SSY281 Model Predictive Control 2023
PSS 3 - Optimization

Rémi Lacombe

February 14th, 2023

Lecture Refresh

Cf PSS notes.

Exercise 1

Norm-1 objective as an LP

Show that the norm-1 optimization problem:

$$\min_x \|Ax\|_1 \tag{1}$$

can be equivalently written as a linear program with the form:

$$\begin{aligned} \min_z \quad & c^\top z && (2a)\\ \text{s.t.} \quad & Fz \leq g. && (2b)\end{aligned}$$

Solution:

Consider any real number $y \in \mathbb{R}$. Its norm $|y|$ can be defined as the smallest non-negative ϵ value such that the inequalities $\epsilon \geq y$ and $\epsilon \geq -y$ both hold. This can equivalently be formulated as a minimization problem:

$$\begin{aligned} |y| = \min_\epsilon \quad & \epsilon && (3a)\\ \text{s.t.} \quad & \epsilon \geq y, && (3b)\\ & \epsilon \geq -y. && (3c)\end{aligned}$$

So if one wanted to, say, minimize the norm $|y|$, one could solve the following LP:

$$\begin{aligned} \min_{\epsilon, y} \quad & \epsilon && (4a)\\ \text{s.t.} \quad & \epsilon \geq y, && (4b)\\ & \epsilon \geq -y. && (4c)\end{aligned}$$

This can be generalized component by component and applied to the problem at hand:

$$\min_{x,\epsilon} \quad \mathbb{1}^\top \epsilon \tag{5a}$$

$$\text{s.t.} \quad \begin{bmatrix} A & -I \\ -A & -I \end{bmatrix} \begin{bmatrix} x \\ \epsilon \end{bmatrix} \leq \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{5b}$$

where it can be easily identified that $z = [x, \epsilon]^\top$, $c = [\mathbb{0}, \mathbb{1}]^\top$, $F = \begin{bmatrix} A & -I \\ -A & -I \end{bmatrix}$, and $g = [0, 0]^\top$.

Exercise 2

Convex QP and KKT conditions

Consider the objective function $V(x) = x_1^2 - x_1 x_2 + x_2^2 - 3x_1$ and the following QP problem:

$$\begin{aligned} \text{minimize} \quad & V(x) = x_1^2 - x_1 x_2 + x_2^2 - 3x_1 \\ \text{subject to} \quad & x_1 \geq 0 \\ & x_2 \geq 0 \\ & x_1 + x_2 \leq 2 \end{aligned} \tag{P-I}$$

1. Rewrite the optimization problem (P-I) in standard form.
2. Show that the problem (P-I) is convex.
3. Show that the solution obtained graphically from looking at Figure 1 satisfies the KKT optimality conditions (and hence that this solution is optimal indeed).

Solution:

As we know, the standard QP form is given by:

$$\begin{aligned} \text{minimize} \quad & V(x) = \frac{1}{2} x^T Q x + p^T x, \quad Q \succ 0 \\ \text{subject to} \quad & Gx \leq h, \quad G \in \mathbb{R}^{m \times n} \\ & Ax = b, \quad A \in \mathbb{R}^{p \times n}, \text{ (Note: } p \text{ in } \mathbb{R}^{p \times n} \text{ is not as same as } p \text{ in } V(x)) \end{aligned} \tag{Standard-QP}$$

1. Now let's write (P-I) in (Standard-QP) form:

$$V(x) = \frac{1}{2} \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} p_1 & p_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{6}$$

$$V(x) = \frac{1}{2} q_{11} x_1^2 + \frac{1}{2} \left(q_{12} + q_{21} \right) x_1 x_2 + \frac{1}{2} q_{22} x_2^2 + p_1 x_1 + p_2 x_2 \tag{7}$$

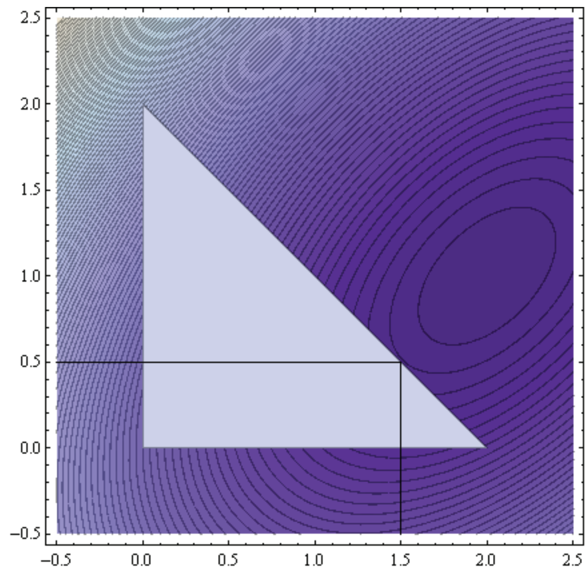

Figure 1: Feasible set overlayed on top of the level curves of the objective function. The level curves decrease in value from the outside to the inside, i.e. the innermost level curve represents the lowest value of the objective function compared to the outer level curves. Here, the constraint $x_1 + x_2 \leq 2$ is tangential to the "optimal" level curve $V(x) = -11/4$.

