SSY281 - Multi Predictive Control

Optimization basics and QP problems

Assignment - 4

ID-Number 43

Fikri Farhan Witjaksono fikrif@student.chalmers.se

February 20, 2020

1 PROBLEM STATEMENT (FIRST PART)

Consider the following statement. The following value have to be used in the question a,b,c of the assignment.

$$\begin{aligned} \min_x \quad & f(x) \\ \text{s.t} \quad & g(x) \leq 0 \\ & h(x) = 0 \end{aligned} \tag{1}$$

where

$$f : R^n \rightarrow \mathrm{IR}, g : \mathrm{IR^n} \rightarrow \mathrm{IR^m}$$

$$h : \mathrm{IR^n} \rightarrow \mathrm{IR^q}, \mathrm{x} \in \mathrm{IR^n}$$

2 QUESTIONS

2.1 Question a: Let x_1 and x_2 be feasible points answer the questions below

2.1.1 Find a simple conditions such that such that z is a feasible solution

Let we have a set C which is convex which happens if the line segment between any two points in C lies in C,i.e., if for any x_1,$x_2$$\in C$ and any θ with $0 \leq \theta \leq 1$, we have

$$\theta x_1 + (1 - \theta)x_2 \in C \tag{2}$$

If we assume that $\theta = 0.5$ and input it into the equation (2) above, we will have the equation transformed into the equation shown as below

$$0.5x_1 + 0.5x_2 \in C \tag{3}$$

Hence, by the convexity of the feasible set, it is proven that the condition $\theta = 0.5$ will give the solution $z = \frac{x_1+x_2}{2}$ as a feasible solution since it is a convex combination of two feasible points. This condition is called as midpoint convex.

2.1.2 Find conditions such that z is never worse than both x_1 and x_2

To do this, we need to make assumption. Suppose S is a nonempty set in $\mathrm{IR^n}$ and suppose that x_2 is a strict local minimum (e.g. *there exists no ϵ-neighbourhood $N_\epsilon(x_2)$ around x_2 such that $f(x) > f(x_2)$ for each $x \in S \bigcap N_\epsilon(x_2)$).* Then, by the fundamental property of convex optimization problem which rules that every locally optimal point is also globally optimal, we can say that x_2 is also a global minimum.

By contradiction, assume that x_2 is not the unique global optimal solution. That is, suppose that there exist an $x_2 \in S$ such that $f(x_1) = f(x_2)$. Then, defining z as we have define in this problem, we have by the convexity of f and S, that $f(z) \leq \lambda f(x_1) + (1 - \lambda)f(x_2) = f(x_2)$, and $z \in S$ for all $0 \leq \lambda \leq 1$. By taking positive value of $\lambda(\lambda \rightarrow 0^+)$, we can make $z \in N_\epsilon(x_2)$ for any $\epsilon > 0$ such that x_2 is not strict local optimal and we can get z which is never worse than x_1 and x_2 by the inequality that we have above.

However, this will contradict the strict local optimality of x_2, hence it is impossible to do, therefore x_2 is the unique global minimum and **the conditions such that z is never worse than both x_1 and x_2 cannot be found**.

2.1.3 Is it the same with $\lambda = 0.5$

If $\lambda = 0.5$, we have z as following

$$z = \frac{1}{2}x_1 + \frac{1}{2}x_2 \tag{4}$$

and by convexity of f and S, we will have the inequality above expressed as

$$f(\frac{1}{2}x_1 + \frac{1}{2}x_2) \leq \frac{1}{2}f(x_1) + \frac{1}{2}f(x_2) = f(x_2) \tag{5}$$

Since S is convex, $\frac{1}{2}x_1 + \frac{1}{2}x_2 \in S$, and the above inequality contradicts global optimality of x_2. Hence, x_2 is the unique global minimum and the notion that by choosing $\lambda = 0.5$, **we could conclude that the target of getting z which is never worse than x_1 and x_2 is impossible to do.**

2.2 Question b: Determine whether these sets are convex

2.2.1 Slab problem

Since a slab is an intersection between two halfspaces, therefore it is a convex set. It is also a polyhedra, which is the intersection of a finite number of halfspaces and hyperplanes.

A hyperplane divides R^n into two halfspaces whereas a halfspace is a set of the form as shown below

$$\{x|a^T x \leq b\}, \tag{6}$$

where $a \neq 0$ which means that the solution of one nontrivial linear inequality. All halfspaces are convex, but not affine. Hence, the solution above is proven.

