SSY281 - Model Predictive Control
Assignment A04 - Optimization basics and QP problems

Lucas Rath - **Group 09**

Question 1

A) If the optimization problem is convex, then for all $x_1, x_2 \in S \Rightarrow z = \theta x_1 + (1-\theta)x_2 \in S,\ 0 \leq \theta \leq 1$, where S is the feasible set $S = \{\ x \mid g_i(x) \leq 0, h_i(x) = 0, \forall i\}$. Therefore, for $\theta = 0.5 \Rightarrow z = (x_1 + x_2)/2$ and z must belong to the feasible set S.

B) If the optimization problem is convex and $a + b = 1,\ \ 0 \leq a \leq 1$ then $z = ax_1 + bx_2$ will be never worse than both x_1 and x_2 at the same time. Moreover, since $a + b = 1$ we can rewrite $a = \theta$ and $b = (1 - \theta)$ and according to the convexity theorem, z will be feasible and also:

$$\underbrace{f(\theta x_1 + (1-\theta)x_2)}_{f(z)} \leq \underbrace{\theta f(x_1) + (1-\theta)f(x_2)}_{\Omega}, \qquad \forall\, 0 \leq \theta \leq 1 \tag{1}$$

Moreover, one can easily verify that:

$$f(x_1) \leq \Omega \leq f(x_2), \quad \text{if} \quad f(x_2) \geq f(x_1) \tag{2}$$
$$f(x_2) \leq \Omega \leq f(x_1), \quad \text{if} \quad f(x_1) \geq f(x_2) \tag{3}$$

Therefore, if $f(z) \leq \Omega$, it follows that necessarily $f(z)$ is less than either $f(x_1)$ or $f(x_2)$. More formally: $f(z) \leq \max(f(x_1), f(x_2))$.

Indeed, it can happen that z is better than both x_1 and x_2. For instance, if x_1 and x_2 are chosen such that $f(x_1) = f(x_2)$, then from (1), it follows that $f(z) \leq f(x_1) = f(x_2)$. In this particular case, $f(z)$ will be necessarily equal or better than both x_1 and x_2.

C) Clearly, this condition is just a particular case from last item, when $\theta = 0.5$. Therefore all the assumptions made in the last item still hold.

Question 2

A certain set is said to be convex if all the line segments between two points of the set are also part of the set:

$$x_1, x_2 \in S \quad \Rightarrow \quad \theta x_1 + (1-\theta)x_2\ \in\ S, \quad 0 \leq \theta \leq 1 \tag{4}$$

A) First, lets reformulate the definition of $S = \{x \in \Re^n \mid \alpha \leq a^T x \leq \beta\}$ as an intersection of two sets S_1 and S_2:

$$\begin{aligned} &S = S_1 \cap S_2 \\ &S_1 = \{x \in \Re^n \mid f_1(x) \leq 0\} \qquad f_1(x) = a^T x - \beta \\ &S_2 = \{x \in \Re^n \mid f_2(x) \leq 0\} \qquad f_2(x) = -a^T x + \alpha \end{aligned} \tag{5}$$

Then, it can be proven that S_1 is a convex set, i.e. if $x_1, x_2 \in S_1$, then also $\theta x_1 + (1-\theta)x_2 \in S_1$:

$$x_1, x_2 \in S_1 \quad \Rightarrow \quad f_1(x_1), f_1(x_2) \leq 0$$

$$f_1\big(\theta x_1 + (1-\theta)x_2\big) = a^T\left(\theta x_1 + (1-\theta)x_2\right) - b = \theta \underbrace{f_1(x_1)}_{\leq 0} + (1-\theta)\underbrace{f_1(x_2)}_{\leq 0} \leq 0$$

$$\Rightarrow \quad \theta x_1 + (1-\theta)x_2 \in S_1 \tag{6}$$

It has then be shown that S_1 is convex. Following the same logic, it can be easily proven that S_2 is also a convex set. Finally, the intersection between convex sets is also convex. Therefore, S is convex.