Comparing with $V(x) = x_1^2 - x_1x_2 + x_2^2 - 3x_1$ we get:

$$q_{11} = 2, \quad q_{22} = 2, \quad p_1 = -3, \quad p_2 = 0$$

$$\frac{1}{2}(q_{12} + q_{21}) = -1$$

Since Q must be symmetric:

$$q_{12} = q_{21} \Rightarrow \frac{1}{2}(q_{12} + q_{12}) = -1 \Rightarrow q_{12} = q_{21} = -1$$

Now the objective function can be written in standard form:

$$V(x) = \frac{1}{2}x^TQx + p^Tx$$

where

$$Q = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$$

$$p = \begin{bmatrix} -3 & 0 \end{bmatrix}^T$$

Now let's write the constraints in standard form:

$$\begin{bmatrix} -1 & 0 \\ 0 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \leq \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} \tag{8}$$

$$G\,x \leq h \tag{9}$$

There are only inequality constraints in this problem, so A and b are zero. Now the problem can be written in standard form:

$$\begin{aligned} \text{minimize} \;\; & V(x) = \frac{1}{2}x^T Q x + p^T x, \quad Q \succ 0 \\ \text{subject to} \;\; & Gx \leq h, \quad G \in \mathbb{R}^{3\times 2} \end{aligned} \tag{P-II}$$

2. Since the eigenvalues eig(Q) $= \{1, 3\}$ are all positive, $Q \succ 0$, which means that the objective function $V(x)$ is a strictly convex function. All the constraint functions are linear, which means that their corresponding linear inequalities define half-spaces, which we know are convex sets. The feasible set of this problem being the intersection of three half spaces, it is a convex set. Hence, the QP (P-I) is convex.

3. Graphically, we can hypothesize that $[1.5, 0.5]^\top$ is the optimal solution of (P-I). Now, using the KKT conditions, we will check this claim. Since the problem is a convex QP, the KKT conditions are both necessary **and** sufficient for global optimality.

   The Lagrangian of (P-II) is given by:

$$L(x, \lambda, \nu) = \frac{1}{2}x^T Q x + p^T x + \lambda^T (Gx - h) + \nu^T (Ax - b) \tag{10}$$

   We have following KKT conditions that must be satisfied for $x^* = \begin{bmatrix} 1.5 & 0.5 \end{bmatrix}^T$

$$\begin{aligned} \text{Equality Constraints:} \;\; & Ax^* - b = 0 && \text{(KKT-I)} \\ \text{Inequality Constraints:} \;\; & Gx^* - h \leq 0 && \text{(KKT-II)} \\ \text{Dual Constraints:} \;\; & \lambda^* \geq 0 && \text{(KKT-III)} \\ \text{Complementary Slackness:} \;\; & \lambda_i^*(g_i x^* - h_i) = 0, \quad \forall i = 1, \cdots, 3, && \text{(KKT-IV)} \\ & \lambda_i^* > 0 \Rightarrow (g_i x^* - h_i) = 0 \\ & (g_i x^* - h_i) < 0 \Rightarrow \lambda_i^* = 0, \\ \text{Gradient of Lagrangian Vanishes:} \;\; & Qx^* + p + G^T \lambda^* = 0 && \text{(KKT-V)} \end{aligned}$$

   Since we don't have any equality constraints in this problem, (KKT-I) needs not be checked here.