2.2.2 The set M

$\|x - y\|$ is the distance between two vectors x and y. Suppose $x, y \in \mathbb{R}^2$.We can express the given set as below

$$\|x - y\| \leq f(y) \Longleftrightarrow \sqrt{< x - y, x - y >} \leq f(y) \tag{7}$$
$$\Longleftrightarrow \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \leq f(y)$$

If we decompose further, we would find that this set could not be reduced to halfspace form. However, it is an affine set since it contains the linear combination . Since an affine set is also convex, therefore it is a convex set.

2.2.3 A set of points closer to a given point than to a given set

Since a norm is always nonnegative, we have $\|x - x_0\|_2 \leq \|x - y\|_2$ if and only if $\|x - x_0\|_2^2 \leq \|x - y\|_2^2$.

$$\|x - x_0\|_2^2 \leq \|x - y\|_2{}^2 \Longleftrightarrow (x - x_0)^T(x - x_0) \leq (x - y)^T(x - y) \tag{8}$$
$$\Longleftrightarrow x^T x - 2x_0^T x + x_0^T x_0 \leq x^T x - 2y^T x + y^T y$$
$$\Longleftrightarrow 2\|y - x_0\|x \leq y^T y - x_0^T x_0$$

The equation above proved that the set is indeed a halfspace. Moreover, it is a convex set since it can be expressed as an intersection of different halfspaces.

$$\bigcap_{y \in S} \{x| \quad \|x - x_0\|_2 \leq \|x - y\|_2\} \tag{9}$$

According to definition of convex set, The intersection of different halfspaces is a convex set.

2.3 Question c: Show that the 1-norm and ∞-norm objective could be expressed in Linear Program (LP) form

2.3.1 1-norm

Let A be m x n matrix. Firstly, we need to add variables $t_1...t_m$ corresponding to rows of m.

$$\begin{bmatrix} min.\Sigma_{i=1}^{m} t_i \\ |A_i\vec{x}\,| \leq t_i \\ t_1...t_m \geq 0 \end{bmatrix} \rightarrow \begin{bmatrix} min.\Sigma_{i=1}^{m} t_i \\ A\vec{x}\ \leq \vec{t} \\ -A\vec{x}\ \leq \vec{t} \\ t_1...t_m \geq 0 \end{bmatrix} \tag{10}$$

After that, the equation (6) above could be written into the general form as shown below

$$\begin{aligned} &\text{maximize} && -\vec{1}^T\vec{t} \\ &\text{s.t} && A\vec{x} - \vec{t} \leq 0 \\ &&& -A\vec{x} - \vec{t} \leq 0 \\ &&& t_1...t_m \geq 0 \end{aligned} \tag{11}$$

2.3.2 ∞-norm

Let A be m x n matrix. Firstly, we need to add variables $t_1...t_m$ corresponding to rows of m.

$$\begin{bmatrix} min.t \\ |A_i\vec{x}\,| \leq t \\ t \geq 0 \end{bmatrix} \rightarrow \begin{bmatrix} min.t \\ A\vec{x}\ \leq t\cdot\vec{1} \\ -A\vec{x}\ \leq t\cdot\vec{1} \\ t \geq 0 \end{bmatrix} \tag{12}$$

After that, the equation (8) above could be written into the general form as shown below

$$\begin{aligned} &\text{maximize} && t \\ &\text{s.t} && A\vec{x} - t\cdot\vec{1}\ \leq \vec{0} \\ &&& -A\vec{x} - t\cdot\vec{1}\ \leq \vec{0} \\ &&& t \geq 0 \end{aligned} \tag{13}$$

2.4 Question d : Solve the following linear regression problem

2.4.1 Solve the first objective

Minimize the following objective function

$$\|Ax^* - b\|_1 \tag{14}$$

This could be done by first converting the equation above into a linear programming minimization problem form. The Linear Programming (LP) form is as shown below

$$\begin{aligned} &\text{minimize} && \mathbf{1}^T s \\ &\text{s.t} && Ax - b \leq s \\ &&& Ax - b \geq -s \end{aligned} \tag{15}$$

in the variables $x \in \mathbb{R}^{\mathrm{n}}$ and $s \in \mathbb{R}$. The constraint according to LP formulation above is

$$-s_k \leq a_k^T x - b_k \leq s_k \tag{16}$$

Hence, for each k imply that

$$s_k \geq |a_k^T x - b_k| \tag{17}$$

which shows the optimum over s could be chosen as shown below due to the fact that the objective function is separable

