



MECH 6v29.002 – Model Predictive Control

L19 – Nonlinear MPC (continued)

Outline



- Nonlinear MPC
- Finding the Sublevel Set for Terminal Constraint Set
- Bounding the Linearization Error
- Inverted Pendulum Numerical Example

Nonlinear MPC (cont.)



- For now, assume that there exists a > 0 such that linearization error is small
- Summary:

$$J_{0}^{*}(x_{0}) = \min_{U_{0}} \sum_{k=0}^{N-1} q(x_{k}, u_{k}) + p(x_{N}) \qquad p(x_{N}) = \frac{1}{2} x_{N}^{T} P x_{N}$$

$$s.t. \qquad A_{K}^{T} P A_{K} - P + \mu Q_{K} = 0$$

$$x_{k+1} = f(x_{k}, u_{k}), k \in \{0, 1, ..., N-1\}$$

$$h(x_{k}, u_{k}) \leq 0, k \in \{0, 1, ..., N-1\}$$

$$x_{N} = \mathcal{X}_{f} \qquad \mathcal{X}_{f} = \text{lev}_{a} p(x_{N}) = \{x_{N} \in \mathbb{R}^{n} \mid p(x_{N}) \leq a\}$$

$$x_{0} = x(0) \qquad = \{x_{N} \in \mathbb{R}^{n} \mid \frac{1}{2} x_{N}^{T} P x_{N} \leq a\}$$

 Assuming initial condition is feasible, the origin is exponentially stable under this MPC controller formulation Ellipsoid centered at the origin

Finding the Sublevel Set



• How do we find *a* > 0 to define the terminal constraint such that the following is true?

$$p(f(x,Kx)) - p(x) + \frac{1}{2}x^T Q_K x \le 0 \quad \forall x \in \mathcal{X}_f = \text{lev}_a \ p(x_N) = \{x_N \in \mathbb{R}^n \mid p(x_N) \le a\}$$

- Could take a direct approach by solving the following nonlinear program.
- Choose a > 0 such that

$$\max_{x \in \text{lev}_a \ p(x_N)} p(f(x, Kx)) - p(x) + \frac{1}{2} x^T Q_K x \le 0$$

• Would likely want to maximize terminal set by maximizing *a*

$$\max_{a>0} \left(\max_{x \in \text{lev}_a \ p(x_N)} p(f(x, Kx)) - p(x) + \frac{1}{2} x^T Q_K x \le 0 \right)$$

s.t.
$$\operatorname{lev}_a p(x_N) \subseteq \mathcal{X}$$

$$K \operatorname{lev}_a p(x_N) \subseteq \mathcal{U}$$

Could be hard to solve directly.



• Alternatively, we could try to analytically find a > 0 such that

$$p(f(x,Kx)) - p(x) + \frac{1}{2}x^T Q_K x \le 0 \quad \forall x \in \mathcal{X}_f = \text{lev}_a \ p(x_N) = \{x_N \in \mathbb{R}^n \mid p(x_N) \le a\}$$

• We have already seen that this is achieved if

$$\left(p(x) = \frac{1}{2}x^T P x\right)$$

$$\Delta p \triangleq p(f(x, Kx)) - p(A_K x) \leq \frac{\mu - 1}{2} x^T Q_K x \quad \forall x \in \text{lev}_a \ p(x)$$

- To find when this is true, define the error between the nonlinear and linear models $e(x) = f(x, Kx) A_K x$
- Algebraic manipulations

$$p(f(x,Kx)) = p(e(x) + A_K x) = \frac{1}{2}(e(x) + A_K x)^T P(e(x) + A_K x)$$

$$= \frac{1}{2}e(x)^T Pe(x) + \frac{1}{2}x^T A_K^T P A_K x + x^T A_K^T P e(x)$$

$$p(A_K x) = \frac{1}{2}x^T A_K^T P A_K x$$

$$\Delta p \triangleq p(f(x,Kx)) - p(A_Kx) = \frac{1}{2}e(x)^T Pe(x) + x^T A_K^T Pe(x)$$



• We need to find a > 0 such that

$$\Delta p = \frac{1}{2}e(x)^T Pe(x) + x^T A_K^T Pe(x) \le \frac{\mu - 1}{2} x^T Q_K x \qquad \forall x \in \text{lev}_a \ p(x)$$

- Goal is to write every term as a function of $||x||_2$
- Note that

Then

$$\Delta p = \frac{1}{2} e(x)^{T} P e(x) + x^{T} A_{K}^{T} P e(x)$$

$$\Delta p \leq \frac{1}{2} \lambda_{\max}(P) \|e(x)\|_{2}^{2} + x^{T} A_{K}^{T} P e(x)$$