   The condition (KKT-II) can be verified as:

$$Gx^* - h = \begin{bmatrix} -1 & 0 \\ 0 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1.5 \\ 0.5 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} (g_1 x^* - h_1) \\ (g_2 x^* - h_2) \\ (g_3 x^* - h_3) \end{bmatrix} = \begin{bmatrix} -1.5000 \\ -0.5000 \\ 0 \end{bmatrix} \leq 0$$

Now from (KKT-IV) we get

$$
\begin{aligned}
\lambda_i^*(g_i x^* - h_i) &= 0, \quad \forall i = 1, \cdots, 3 && (11)\\
\Rightarrow -1.5\lambda_1^* = 0, \quad \because (g_1 x^* - h_1) < 0 &\Rightarrow \lambda_1^* = 0, && (12)\\
-0.5\lambda_2^* = 0, \quad \because (g_2 x^* - h_2) < 0 &\Rightarrow \lambda_2^* = 0 && (13)\\
0\lambda_3^* = 0, \quad \because (g_3 x^* - h_3) = 0 &\Rightarrow \lambda_3^* > 0 && (14)\\
\Rightarrow \lambda^* &= \begin{bmatrix} 0 & 0 & \lambda_3^* \end{bmatrix}^T && (15)
\end{aligned}
$$

Now we need to find λ_3^* using (KKT-V). Since $Qx^* + p = \begin{bmatrix} -0.5 \\ -0.5 \end{bmatrix}$, so from (KKT-V) we can write

$$
\begin{aligned}
\begin{bmatrix} -0.5 \\ -0.5 \end{bmatrix} + \begin{bmatrix} -1 & 0 & 1 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \lambda_3^* \end{bmatrix} &= 0 && (16)\\
\Rightarrow -0.5 + \lambda_3^* &= 0 && (17)\\
-0.5 + \lambda_3^* &= 0 && (18)\\
\Rightarrow \lambda_3^* &= 0.5 && (19)\\
\lambda^* &= \begin{bmatrix} 0 \\ 0 \\ 0.5 \end{bmatrix} && (20)
\end{aligned}
$$

So $\lambda^* \geq 0$ so (KKT-III) is also satisfied. Since all the KKT conditions are satisfied, $x^* = \begin{bmatrix} 1.5 & 0.5 \end{bmatrix}^T$ is the optimal solution indeed.

Exercise 3

Softening the constraints

Consider the following standard QP problem with *hard* linear inequality constraints on the optimization variable x:

$$
\begin{aligned}
\text{minimize} \quad & V(x) = \frac{1}{2} x^\top Q x + p^\top x \\
\text{subject to} \quad & Ax \leq b
\end{aligned}
\tag{21}
$$

In some practical applications, we can relax the hard constraints by introducing slack variables ϵ (this is called *constraint softening*). ϵ can be seen as the amounts by which constraints are allowed to be violated. These amounts should be kept small when possible, which can be done by adding a penalty on the size of ϵ in the cost function.

Following these guidelines, rewrite the QP (21) with soft constraints only and show that the new optimization problem is also a standard QP problem. You can use a quadratic penalty term for ϵ.

Solution:

By adding a non-negative vector of slack variables ϵ to the linear constraints of (21), and a quadratic penalty term on ϵ to the objective function, we get the following QP:

$$\begin{aligned} \text{minimize} \;\; & V(x,\epsilon) = \frac{1}{2}x^\top Qx + p^\top x + \rho\|\epsilon\|_2^2 \\ \text{subject to} \;\; & Ax \leq b + \epsilon, \\ & \epsilon \geq 0. \end{aligned} \tag{22}$$

Note that even if $Ax \leq b$ is not feasible, there always exists some $\epsilon \geq 0$ such that $Ax \leq b + \epsilon$, effectively making the QP (22) feasible.
This optimization problem can be rewritten as

$$\begin{aligned} \text{minimize} \;\; & V(x,\epsilon) = \frac{1}{2}\begin{bmatrix} x^T & \epsilon^T \end{bmatrix} \begin{bmatrix} Q & 0 \\ 0 & \rho I \end{bmatrix} \begin{bmatrix} x \\ \epsilon \end{bmatrix} + \begin{bmatrix} p^T & 0 \end{bmatrix} \begin{bmatrix} x \\ \epsilon \end{bmatrix} \\ \text{subject to} \;\; & \begin{bmatrix} A & -I \\ 0 & -I \end{bmatrix} \begin{bmatrix} x \\ \epsilon \end{bmatrix} \leq \begin{bmatrix} b \\ 0 \end{bmatrix} \end{aligned} \tag{23}$$

which is in the standard QP form in terms of the optimization variables x and ϵ.

Remark. *Note that price to pay for making the original problem feasible was that additional decision variables had to be added to the problem. In mathematical programming, this process is called lifting. Lifting can be carried out to soften hard constraints and facilitate the search for feasible solutions, as was done here, or can also be used to modify the structure of an optimization problem in order to make it easier to solve (e.g. by "changing" equality constraints into inequality constraints).*