.625 \end{bmatrix} \tag{12}$$

11 Question 5b

The KKT conditions for the given optimization problem are as follows,

$$\begin{aligned} \nabla f(x^*) + \nabla g(x^*)\mu^* + \nabla h(x^*)\lambda^* = 0 \\ \mu^* \geq 0 \\ g(x^*) \leq 0; \quad h(x^*) = 0 \end{aligned} \tag{13}$$

where μ^* is the lagrangian multiplier of inequality constraints function $g(x^*)$ and λ^* is the lagrangian multiplier of the equality constraint function $h(x^*)$, which are as shown in the Equation 14

$$\mu^* = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 4.3125 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad \lambda^* = \begin{bmatrix} 1.8125 \\ -0.6250 \end{bmatrix} \tag{14}$$

The code in the MATLAB considers the lower bound and upper bound as a set of two inequality constraints hence the 8 elements vector μ^* where the upper for correspond to the upper bound constraints and the lower 4 elements correspond to the lower bound constraints. These obtained values when substituted in Equation 13 satisfies all the mentioned conditions of KKT. The non-zero positive value in the μ^* which is 4.3125 corresponds to the state x_1 which implies that the inequality constraint is strictly active at $x_1 \leq 2.5$

12 Question 5c

The cost function on removing the lower bound was found to be 0.2647 which is lesser than that with the lower and upper bound. This is obvious because the states are allowed to explore more in the feasible direction in order to optimize the objective function, where as no change in cost function or the optimal variables were observed on removing the upper bound. This makes sense because the very nature of the process of optimization is to pull the center of mass of the feasible space to as minimum as possible so its the lower bound that will be restricting this process more than the upper bounds.