

pyMPC Documentation

Marco Forgiione

February 12, 2021

1 Mathematical formulation

Let us consider a discrete-time linear dynamical system with n_u inputs and n_x states:

$$x_{k+1} = Ax_k + Bu_k, \quad (1)$$

with $A \in \mathbb{R}^{n_x \times n_x}$ and $B \in \mathbb{R}^{n_x \times n_u}$.

The MPC problem solved by pyMPC is:

$$\begin{aligned} \arg \min_U & \overbrace{\frac{1}{2} (x_N - x_{\text{ref}})^\top Q_{x_N} (x_N - x_{\text{ref}})}^{=J_{Q_{x_N}}} + \overbrace{\frac{1}{2} \sum_{k=0}^{N_p-1} (x_k - x_{\text{ref}})^\top Q_x (x_k - x_{\text{ref}})}^{J_{Q_x}} + \\ & + \overbrace{\frac{1}{2} \sum_{k=0}^{N_p-1} (u_k - u_{\text{ref}})^\top Q_u (u_k - u_{\text{ref}})}^{J_u} + \overbrace{\frac{1}{2} \sum_{k=0}^{N_p-1} \Delta u_k^\top Q_{\Delta u} \Delta u_k}^{J_{\Delta u}} \end{aligned} \quad (2a)$$

subject to :

$$x_{k+1} = Ax_k + Bu_k \quad (2b)$$

$$u_{\min} \leq u_k \leq u_{\max} \quad (2c)$$

$$x_{\min} \leq x_k \leq x_{\max} \quad (2d)$$

$$\Delta u_{\min} \leq \Delta u_k \leq \Delta u_{\max} \quad (2e)$$

$$x_0 = \bar{x} \quad (2f)$$

$$u_{-1} = \bar{u}, \quad (2g)$$

where $\Delta u_k = u_k - u_{k-1}$ and the optimization variables are the elements of the input sequence:

$$U = \{u_0, u_1, \dots, u_{N_p-1}\}. \quad (3)$$

According to the solution strategy, the elements of the state sequence:

$$X = \{x_0, x_1, \dots, x_{N_p}\} \quad (4)$$

may be either included as optimization variables (sparse implementation), or eliminated (dense implementation).

1.1 Notation

The column vector $\mathbf{u} \in \mathbb{R}^{N_p n_u \times 1}$ contains all the stacked entries of the input sequence U :

$$\mathbf{u}^\top = [u_0^\top \ u_1^\top \ \dots \ u_{N_p-1}^\top]^\top. \quad (5)$$

Similarly, the column vector $\mathbf{x} \in \mathbb{R}^{N_p n_x \times 1}$ is defined as:

$$\mathbf{x}^\top = [x_0^\top \ x_1^\top \ \dots \ x_{N_p}^\top]^\top, \quad (6)$$

the column vector $\mathbf{x}_{\text{ref}} \in \mathbb{R}^{N_p n_x \times 1}$ is defined as:

$$\mathbf{x}_{\text{ref}}^\top = [x_{\text{ref}}^\top \ x_{\text{ref}}^\top \ \dots \ x_{\text{ref}}^\top]^\top, \quad (7)$$

and finally the column vector \mathbf{u}_{ref} is defined as:

$$\mathbf{u}_{\text{ref}}^\top = [u_{\text{ref}}^\top \ u_{\text{ref}}^\top \ \dots \ u_{\text{ref}}^\top]^\top. \quad (8)$$

Note that we consider a constant state reference x_{ref} over the prediction horizon for notation simplicity. The extension to a varying reference is straightforward (and actually implemented in pyMPC).

1.2 Receding horizon implementation

In a typical implementation, the MPC input is applied in *receding horizon*. At each time step i , the problem (2) is solved with $x_0 = x[i]$, $u_{-1} = u[i-1]$ and an optimal input sequence u_0, \dots, u_{N_p} is obtained. The first element of this sequence u_0 is the control input that is actually applied at time instant i . At time instant $i+1$, a new state $x[i+1]$ is measured (or estimated), and the process is iterated.

