SSY281 - Model predictive control
Assignment 3

Nicholas Granlund

February 20, 2023

Question 1: Constrained optimization

(a) What is a convex/strictly convex function?
A convex function is a function where any line segment connecting two points on the graph of the function lies entirely above or on the graph. i.e for a funtion f the following condition must hold for any $\theta \in [0, 1]$ to imply convexity:

$$f\big(\theta x_1 + (1-\theta)x_2\big) \;\; \leq \;\; \theta f(x_1) + (1-\theta)f(x_2)$$

For a strictly convex function, the above condition must hold, but instead for strict-inequality (i.e $<$ instead of $\leq$). An illustration of convex functions can be seen in Figure 1.

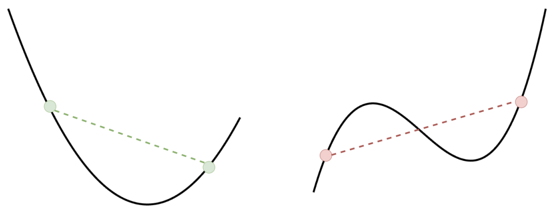

Figure 1: Two functions. One convex and one non-convex. The rightmost function has a line segment where not all points are above the graph between the points.

(b) What is a convex set?
A convex set is a set which contains line segmemts between every two points in the set (meaning all point between are part of the set). The formal mathematical definition is:

$$x_1, x_2 \in \mathcal{S} \implies \theta x_1 + (1-\theta)x_2 \in \mathcal{S},\; 0 \leq \theta \leq 1$$

This is illustrated by Figure 15 in the lecture notes. however that figure has been edited and are shown in figure 2 for a visual interpretation of convex sets.

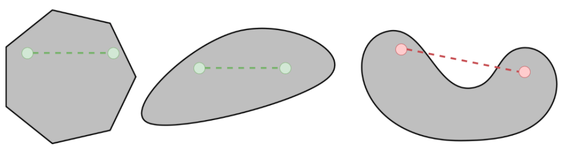

Figure 2: *Figure (15) revised from L.N p.70.* Two convext sets and one non-convext set with line segment between two arbitrary points to illustrate convexity. Rightmost set is non-convex since a linesegment are not entierly in the set.

(c) Under what conditions on function f, g, h does the optimization problem becomes a convex optimization problem?
The optimization problem becomes convex if the following holds:

1. The objective function $f(x)$ has to be convex.
2. The inequality constraints defined by $g(x)$ has to be convex.
3. The equality constraints defined by $h(x)$ has to be affine.

Question 2: Convexity

(a) Is set S_1 convex?
Yes, this set is convex. This can be shown by splitting the set S into an intersection between two separate sets.

$$S_1 = \{x \in \mathbb{R}^n | \alpha \leq a^\top x \leq \beta\} = S_{11} \cap S_{12}$$

where

$$S_{11} = \{x \in \mathbb{R}^n | \underbrace{a^\top x - \beta}_{f_{11}(x)} \leq 0\}$$
$$S_{12} = \{x \in \mathbb{R}^n | \underbrace{-a^\top x + \alpha}_{f_{12}(x)} \leq 0\}$$

Now we can prove that both of these separate sets are convex. Lets pick two arbitrary points x_1 and x_2 that belongs to the set S_{11}. Now we can prove the set convexity by using the function $f_{11}(x)$ which describes the points in the set S_{11}. We get:

$$f_{11}\big(\theta x_1 + (1-\theta)x_2\big) = a^\top \big(\theta x_1 + (1-\theta)x_2\big) - \beta$$
$$a^\top \big(\theta x_1 + (1-\theta)x_2\big) - \beta \leq \underbrace{\theta \underbrace{f_{11}(x_1)}_{\leq 0} + (1-\theta) \underbrace{f_{11}(x_2)}_{\leq 0}}_{\leq 0}$$

Hence we have that:

$$\theta x_1 + (1-\theta)x_2 \in S_{11}$$

Which implies that every point on the line segment belongs to the set S_{11}. I.e the set is convex. The set S_{12} can be shown to be convex by following the same derivations as for set S_{11}. Moreover, since it is known that the intersection of two convex sets also is a convex set, we can conclude that the set S_1 is indeed convex.