$$s = |a_k^T x - b_k| \tag{18}$$

Therefore, by fixing x, the optimal value of the LP is $p^*(x) = \|Ax - b\|_1$. Hence, we could prove that minimizing eq.(6) is equal to the LP form (7) if we optimize x and t at the same time. The optimal x and the optimal s are the minimizer and the minimum of the original problem, respectively.
In MATLAB, we could form an augmented matrix for the 1-norm as shown below

$$A_{aug} = \begin{bmatrix} A & -\mathbf{I_m} \\ -A & -\mathbf{I_m} \end{bmatrix} \tag{19}$$

where m is the row size of matrix input A. Moreover, we could form the objective function to solve with linprog as expressed below

$$f = \begin{bmatrix} 0_{nx1} & 1_{nx1} \end{bmatrix} \tag{20}$$

where n is the row size of the input b matrix. If we take the following matrices as an input

$$A = \begin{bmatrix} 2 & 5 \\ 0 & 1 \end{bmatrix} \tag{21}$$

$$b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Through the process explained above, we will get the minimization solution using LP expressed as below

$$\begin{bmatrix} x_1 \\ x_2 \\ s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} \tag{22}$$

as well as optimal solution x

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

2.4.2 Solve the second objective

Minimize the following objective function

$$\|Ax^* - b\|_\infty \tag{23}$$

This could be done by first converting the equation above into a linear programming minimization problem form. The Linear Programming form is as shown below

$$\begin{aligned} \text{minimize} \quad & t \\ \text{s.t} \quad & Ax - b \leq t\mathbf{1} \\ & Ax - b \geq t\mathbf{1} \end{aligned} \tag{24}$$

in the variables $x \in \mathbb{R}^n$ and $t \in \mathbb{R}$, where $\mathbf{1}$ symbolize row vector of ones. The constraint according to LP formulation above is

$$-t \leq a_k^T x - b_k \leq t \tag{25}$$

To prove that these are indeed equivalent, Assume that only t is optimized in this problem by excluding x in the optimization (fixing x). Then, we could notice that for each k which imply

$$t \geq |a_k^T x - b_k| \rightarrow t \geq \max_k |a_k^T x - b_k| = \|Ax - b\|_\infty \tag{26}$$

Therefore, by fixing x, the optimal value of the LP is $p^*(x) = \|Ax - b\|_\infty$. If we optimize over x and t simultaneously will prove that the LP formulation is the same thing with the initial objective function. The optimal x and the optimal t are the minimizer and the minimum of the original problem, respectively. The equation (16) above can be written as below

$$\begin{aligned} &\text{minimize} \quad t \\ &\text{s.t} \qquad\quad t\mathbf{1_m} \leq Ax - b \leq t\mathbf{1_m} \end{aligned} \tag{27}$$

Moreover, in MATLAB, we could expressed the inequality

$$Ax - t1_m \leq b \tag{28}$$

as

$$\begin{bmatrix} A & -1_m \end{bmatrix} \begin{bmatrix} x \\ t \end{bmatrix} \leq b$$

Therefore, the augmented matrix A for ∞-norm become as shown below

$$A_{aug} = \begin{bmatrix} A & -\mathbf{1_m} \\ -A & -\mathbf{1_m} \end{bmatrix} \tag{29}$$

Moreover, we could form the objective function to solve with linprog as expressed below

$$f = \begin{bmatrix} 0_{nx1} \\ 1 \end{bmatrix} \tag{30}$$

where n is the row size of the input b matrix. Next, if we take the input value as defined above in equation (17),through the process explained above, we will get the minimization solution using LP expressed as below

$$\begin{bmatrix} x_1 \\ x_2 \\ t \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} \tag{31}$$

as well as optimal solution x could be defined as

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

2.5 Question e: Solve the Optimization problem of SISO Constrained optimal control using Quadratic Programming

2.5.1 Solve the QP

Let we define the SISO process as below

$$x_{k+1} = 0.5x_k + u_k \tag{32}$$

with initial state $x_0 = 2$. Then from here we could construct the equality constraint by initially inputting k = 0,1 to the state space equation (28)

1. for k = 0
$$x_1 = 0.5x_0 + u_0 = 1 + u_0 \tag{33}$$
2. for k = 1
$$x_2 = 0.5x_1 + u_1 \tag{34}$$

From these two conditions, we could construct the matrix A_{eq} and b_{eq} to be inputted to the quadprog solver in MATLAB.