Thus, formally, the MPC control law is a (static) function of the current state and the previous input:

$$u_{MPC} = K(x[i], u[i-1]). \quad (9)$$

Note that this function also depends on the references x_{ref} and u_{ref} and on the system matrices A and B .

2 Quadratic Programming Formulation

The OSQP Quadratic Programming (QP) solver expects a problem with form:

$$\min \frac{1}{2} \mathbf{z}^\top \mathbf{P} \mathbf{z} + \mathbf{q}^\top \mathbf{z} \quad (10a)$$

subject to

$$l \leq \mathbf{A} \mathbf{z} \leq u \quad (10b)$$

To implement the MPC controller using the OSQP solver, we need to re-write the MPC optimization problem (2) in form (10).

3 Sparse Implementation

In the sparse implementation, the dynamic constraints given by the linear dynamics equations (2b) are included explicitly and the state sequence X is considered as an optimization variable, along with the input sequence U .

3.1 Cost function

3.1.1 Terms in Q_x and Q_{x_N}

By direct inspection, the non-constant terms of the cost function in Q_x and Q_{x_N} are:

$$\begin{aligned} J_{Q_x} = \frac{1}{2} \begin{bmatrix} x_0^\top & x_1^\top & \dots & x_{N_p}^\top \end{bmatrix}^\top \overbrace{\text{blkdiag}(Q_x, Q_x, \dots, Q_{x_N})}^{=Q_x} \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{N_p} \end{bmatrix} + \\ + \begin{bmatrix} -x_{\text{ref}}^\top Q_x & -x_{\text{ref}}^\top Q_x & \dots & -x_{\text{ref}}^\top Q_{x_N} \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{N_p} \end{bmatrix}^\top. \quad (11) \end{aligned}$$

More compactly, this is equivalent to:

$$J_{Q_x} = \frac{1}{2} \mathbf{x}^\top Q_x \mathbf{x} - \mathbf{x}_{\text{ref}}^\top Q_x \mathbf{x}. \quad (12)$$

3.1.2 Terms in Q_u

Similarly, for the term J_{Q_u} :

$$J_{Q_u} = \frac{1}{2} \begin{bmatrix} u_0^\top & u_1^\top & \dots & u_{N_p-1}^\top \end{bmatrix} \overbrace{\text{blkdiag}(Q_u, Q_u, \dots, Q_u)}^{Q_u} \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{N_p-1} \end{bmatrix} + \\ + \begin{bmatrix} -u_{\text{ref}}^\top Q_u & -u_{\text{ref}}^\top Q_u & \dots & -u_{\text{ref}}^\top Q_u \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{N_p-1} \end{bmatrix} \quad (13)$$

More compactly, this is equivalent to:

$$J_{Q_u} = \frac{1}{2} \mathbf{u}^\top \mathcal{Q}_u \mathbf{u} - \mathbf{u}_{\text{ref}}^\top \mathcal{Q}_u \mathbf{u}. \quad (14)$$

3.1.3 Terms in $Q_{\Delta u}$

As for the terms in $Q_{\Delta u}$ we have instead:

$$J_{\Delta u} = \frac{1}{2} \begin{bmatrix} u_0 & u_1 & \dots & u_{N_p-1} \end{bmatrix}^\top \begin{bmatrix} 2Q_{\Delta u} & -Q_{\Delta u} & 0 & \dots & \dots & 0 \\ -Q_{\Delta u} & 2Q_{\Delta u} & -Q_{\Delta u} & 0 & \dots & 0 \\ 0 & -Q_{\Delta u} & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & -Q_{\Delta u} & 2Q_{\Delta u} & -Q_{\Delta u} \\ 0 & 0 & 0 & 0 & -Q_{\Delta u} & Q_{\Delta u} \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{N_p-1} \end{bmatrix}^\top \\ - u_{-1}^\top Q_{\Delta u} u_0 \quad (15)$$

More compactly, this is equivalent to:

$$J_{\Delta u} = \frac{1}{2} \mathbf{u}^\top \mathcal{Q}_{\Delta u} \mathbf{u} - u_{-1}^\top Q_{\Delta u} u_0$$