(b) Is set S_2 convex?
Yes, this set is also convex. The set

$$S_2 = \{x| \ ||x-y|| \leq f(y), \ \forall y \in S\}, \quad S \in \mathbb{R}^n$$

can be rewritten as

$$S_2 = \{x| \ ||x-y|| - f(y) \leq 0, \ \forall y \in S\}, \quad S \in \mathbb{R}^n$$

And with this we can yet again determine the convexity by analysing the function which defines the set S_2, i.e $S_2 = \{x| \ g(x) \leq 0\}$ where $g(x) = ||x-y|| - f(y)$. The set S_2 is then convex if for arbitrary points x_1 and x_2, every point in between $\theta x_1 + (1-\theta)x_2$ also belongs to the set. We assume two points $x_1, x_2 \in S_2$ and get:

$$g(\theta x_1 + (1-\theta)x_2) = ||(\theta x_1 + (1-\theta)x_2) - y|| - f(y)$$

Here we can apply the traingel inequaity given as hint in the assignment $||a+b|| \leq ||a|| + ||b||$. and expand the expression to:

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

$$g(\theta x_1 + (1-\theta)x_2) \ \leq \ \theta g(x_1) + (1-\theta) g(x_2)$$

Hence we have proven that the function which defines the points in the set is convex. Thereby the set itself is convex.

(c) Is set S_3 convex?
Yes, this final set is also convex. The set

$$S_3 = \{(x, y) \mid \ y \leq 2^x, \ \ \forall (x, y) \in \mathbb{R}^2\}$$

Can be rewritten as

$$S_3 = \{(x, y) \mid \ \underbrace{\log(y) - x\log(2)}_{g(x,y)} \leq 0, \ \ \forall (x, y) \in \mathbb{R}^2\}$$

Again, lets assume two arbitrary points (x_1, y_1) and (x_2, y_2) which belongs to S_3. We then know that:

$$g(x_1, y_1), \ \ g(x_2, y_2) \leq 0$$

Similarly to the previous sets, we want to show that any point in the line segment also belongs to the set. Here goes...

$$y = \theta y_1 + (1 - \theta)y_2$$

$$\log(y) = \log(\theta y_1 + (1 - \theta)y_2)$$

$$\log(y) = \log(\theta e^{\log(y_1)} + (1 - \theta)e^{\log(y_2)})$$

$$\log(y) \leq \log(\theta e^{\log(y_1)}) + \log((1 - \theta)e^{\log(y_2)})$$

$$\log(y) \leq \theta \log(y_1) + (1 - \theta) \log(y_2)$$

Now we get with the help of $g(x, y)$

$$\log(y) - x\log(2) \leq \theta\Big(\log(y_1) - x_1\log(2)\Big) + (1 - \theta)\Big(\log(y_2) - x_2\log(2)\Big)$$

Since (x_1, y_2) and (x_2, y_2) belongs to S_3 we now know that $\log(y_1) - x_1\log(2) \leq 0$ and that $\log(y_2) - x_2\log(2) \leq 0$ Thereby we can conclude that any point in the line segment belongs to the set S_3

Question 3: Norm problems as linear programs

(a) Explain briefly why (5) and (3) yields the same result
The problems give same result because the two optimization problems are essentially minimizing the maximum absolute deviation between the vector Ax and b.

Problem (3), min x $||Ax - b||_\infty$, is to find which x that minimizes the largest deviation between the i-th component of Ax and b (infinity norm).

Problem (5), min ϵ s.t $-\epsilon \leq (Ax - b)_i \leq \epsilon$, is minimizing the size of the deviation ϵ of the i-th component of $(Ax - b)_i$, whilst the deviation is no larger than the value of ϵ.