B) The set M is convex if for all $x_1, x_2 \in M$, then also $\theta x_1 + (1-\theta)x_2 \in M$:

$$M = \{x \mid g(x) \leq 0\}, \qquad g(x) = \|x - y\| - f(y) \quad \forall y \in S \subseteq \mathbb{R}^n$$

$$x_1, x_2 \in M \quad \Rightarrow \quad g(x_1), g(x_2) \leq 0$$

$$g\big(\theta x_1 + (1-\theta)x_2\big) = \|\theta x_1 + (1-\theta)x_2 - y\| - f(y) \tag{7}$$

Now, by using the triangle inequality $|a+b| \leq |a| + |b|$:

$$g\big(\theta x_1 + (1-\theta)x_2\big) \leq \theta\big(\|x_1 - y\| - f(y)\big) + (1-\theta)\big(\|x_2 - y\| - f(y)\big) = \theta \underbrace{g(x_1)}_{\leq 0} + (1-\theta)\underbrace{g(x_2)}_{\leq 0} \leq 0$$

$$\Rightarrow \quad g\big(\theta x_1 + (1-\theta)x_2\big) \leq 0 \tag{8}$$

Equation 8 has then proven that the set M is convex for all $y \in S$.

C) First we can rewrite the definition of the convex set in the following way:

$$M = \{x \mid g(x) \leq 0\} \tag{9}$$

where:

$$\begin{aligned} g(x) &= \|x - x_0\| - \|x - y\| \\ &= (x - x_0)^T(x - x_0) - (x - y)^T(x - y) \\ &= 2(y - x_0)x - (y^T y - x_0^T x_0) \end{aligned} \tag{10}$$

As can be notice, for a fixed y, the set M is a halfspace, which as known to be convex. Therefore, since $y \in S$, the set M can be reformulated as an intersection of halfspaces:

$$\bigcap_{y \in S} \{x \mid \|x - x_0\| - \|x - y\| \leq 0\} \tag{11}$$

and as explained before, the intersection of convex sets is also an convex set. Therefore M is convex.

Question 3

1-Norm)

The 1-norm objective is given by:

$$\begin{aligned} \min_{x} \quad & |Ax|_1 \\ s.t. \quad & Fx \leq g \end{aligned} \tag{12}$$

whose cost function can be rewritten as:

$$|Ax|_1 = \sum_i |a^i x| \tag{13}$$

where a^i is the i-th row of the matrix A. Each term $|a^i x|$ can then be read as equal to the smaller number σ_i that satisfies the following equation:

$$|a^i x| = \begin{cases} \min\limits_{\sigma_i} & \sigma_i \\ s.t. & \sigma_i \geq |a^i x| \end{cases} \tag{14}$$

Coupling this statement above in (12) and (13), we have:

$$\begin{aligned} \min_{x} \quad & \sum_i \min_{\sigma_i} \sigma_i \\ s.t. \quad & \begin{array}{l} Fx \leq g \\ \sigma_i \geq +a^i x \quad 1 \leq i \leq N \\ \sigma_i \geq -a^i x \end{array} \end{aligned} \tag{15}$$

which can be finally rewritten in the standard form as:

$$\begin{aligned} \min_{x,\sigma} \quad & \begin{bmatrix} \mathbf{0} \\ \mathbf{1} \end{bmatrix}^T \begin{bmatrix} x \\ \sigma \end{bmatrix} \\ s.t. \quad & \begin{bmatrix} F & 0 \\ A & -I \\ -A & -I \end{bmatrix} \begin{bmatrix} x \\ \sigma \end{bmatrix} \leq \begin{bmatrix} g \\ 0 \\ 0 \end{bmatrix} \end{aligned} \tag{16}$$

where $\mathbf{0}$ and $\mathbf{1}$ are column vectors of zeros and ones respectively, with the same dimension as $x, \sigma \in \mathbb{R}^n$. Note that now the optimization variables are x and σ and the problem is clearly a linear optimization problem.

∞-Norm)

The infinity norm optimization problem is given bellow:

$$\begin{aligned} \min_{x} \quad & |Ax|_\infty \\ s.t. \quad & Fx \leq g \end{aligned} \tag{17}$$

which cost function can be also defined as:

$$|Ax|_\infty = \max_i |a^i x| = \begin{cases} \min\limits_{\gamma} & \gamma \\ s.t. & \mathbf{1}\gamma \geq |Ax| \end{cases} \tag{18}$$

where again a^i is the i-th row of the matrix A and $\gamma \in \mathbb{R}^1$ is an auxiliary variable. Coupling (18) in (17), and using matrix representation, we can rewrite the optimization function:

$$\begin{aligned} \min_{x,\gamma} \quad & \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix}^T \begin{bmatrix} x \\ \gamma \end{bmatrix} \\ s.t. \quad & \begin{bmatrix} F & 0 \\ A & -\mathbf{1} \\ -A & -\mathbf{1} \end{bmatrix} \begin{bmatrix} x \\ \gamma \end{bmatrix} \leq \begin{bmatrix} g \\ 0 \\ 0 \end{bmatrix} \end{aligned} \tag{19}$$