It is convenient to write the above expression using stacked vectors only:

$$J_{\Delta u} = \frac{1}{2} \mathbf{u}^\top \mathcal{Q}_{\Delta u} \mathbf{u} + \begin{bmatrix} -u_{-1}^\top Q_{\Delta u} & 0 & \dots & 0 \end{bmatrix} \mathbf{u} \quad (16)$$

3.2 Constraints

3.2.1 Linear dynamics

Let us consider the linear equality constraints (2b) representing the system dynamics. These can be written in matrix form as:

$$\begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{N_p-1} \\ x_{N_p} \end{bmatrix} = \overbrace{\begin{bmatrix} 0 & 0 & \dots & 0 \\ A_d & 0 & \dots & 0 \\ 0 & A_d & \dots & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & \dots & A_d \end{bmatrix}}^{=\mathcal{A}} \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{N_p-1} \\ x_{N_p} \end{bmatrix} + \overbrace{\begin{bmatrix} 0 & 0 & \dots & 0 \\ B_d & 0 & \dots & 0 \\ 0 & B_d & \dots & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & \dots & B_d \end{bmatrix}}^{=\mathcal{B}} \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{N_p-2} \\ u_{N_p-1} \end{bmatrix} + \overbrace{\begin{bmatrix} \bar{x} \\ 0 \\ \vdots \\ \vdots \\ 0 \end{bmatrix}}^{=\mathcal{C}}. \quad (17)$$

Thus, we get a set of linear equality constraints representing the system dynamics (2b). These constraints can be written as

$$[(\mathcal{A} - I) \quad \mathcal{B}] \begin{bmatrix} \mathbf{x} \\ \mathbf{u} \end{bmatrix} = \mathcal{C}. \quad (18)$$

3.2.2 Variable bounds: x and u

The bounds on x and u are readily implemented as:

$$\begin{bmatrix} x_{\min} \\ u_{\min} \end{bmatrix} \leq \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{u} \end{bmatrix} \leq \begin{bmatrix} x_{\max} \\ u_{\max} \end{bmatrix}. \quad (19)$$

3.2.3 Variable bounds: Δu

$$\begin{bmatrix} u_{-1} + \Delta u_{\min} \\ \Delta u_{\min} \\ \vdots \\ \Delta u_{\min} \end{bmatrix} \leq \begin{bmatrix} I & 0 & \dots & \dots & 0 & 0 \\ -I & I & 0 & \dots & 0 & 0 \\ 0 & -I & I & \dots & 0 & 0 \\ \vdots & & & & & \\ 0 & 0 & \dots & 0 & -I & I \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{N_c-1} \end{bmatrix} \leq \begin{bmatrix} u_{-1} + \Delta u_{\max} \\ \Delta u_{\max} \\ \vdots \\ \Delta u_{\max} \end{bmatrix} \quad (20)$$

3.3 Slack variables

Bounds on x may result in an problem unfeasible! A common solution is to transform the hard constraints in x into soft constraints by means of *slack*

variables ϵ . In the current implementation, there are as many slack variables as state variables, i.e. $\epsilon \in \mathbb{R}^{N_p n_x \times 1}$. We use the constraint:

$$\begin{bmatrix} x_{\min} \\ u_{\min} \end{bmatrix} \leq \begin{bmatrix} I & 0 & I \\ 0 & I & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{u} \\ \epsilon \end{bmatrix} \leq \begin{bmatrix} x_{\max} \\ u_{\max} \end{bmatrix}. \quad (21)$$

In order to penalize constraint violation, we have a penalty term in the cost function:

$$J_\epsilon = \frac{1}{2} \epsilon^\top Q_\epsilon \epsilon, \quad (22)$$

where

$$Q_\epsilon = \text{blkdiag}(Q_\epsilon, Q_\epsilon, \dots, Q_\epsilon) \quad (23)$$

and Q_ϵ is σI_{n_x} , with σ a “large” constant (e.g. $\sigma = 10^4$).