$$A_{eq}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = b_{eq} \tag{35}$$

$$A_{eq} = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0.5 & 0 & 0 & 1 \end{bmatrix}, b_{eq} = \begin{bmatrix} 1 \\ 0.5 \end{bmatrix}$$

After that, we need to define the inequality constraint with the lower bound (lb) and upper bound (ub) as below

$$\begin{bmatrix} 2.5 \\ -1 \\ -2 \\ -2 \end{bmatrix} \leq \begin{bmatrix} x_1 \\ x_2 \\ u_0 \\ u_1 \end{bmatrix} \leq \begin{bmatrix} 5 \\ 1 \\ 2 \\ 2 \end{bmatrix} \tag{36}$$

$$l_b = \begin{bmatrix} 2.5 \\ -1 \\ -2 \\ -2 \end{bmatrix}, u_b = \begin{bmatrix} 5 \\ 1 \\ 2 \\ 2 \end{bmatrix}$$

Then, defining the H matrix which is the Hessian matrix in the quadprog function

$$\frac{1}{2}x^T Hx = \frac{1}{2}\begin{bmatrix} x_1 \\ x_2 \\ u_0 \\ u_1 \end{bmatrix} H \begin{bmatrix} x_1 & x_2 & u_0 & u_1 \end{bmatrix} = \frac{1}{2}(x_1^2 + x_2^2 + u_0^2 + u_1^2) \tag{37}$$

Hence,

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

After that, the matrices defined in equation (31),(32),(33) are then used in MATLAB to solve QP problem. The optimal solution to this problem is as below

$$x^o = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2.5000 \\ 0 \end{bmatrix}, u^o = \begin{bmatrix} u_0 \\ u_1 \end{bmatrix} = \begin{bmatrix} 1.5000 \\ -0.7500 \end{bmatrix} \tag{38}$$

2.5.2 Define whether the KKT condition hold

The KKT conditions for QP with inequality constraints could be shown as below For the problems of the form :

$$\begin{aligned} \min_x \quad & f(x) \\ \text{s.t} \quad & g(x) \leq 0 \\ & h(x) = 0 \end{aligned} \tag{39}$$

The KKT conditions are :

$$\nabla f(x^*) + \Sigma_{i=1}^m \nabla g_i(x^*)\mu_i^* + \Sigma_{j=1}^l \nabla h_j(x^*)\lambda_j^* = 0 \tag{40}$$

$$\mu^* \geq 0 \tag{41}$$

$$g_i(x^*) \leq 0, h_j(x^*) = 0 \tag{42}$$

$$\mu_i g_i(x^*) = 0, i = 1, ..., m \tag{43}$$

The Lagrangian is :

$$\mathcal{L}(x, \mu, \lambda) = \frac{1}{2}x^T Hx + \mu^T g(x) + \lambda^T h(x) \tag{44}$$

To know if the problems that we have in section (2.5.1) are satisfying the KKT conditions, we could define the KKT conditions paramater as below

1. Write out the Lagrangian multipliers

$$\mu_i(lb) = \begin{bmatrix} 4.375 \\ 0 \\ 0 \\ 1.127 \cdot 10^{-11} \end{bmatrix}, \mu_i(ub) = \begin{bmatrix} 0 \\ 2.46 \cdot 10^{-11} \\ 4.54 \cdot 10^{-10} \\ 0 \end{bmatrix}, \lambda_j = \begin{bmatrix} 1.5 \\ 0.75 \end{bmatrix} \tag{45}$$

2. Assume that each lower bound and upper bound corresponds to an inequality constraint g and the equality constraints correspond to h

$$g_1(lb) = 2.5, g_2(lb) = -1, g_3(lb) = -2, g_4(lb) = -2, \tag{46}$$

$$g_1(ub) = 5, g_2(ub) = 1, g_3(ub) = 2, g_4(ub) = 2$$

$$h_1 = 1, h_2 = 0.5$$

3. Check for active and inactive constraints
   Based on eq (41) and (42) above, it could be said that the inequality constraints $(\mathbf{g_2(lb)}, \mathbf{g_3(lb)}, \mathbf{g_4 lb})$ are inactive since $g_i(x) < 0$. The rest of the constraints are active since $g_i(x) \geq 0$. Moreover, for active constraints, the lagrange multiplier μ can be divided into two groups, either it is $\mu_i > 0$ *strictly active* or it is $\mu_i = 0$ *not strictly active*. From this definition, $\mu_{1,4}(ub)$ are *not strictly active* while $\mu_1(lb)$ and $\mu_{2,3}(ub)$ are *strictly active*.