Question 4

A) To adapt to this new cost function, trivial changes must be made in equation (16):

$$\begin{aligned} \min_{x,\sigma} \quad & \begin{bmatrix} \mathbf{0} \\ \mathbf{1} \end{bmatrix}^T \begin{bmatrix} x \\ \sigma \end{bmatrix} \\ s.t. \quad & \begin{bmatrix} A & -I \\ -A & -I \end{bmatrix} \begin{bmatrix} x \\ \sigma \end{bmatrix} \leq \begin{bmatrix} b \\ -b \end{bmatrix} \end{aligned} \tag{20}$$

B) The same can be made with equation (19):

$$\begin{aligned} \min_{x,\gamma} \quad & \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix}^T \begin{bmatrix} x \\ \gamma \end{bmatrix} \\ s.t. \quad & \begin{bmatrix} A & -\mathbf{1} \\ -A & -\mathbf{1} \end{bmatrix} \begin{bmatrix} x \\ \gamma \end{bmatrix} \leq \begin{bmatrix} b \\ -b \end{bmatrix} \end{aligned} \tag{21}$$

Question 5

Solving this problem in Matlab requires reformulating the constraints in the following form:

$$\begin{aligned} x_1 &\leq 5 \\ x_2 &\leq 1 \\ u_0 &\leq 2 \\ u_0 &\leq 2 \\ -x_1 &\leq -2.5 \\ -x_2 &\leq 1 \\ -u_0 &\leq 2 \\ -u_0 &\leq 2 \end{aligned} \qquad \text{and} \qquad \begin{aligned} x_1 - u_0 &= 1 \\ x_2 - 0.5x_1 - u_1 &= 0 \end{aligned} \tag{22}$$

After that, we can put those constraints in the standard Matlab format ($Ax \leq b$ and $A_{eq}x = b_{eq}$) and solve for the given cost function.

A) Solving the quadratic optimization problem leads to the following results:

```
x =
    2.5000      % x1
    0.6250      % x2
    1.5000      % u0
   -0.6250      % u1
fval =
    4.6406      % cost function at x*
```

B) The lagrange multipliers μ^* calculated with `quadprog` are:

```
lambda.ineqlin =
         0
         0
         0
         0
    4.3125
         0
         0
         0
```

To check if the solution satisfies the KKT conditions we need to verify 5 conditions. Very straightforward, the first 4 conditions can be confirmed to hold:

$$\begin{aligned} \mu^* &\geq 0 \\ h(x^*) &= 0 \\ g(x^*) &\leq 0 \\ \mu_i \; g_i(x^*) &= 0 \end{aligned} \tag{23}$$

We can easily see that all μ^* are greater than zero, all the constraints $h_i(x^*)$ and $g_i(x^*)$ are satisfied, and that either μ^* or $g_i(x^*)$ are zero for each constraint.

The last KTT condition follows:

$$\begin{aligned} &\nabla f(x^*) + \nabla g(x^*)\mu^* + \nabla h(x^*)\lambda^* \\ = &\begin{bmatrix} x1^* \\ x_2^* \\ u_0^* \\ u_1^* \end{bmatrix} + A^T\mu^* + A_{eq}^T\lambda^* \quad = \begin{bmatrix} 2.5 \\ 0.625 \\ 1.5 \\ -0.625 \end{bmatrix} + \begin{bmatrix} -4.3125 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1.8125 \\ -0.6250 \\ -1.5000 \\ 0.6250 \end{bmatrix} = \mathbf{0} \end{aligned} \tag{24}$$

Therefore, all the KKT conditions are satisfied.

As can be easily confirmed, the only active constraint is $-x_1 \leq -2.5$, since $\mu^* \neq 0$ for this constraint.

C) If the lower bound of x_1 is removed then another solution is found and the cost function now is even lower. Mainly, what was avoiding the cost function to reduce in the last question was exactly the lower bound of x_1, since on the optimal point its Lagrangian multiplier was different than zero.

On the other hand, removing the upper bound constraint of x_1 would in this case not change the result and the optimal value is bounded.