The optimal solution for in (5) is the same as the optimal solution of the infinity norm of $(Ax - b)$, which essentially is optimization problem (3). Therefore the two problems are equivalent.

(b) Assuming $z^\top = [x^\top \quad \epsilon]$, what are $c^\top$, F and g if one wants to represent (5) as in (2)

Lets say that the vector $x^\top$ has length $n \times 1$. Then the concatenated vector $z^\top$ will have length $n + 1$. Since we want to turn the optimization problem into minimizing ϵ, we can simply put the vector $c^\top$ to contain n zeros and a 1 at the last position (pos $n + 1$). i.e $c^\top = [0, 0, ..., 0, 1]$ Hence we get :

$$\min_{[x^\top \epsilon]} \quad c^\top [x^\top \epsilon]^\top = \underbrace{\begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \end{bmatrix}}_{c^\top} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \\ \epsilon \end{bmatrix} = \epsilon$$

Now we can create F and g with the given constraint. Since we can reformulate the constraint to equal two separate constraints we get:

$$Ax_i - b \leq \epsilon \quad \text{and} \quad -\epsilon \leq Ax_i - b$$

These two constraints can be rewritten as:

$$Ax_i - \epsilon \leq b \quad \text{and} \quad -Ax_i - \epsilon \leq -b$$

And these can be expressed together in a matrix format as:

$$\underbrace{\begin{bmatrix} A & -1 \\ -A & -1 \end{bmatrix}}_{F_i} \underbrace{\begin{bmatrix} x_i \\ \epsilon \end{bmatrix}}_{z_i} \leq \underbrace{\begin{bmatrix} b \\ -b \end{bmatrix}}_{g_i}$$

The expression above defines the constraints for the i-th x. Now since we have n number of x, we have to enlarge the matrix quite a bit. We arrive at:

$$\underbrace{\begin{bmatrix} A & 0 & 0 & \cdots & 0 & -1 \\ -A & 0 & 0 & \cdots & 0 & -1 \\ 0 & A & 0 & \cdots & 0 & -1 \\ 0 & -A & 0 & \cdots & 0 & -1 \\ 0 & 0 & A & \cdots & 0 & -1 \\ 0 & 0 & -A & \cdots & 0 & -1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & A & -1 \\ 0 & 0 & 0 & \cdots & -A & -1 \end{bmatrix}}_{F} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \\ \epsilon \end{bmatrix}}_{z} \leq \underbrace{\begin{bmatrix} b \\ -b \\ b \\ \vdots \\ b \\ -b \end{bmatrix}}_{g}$$

Or in short:

$$Fz \leq g$$

Now I have to clarify that the 0 and -1 in the expression above have to fit the dimensions of A. Lets say that matrix A is of size $m \times n$, then every 0 is a matrix $0_{m\times n}$ and every -1 is a vector of $-1_{m\times 1}$ Every two rows of F defines the constraints for the i-th x. Hence F has $2n$ rows and $nm + 1$ columns. The g vector contains alternating values of b and $-b$ and has $2mn$ elements.

(c) Consider A and b. Solve the linear programs.

Simple enough this is done in Matlab, a function `solve_linear_prog()` was created which solves the linear program given matrices A and b. The computations was as follow;

$$\min_{x,\epsilon} \ \epsilon \qquad \text{s.t} \begin{bmatrix} A & -1 \\ -A & -1 \end{bmatrix} \begin{bmatrix} x \\ \epsilon \end{bmatrix} \leq \begin{bmatrix} b \\ -b \end{bmatrix}$$