3.4 Control Horizon

Sometimes, we may want to use a control horizon $N_c < N_p$ instead of the standard $N_c = N_p$. In this case, the input considered is constant for $N_c \geq N_p$.

$$\begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{N_p-1} \\ x_{N_p} \end{bmatrix} = \overbrace{\begin{bmatrix} 0 & 0 & \dots & 0 \\ A_d & 0 & \dots & 0 \\ 0 & A_d & \dots & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & \dots & A_d \end{bmatrix}}^{=A} \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{N_p-1} \\ x_{N_p} \end{bmatrix} + \overbrace{\begin{bmatrix} 0 & 0 & \dots & 0 \\ B_d & 0 & \dots & 0 \\ 0 & B_d & \dots & 0 \\ \vdots & 0 & \ddots & 0 \\ 0 & 0 & \dots & B_d \\ 0 & 0 & \dots & \vdots \\ 0 & 0 & \dots & B_d \end{bmatrix}}^{=B} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ \vdots \\ u_{N_c-1} \\ \vdots \\ u_{N_c-1} \end{bmatrix} + \overbrace{\begin{bmatrix} \bar{x} \\ 0 \\ \vdots \\ 0 \end{bmatrix}}^{=C} \quad (24)$$

The contributions J_{Q_u} of the cost function also changes:

$$\begin{aligned}
J_{Q_u} = & \frac{1}{2} \begin{bmatrix} u_0^\top & u_1^\top & \dots & u_{N_p-1}^\top \end{bmatrix} \text{blkdiag} \left(Q_u, Q_u, \dots, (N_p - N_c + 1) Q_u \right) \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{N_p-1} \end{bmatrix} + \\
& + \begin{bmatrix} -u_{\text{ref}}^\top Q_u & -u_{\text{ref}}^\top Q_u & \dots & -(N_p - N_c + 1) u_{\text{ref}}^\top Q_u \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{N_p-1} \end{bmatrix} \quad (25)
\end{aligned}$$

Instead, $J_\Delta u$ does not change (because the input is constant for $k \geq N_c$!

4 Dense Implementation

In the dense implementation, the elements of the input sequence U are the only optimization variables. The state variables are eliminated using the so-called Lagrange equations:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{N_p-1} \\ x_{N_p} \end{bmatrix} = \overbrace{\begin{bmatrix} A_d \\ A_d^2 \\ \vdots \\ A_d^{N_p-1} \\ A_d^{N_p} \end{bmatrix}}^{=\mathcal{A}} x_0 + \overbrace{\begin{bmatrix} B_d & 0 & 0 & \dots & 0 & 0 \\ A_d B_d & B_d & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & 0 & 0 \\ A_d^{N_p-2} B_d & A_d^{N_p-3} B_d & A_d^{N_p-4} B_d & \dots & B_d & 0 \\ A_d^{N_p-1} B_d & A_d^{N_p-2} B_d & A_d^{N_p-3} B_d & \dots & A_d B_d & B_d \end{bmatrix}}^{=\mathcal{B}} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ \vdots \\ u_{N_p-1} \end{bmatrix}. \quad (26)$$

In vector notation, this is simply:

$$\mathbf{x} = \mathcal{A}x_0 + \mathcal{B}\mathbf{u} \quad (27)$$

The elements of the cost function may be written compactly as:

$$J_{Q_x} = \frac{1}{2} (\mathcal{A}x_0 + \mathcal{B}\mathbf{u} - \mathbf{x}_{\text{ref}})^\top \mathcal{Q}_x (\mathcal{A}x_0 + \mathcal{B}\mathbf{u} - \mathbf{x}_{\text{ref}}) \quad (28)$$

$$J_{Q_u} = \frac{1}{2} (\mathbf{u} - \mathbf{u}_{\text{ref}})^\top \mathcal{Q}_u (\mathbf{u} - \mathbf{u}_{\text{ref}}) \quad (29)$$

$$J_{Q_{\Delta u}} = \frac{1}{2} \mathbf{u}^\top \mathcal{Q}_{\Delta u} \mathbf{u} + [-u_{-1}^\top Q_{\Delta u} \quad 0 \quad \dots \quad 0] \mathbf{u} \quad (30)$$