- Note that $x^T A_K^T Pe(x) \le \|PA_K x\|_2 \|e(x)\|_2 \le \|PA_K\|_2 \|x\|_2 \|e(x)\|_2$
- Therefore

$$\Delta p \le \frac{1}{2} \lambda_{\max}(P) \|e(x)\|_{2}^{2} + \|PA_{K}\|_{2} \|x\|_{2} \|e(x)\|_{2}$$



• We have $\Delta p \leq \frac{1}{2} \lambda_{\max}(P) \|e(x)\|_{2}^{2} + \|PA_{K}\|_{2} \|x\|_{2} \|e(x)\|_{2}$

• Let's assume the error is small (we will prove this later)

$$\|e(x)\|_{2} \le \frac{1}{2}c_{\delta}\|x\|_{2}^{2}$$
 $\forall x \in \delta\mathcal{B} = \{x \in \mathbb{R}^{n} \mid x^{T}x = \|x\|_{2}^{2} \le \delta^{2}\}$

- Then $\Delta p \leq \frac{1}{8} c_{\delta}^2 \lambda_{\max}(P) \|x\|_2^4 + \frac{1}{2} c_{\delta} \|PA_K\|_2 \|x\|_2^3$
- Since we want $\Delta p \leq \frac{\mu 1}{2} x^T Q_K x$
- And we have

$$\Delta p \leq \frac{1}{8} c_{\delta}^{2} \lambda_{\max} (P) \|x\|_{2}^{4} + \frac{1}{2} c_{\delta} \|PA_{K}\|_{2} \|x\|_{2}^{3} \qquad \lambda_{\min} (Q_{K}) \|x\|_{2}^{2} \leq x^{T} Q_{K} x$$

We need

$$\frac{1}{8}c_{\delta}^{2}\lambda_{\max}(P)\|x\|_{2}^{4} + \frac{1}{2}c_{\delta}\|PA_{K}\|_{2}\|x\|_{2}^{3} \leq \frac{\mu - 1}{2}\lambda_{\min}(Q_{K})\|x\|_{2}^{2}$$



• We need

$$\frac{1}{8}c_{\delta}^{2}\lambda_{\max}(P)\|x\|_{2}^{4} + \frac{1}{2}c_{\delta}\|PA_{K}\|_{2}\|x\|_{2}^{3} \leq \frac{\mu - 1}{2}\lambda_{\min}(Q_{K})\|x\|_{2}^{2}$$

• Since $||x||_2 > 0 \quad \forall x \neq 0$

$$\frac{1}{8}c_{\delta}^{2}\lambda_{\max}(P)\|x\|_{2}^{2} + \frac{1}{2}c_{\delta}\|PA_{K}\|_{2}\|x\|_{2} \leq \frac{\mu - 1}{2}\lambda_{\min}(Q_{K})$$

- We currently have the condition that $x \in \delta \mathcal{B} = \left\{ x \in \mathbb{R}^n \mid x^T x = ||x||_2^2 \le \delta^2 \right\}$
- We want to impose the condition

$$x \in \mathcal{X}_f = \text{lev}_a \ p(x) = \left\{ x \in \mathbb{R}^n \mid p(x) \le a \right\} = \left\{ x \in \mathbb{R}^n \mid \frac{1}{2} x^T P x \le a \right\}$$

• Since $\frac{1}{2} \lambda_{\min}(P) ||x||_2^2 \le \frac{1}{2} x^T P x$

$$x \in \mathcal{X}_f \Rightarrow \frac{1}{2} \lambda_{\min}(P) \|x\|_2^2 \le a \Rightarrow \|x\|_2 \le \sqrt{\frac{2a}{\lambda_{\min}(P)}}$$



• Now we can finally choose a > 0 such that x being in the terminal constraint set results in

$$\|x\|_2 \le \sqrt{\frac{2a}{\lambda_{\min}(P)}} \le \delta$$

Since $x \in \delta \mathcal{B} = \left\{ x \in \mathbb{R}^n \mid x^T x = ||x||_2^2 \le \delta^2 \right\}$ is required for the linearization error to be small

• and
$$\frac{1}{8}c_{\delta}^{2}\lambda_{\max}(P)\|x\|_{2}^{2} + \frac{1}{2}c_{\delta}\|PA_{K}\|_{2}\|x\|_{2} \leq \frac{\mu-1}{2}\lambda_{\min}(Q_{K})$$

which is satisfied if the following condition is satisfied

$$\frac{1}{8}c_{\delta}^{2}\lambda_{\max}(P)\frac{2a}{\lambda_{\min}(P)} + \frac{1}{2}c_{\delta}\|PA_{K}\|_{2}\sqrt{\frac{2a}{\lambda_{\min}(P)}} \leq \frac{\mu - 1}{2}\lambda_{\min}(Q_{K})$$