Which with the given numerical values gives the problem:

$$\begin{bmatrix} \begin{bmatrix} 0.4889 & 0.2939 \\ 1.0347 & -0.7873 \\ 0.7269 & 0.8884 \\ -0.3034 & -1.1471 \end{bmatrix} & \begin{bmatrix} -1 \\ -1 \\ -1 \\ -1 \end{bmatrix} \\ - \begin{bmatrix} 0.4889 & 0.2939 \\ 1.0347 & -0.7873 \\ 0.7269 & 0.8884 \\ -0.3034 & -1.1471 \end{bmatrix} & \begin{bmatrix} -1 \\ -1 \\ -1 \\ -1 \end{bmatrix} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \epsilon \end{bmatrix} \leq \begin{bmatrix} \begin{bmatrix} -1.0689 \\ -0.8095 \\ -2.9443 \\ 1.4384 \end{bmatrix} \\ - \begin{bmatrix} -1.0689 \\ -0.8095 \\ -2.9443 \\ 1.4384 \end{bmatrix} \end{bmatrix}$$

Solving this linear program with Matlab yields the following result:

$$z^* = \begin{bmatrix} x_1^* \\ x_2^* \\ \epsilon^* \end{bmatrix} = \begin{bmatrix} -2.0674 \\ -1.1067 \\ 0.4583 \end{bmatrix}$$

(d) Write the dual of (2)
We have the problem:

$$\min_{x,\epsilon} \quad c^\top z \qquad \text{s.t} \qquad \underbrace{Fz - h}_{g(z)} \leq 0$$

Which we can rewrite. By moving the constraint to the objective function we get the familiar lagrangian

$$\mathcal{L}(z, \mu, \lambda) = c^\top z + \mu^\top (Fz - h) + \lambda^\top 0$$

$$\mathcal{L}(z, \mu, \lambda) = c^\top z + \mu^\top (Fz - h)$$

$$q(\mu, \lambda) = \inf_z \mathcal{L}(z, \mu, \lambda)$$

Now with this Lagrangian we can define the Dual function:

$$\max_{\mu} \quad -h^\top \mu$$

$$\text{s.t} \qquad \mu \geq 0$$
$$F^\top \mu + c^\top = 0$$

(e) Use the numerical values of A and b and solve the dual problem
Solving this linear program with Matlab gives the optimal solution for μ as given below. This result can also be confirmed by the solution for (d). In Matlab, it is possible to obtain μ as an output of linprog() for the initial linear program. This confirms that this optimal μ^* is correct.

$$\mu^* = \begin{bmatrix} 0.0000 \\ 0.0000 \\ 0.4095 \\ 0.4284 \\ 0.0000 \\ 0.1621 \\ 0.0000 \\ 0.0000 \end{bmatrix}$$

(f) Use the solution to find the solution of the primal one.

Question 4: Quadratic programming

(a) Solve the QP with Matlab

for this problem we want to solve the quadratic program according to:

$$\min_{x,u} \ \frac{1}{2}(x_1^2 + x_2^2 + u_0^2 + u_1^2)$$

$$\begin{aligned} \text{s.t} \qquad 2.5 \le x_1 &\le 5 \\ -0.5 \le x_2 &\le 0.5 \\ -2 \le u_0 &\le 2 \\ -2 \le u_1 &\le 2 \end{aligned}$$

$$x_{k+1} = Ax_k + bu_k$$

In this problem we have that $A = 0.4$ and $b = 1$. By considering the typical expression for a quadratic program we can rewrite this according to:

$$\min_{x,u} \ \frac{1}{2}\begin{bmatrix} x_1 & x_2 & u_0 & u_1 \end{bmatrix} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{Q} \begin{bmatrix} x_1 \\ x_2 \\ u_0 \\ u_1 \end{bmatrix} + \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}}_{p^\top} \begin{bmatrix} x_1 \\ x_2 \\ u_0 \\ u_1 \end{bmatrix}$$

$$\text{s.t} \qquad \underbrace{\begin{bmatrix} -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{G} \begin{bmatrix} x_1 \\ x_2 \\ u_0 \\ u_1 \end{bmatrix} \le \underbrace{\begin{bmatrix} -2.5 \\ 5 \\ 0.5 \\ 0.5 \\ 2 \\ 2 \\ 2 \\ 2 \end{bmatrix}}_{h}$$

$$\underbrace{\begin{bmatrix} 1 & 0 & -b & 0 \\ -A & 1 & 0 & -b \end{bmatrix}}_{A_{eq}} \begin{bmatrix} x_1 \\ x_2 \\ u_0 \\ u_1 \end{bmatrix} = \underbrace{\begin{bmatrix} Ax_0 \\ 0 \end{bmatrix}}_{b_{eq}}$$