Summing up and expanding terms:

$$\begin{aligned} J = & \frac{1}{2} \mathbf{u}^\top \mathcal{B}^\top \mathcal{Q}_x \mathcal{B} \mathbf{u} + \frac{1}{2} (\mathcal{A}x_0 - \mathbf{x}_{\text{ref}})^\top \mathcal{Q}_x (\mathcal{A}x_0 - \mathbf{x}_{\text{ref}}) + (\mathcal{A}x_0 - \mathbf{x}_{\text{ref}})^\top \mathcal{Q}_x \mathcal{B} \mathbf{u} + \\ & + \frac{1}{2} \mathbf{u}^\top \mathcal{Q}_u \mathbf{u} + \frac{1}{2} \mathbf{u}_{\text{ref}}^\top \mathcal{Q}_u \mathbf{u}_{\text{ref}} + -\mathbf{u}_{\text{ref}}^\top \mathcal{Q}_u \mathbf{u} + \\ & + \frac{1}{2} \mathbf{u}^\top \mathcal{Q}_{\Delta u} \mathbf{u} + [-u_{-1}^\top Q_{\Delta u} \quad 0 \quad \dots \quad 0] \mathbf{u} \quad (31) \end{aligned}$$

Neglecting constant terms and collecting:

$$\begin{aligned} J = C + \frac{1}{2} \mathbf{u}^\top \overbrace{(\mathcal{B}^\top \mathcal{Q}_x \mathcal{B} + \mathcal{Q}_u + \mathcal{Q}_{\Delta u})}^{=\mathbf{P}} \mathbf{u} + \\ \overbrace{\left[(\mathcal{A}x_0 - X_{\text{ref}})^\top \mathcal{Q}_x \mathcal{B} - U_{\text{ref}}^\top \mathcal{Q}_u - [u_{-1}^\top Q_{\Delta u} \quad 0 \quad \dots \quad 0] \right]}^{=\mathbf{q}^\top} \mathbf{u} \quad (32) \end{aligned}$$

Thus, we have

$$\mathbf{q} = \mathcal{B}^\top \mathcal{Q}_x (\mathcal{A}x_0 - X_{\text{ref}}) - \mathcal{Q}_u U_{\text{ref}} + \begin{bmatrix} -Q_{\Delta u} u_{-1} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (33)$$

expanding

$$\mathbf{q} = \overbrace{\mathcal{B}^\top \mathcal{Q}_x \mathcal{A} x_0}^{=p_{x_0}} + \overbrace{-\mathcal{B}^\top \mathcal{Q}_x X_{\text{ref}}}^{p_{X_{\text{ref}}}} + \overbrace{-\mathcal{Q}_u U_{\text{ref}}}^{=p_{U_{\text{ref}}}} + \overbrace{\begin{bmatrix} -Q_{\Delta u} \\ 0 \\ \vdots \\ 0 \end{bmatrix}}^{=p_{u_{-1}}} u_{-1} \quad (34)$$

4.1 Unconstrained case

For an unconstrained problem, the optimal value of \mathbf{u} is:

$$\mathbf{u}^{\text{opt}} = -\mathbf{P}^{-1} \mathbf{q} \quad (35)$$

Expanding, we obtain:

$$\mathbf{u}^{\text{opt}} = k_{x_0}x_0 + k_{\mathbf{x}_{\text{ref}}}\mathbf{x}_{\text{ref}} + k_{\mathbf{u}_{\text{ref}}}\mathbf{u}_{\text{ref}}k_{u_{-1}}u_{-1} \quad (36)$$

with:

$$k_{x_0} = -\mathbf{P}^{-1}p_{x_0} \quad (37\text{a})$$

$$k_{X_{\text{ref}}} = -\mathbf{P}^{-1}p_{X_{\text{ref}}} \quad (37\text{b})$$

$$k_{U_{\text{ref}}} = -\mathbf{P}^{-1}p_{U_{\text{ref}}} \quad (37\text{c})$$

$$k_{u_{-1}} = -\mathbf{P}^{-1}p_{u_{-1}}. \quad (37\text{d})$$

Thus, the control law is a simple linear function of x_0 , X_{ref} , U_{ref} , and u_{-1} .