4. Check the KKT conditions
   The 1st condition is called the stationary condition. Assume all constraints are binding

$$x_1 = 2.5 \longrightarrow x_1 + \mu_1(lb) \cdot 1 + \mu_1(ub) \cdot 1 + \lambda_1 \cdot 1 + \lambda_2 \cdot 0.5 = 8.75 \neq 0 \tag{47}$$

$$x_2 = 0 \longrightarrow x_2 + \mu_2(lb) \cdot 1 + \mu_2(ub) \cdot 1 = 2.46 \cdot 10^{-11} \approx 0 \tag{48}$$

$$u_0 = 1.5 \longrightarrow u_0 + \mu_3(lb) \cdot 1 + \mu_3(ub) \cdot 1 + \lambda_1 \cdot -1 = 4.54 \cdot 10^{-10} \approx 0 \tag{49}$$

$$u_1 = -0.75 \longrightarrow u_1 + \mu_4(lb) \cdot 1 + \mu_4(ub) \cdot 1 + \lambda_2 \cdot 0.5 = 1.127 \cdot 10^{-11} \approx 0 \tag{50}$$

   Hence, we could conclude that the solution generally satisfy the 1st condition except for the x1 value which explicitly did not satisfy the 1st condition. Therefore, improvement towards the first inequality constraint is possible to get more optimal solution.

   The 2nd condition is called the dual feasibility condtion. This condition is clearly satisfied if we observe the matrices in eq.(41).

   The 3rd condition is called the primal feasibility condition. This condition is clearly satisfied due to the fact that *the MATLAB quadprog exitflag shows the value of '1'* indicating that the function converged to the solution x. On the other hand, if exit flag output is -5, it means that the primal and dual feasibility are both not satisfied.

   The 4th condition is called the complementary slackness. In MATLAB, this could be expressed as the complementarity measures($\Sigma_i \mu_i g_i$). It is lower than the optimality tolerance ($< 10^{-8}$), we could generally say that the system satisfy the 4th condition with slight tolerable breach. However, if we want to be specific, the first multiplier ($\mu_1 \cdot g1 = 4.375 \cdot 2.5 = 10.94 >> 0$) explicitly did not satisfy the 4th KKT condition. Hence, there is a possibility for improvement for the 1st inequality constraint here.

2.5.3 Determine what happen if lower bound and upper bound removed

1. For the case when lower bound is removed :

$$x^o = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0.555 \\ 0 \end{bmatrix}, u^o = \begin{bmatrix} u_0 \\ u_1 \end{bmatrix} = \begin{bmatrix} -0.444 \\ 0.222 \end{bmatrix} \tag{51}$$

$$\mu_i(lb) = \begin{bmatrix} 0.2568 \cdot 10^{-10} \\ 0 \\ 0.2568 \cdot 10^{-10} \\ 0 \end{bmatrix}, \mu_i(ub) = \begin{bmatrix} 0 \\ 0.1916 \cdot 10^{-9} \\ 0 \\ 0.0076 \cdot 10^{-9} \end{bmatrix}, \lambda_i = \begin{bmatrix} -0.444 \\ -0.222 \end{bmatrix} \tag{52}$$

   It could be observed that removing the lower bound will change the optimal state and input as well as changing the lagrangian multipliers. If we compare to the initial condition, we will have *no strictly active Lagrange multiplier at all.* Moreover, the relative maximum (constraint violation) parameter in the quadprog solver seems to be decreased ($1.24 \cdot 10^{-11} \rightarrow 5.55 \cdot 10^{-17}$) as well as the objective function at the solution **'fval'** which was decreased as well ($4.5313 \rightarrow 0.2778$) which indicates when the lower bound of x_1 is removed, better minimization achieved. Moreover, the first order optimality also decreased when lower bound is removed, which means improved performance. In addition, the first multiplier becomes approximately zero ($\mu_1 \cdot g1 = (0.2568 \cdot 10^{-10}) \cdot 2.5 = 6.4 \cdot 10^-11 \approx 0$), hence satisfying the complementary slackness condition. Finally, it could also be observed that the 1st KKT condition is all satisfied for all optimal solution found.

2. For the case when upper bound is removed : The KKT conditions was not satisfied due to the fact that the solver cannot solve the problem when there is any lower bound which exceed its corresponding upperbound since MATLAB automatically set the upperbound to infinity if we remove the upperbound for x_1, hence creating infeasibility in the solver.