Solving this with the initial condition $x_0 = 1.5$ with Matlab gives us the following results:

$$x^* = \begin{bmatrix} x_1^* \\ x_2^* \\ u_0^* \\ u_1^* \end{bmatrix} = \begin{bmatrix} 2.5000 \\ 0.5000 \\ 1.9000 \\ -0.5000 \end{bmatrix}$$

(b) Do the KKT hold for the solution? Which constraints are active?

To confirm if the KKT conditions hold for the solution, we need to verify 5 conditions. The first four conditions are easily determined:

$$\begin{aligned} &\mu^* \ge 0 \\ &g(x^*) \le 0 \\ &h(x^*) = 0 \\ &\mu_i^* g_i(x^*) = 0 \end{aligned}$$

All μ^* for the solution are either zero or positive. The inequality constraints are upheld at the solution x^* giving that $g(x^*) \leq 0$. The equality constraint is upheld for the solution implying that $h(x^*) = 0$. Also we can verify that $\mu_i^* g_i(x^*) = 0$ for every constraint. In this case, $\mu_1, \mu_4 \neq 0$ but luckily we have that $g_1(x^*), g_4(x^*) = 0$.

Now we have to determine if the main KKT condition is satisfied. That condition is:

$$\nabla f(x^*) + \nabla g(x^*)\mu^* + \nabla h(x^*)\lambda^* = 0$$

By calculating the respective gradients we get that:

$$\nabla f(x^*) = x^*$$

$$\nabla g(x^*)\mu^* = G^\top \mu^*$$

$$\nabla h(x^*)\lambda^* = A_{eq}^\top \lambda^*$$

Which shows that :

$$\begin{bmatrix} 2.5000 \\ 0.5000 \\ 1.9000 \\ -0.5000 \end{bmatrix} + \begin{bmatrix} -4.6000 \\ 0.0001 \\ 0.0000 \\ -0.0000 \end{bmatrix} + \begin{bmatrix} 2.1000 \\ -0.5000 \\ -1.9000 \\ 0.5000 \end{bmatrix} = 0$$

The solution thus satisifes all of the KKT conditions. Please note that if one uses many decimal places, the solution for $x_2 = 4.9999$ and therefor $\mu_4^* = 0.0001$. Either way the conditions are satisfied.

As can be seen by the results, two constraints are active. x_1 are at its lower bound and x_2 is at its upper bound. This is confirmed since the lagrange multipliers $\mu_1^*, \mu_4^* \neq 0$ for the lower constraint on x_1 and higher constraint on x_2.

(b) What would happen if we remove lower bound on x_1? What about the upper bound?

If we remove the lower constraint for x_1 the quadratic program finds another better solution which has a lower value for the objective function. It founds the better solution:

$$z^* = \begin{bmatrix} x_1^* \\ x_2^* \\ u_0^* \\ u_1^* \end{bmatrix} = \begin{bmatrix} 0.2885 \\ 0.0577 \\ -0.3115 \\ -0.0577 \end{bmatrix}$$

Evidently this solution has a lower value for x_1 which explains why the lower constraint on x_1 was active for the previous example. The objective function wanted to move x_1 against the constraint and now when there is no lower bound, a better solution could be obtained. We can also confirm that no constraints are active at this solution since all lagrange multipliers for the inequality constraints are 0.

On the other hand, when the upper bound on x_1 is removed, no better solution is found. The optimal solution is still unobtainable due to the lower constraint, so the upper bound does not affect the minimization of the objective function.

/ Nicholas Granlund