Bounding the Linearization Error



Now we can derive/verify our assumption that the error was small

$$\|e(x)\|_{2} \le \frac{1}{2}c_{\delta}\|x\|_{2}^{2}$$
 $\forall x \in \delta\mathcal{B} = \{x \in \mathbb{R}^{n} \mid x^{T}x = \|x\|_{2}^{2} \le \delta^{2}\}$

- Linearization error was defined as $e(x) = f(x, Kx) A_K x$
- By definition, no linearization error at the origin $e(0) = f(0,0) A_K 0 = 0$
- Derivative of linearization error is

$$\frac{d}{dx}e(x) = \frac{\partial}{\partial x}f(x,Kx) + \frac{\partial}{\partial u}f(x,Kx)\frac{du}{dx} - A_K \qquad u = Kx$$

$$\frac{d}{dx}e(x) = \frac{\partial}{\partial x}f(x,Kx) + \frac{\partial}{\partial u}f(x,Kx)K - A_K \qquad \frac{du}{dx} = K$$

Derivative of linearization error is zero at the origin

$$\frac{d}{dx}e(0) = A + BK - A_K = 0 A = \frac{\partial f}{\partial x}\Big|_{x,u=0} B = \frac{\partial f}{\partial u}\Big|_{x,u=0}$$

Bounding the Linearization Error (cont.)



- Will use result from Mean Value Theorem to bound linearization error (Proposition A.11 of [1])
- Mean Value Theorem for Vector Functions If $f: \mathbb{R}^n \to \mathbb{R}^m$ has continuous partial derivatives at each point of $x \in \mathbb{R}^n$, then for any $x, y \in \mathbb{R}^n$,

$$f(y) = f(x) + \frac{\partial}{\partial x} f(x)(y-x) + \int_{0}^{1} (1-s)(y-x)^{T} \frac{\partial^{2}}{\partial x^{2}} f(x+s(y-x))(y-x) ds$$

- Think about f(y) = e(x) f(x) = e(0)
- Since

$$e(0) = 0 \qquad \frac{d}{dx}e(0) = 0$$

$$e(x) = 0 + 0 + \int_{0}^{1} (1 - s) x^{T} \frac{\partial^{2}}{\partial x^{2}} e(sx) x ds$$

Watch out for the change in variable *x*

Bounding the Linearization Error (cont.)



- Now we can use the fact that *f* is twice continuously differentiable to bound the linearization error
- Since f is twice continuously differentiable for any $\delta > 0$, there exists $c_{\delta} > 0$ such that

$$\left\| \frac{\partial^2}{\partial x^2} e(x) \right\|_2 \le c_{\delta}, \quad \forall x \in \delta \mathcal{B} = \left\{ x \in \mathbb{R}^n \mid x^T x = \|x\|_2^2 \le \delta^2 \right\}$$

• Therefore,

$$e(x) = \int_{0}^{1} (1-s)x^{T} \frac{\partial^{2}}{\partial x^{2}} e(sx)xds$$

$$\|e(x)\|_{2} \leq \left\|\int_{0}^{1} (1-s)x^{T} c_{\delta}xds\right\|_{2}$$
Note that s is between 0 and 1
$$\|e(x)\|_{2} \leq c_{\delta} \int_{0}^{1} (1-s)ds \|x^{T}x\|_{2}$$
This is exactly what we

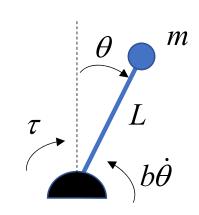
This is exactly what we used when we assumed the error was small

$$\|e(x)\|_2 \le \frac{c_{\delta}}{2} \|x\|_2^2 \quad \forall x \in \delta \mathcal{B}$$

Inverted Pendulum Example



- Consider the inverted pendulum
- Dynamic model: $I\ddot{\theta} = \sum M$ $mL^2\ddot{\theta} = \tau b\dot{\theta} + mgL\sin\theta$ $\ddot{\theta} = \frac{1}{mI^2}\tau \frac{b}{mI^2}\dot{\theta} + \frac{g}{I}\sin\theta$



• State-space model:

$$x_1 = \theta$$

$$\dot{x}_1 = x_2$$

$$x_2 = \dot{\theta}$$

$$\dot{x}_2 = \frac{g}{L}\sin x_1 - \frac{b}{mL^2}x_2 + \frac{1}{mL^2}u$$

$$u = \tau$$

• Forward-Euler discretization $\dot{x} \approx \frac{x^+ - x}{\Delta t}$

$$x_1^+ = x_1 + \Delta t x_2$$

$$x_2^+ = \Delta t \frac{g}{L} \sin x_1 + \left(1 - \frac{\Delta t b}{mL^2}\right) x_2 + \frac{\Delta t}{mL^2} u$$

$$x_{k+1} = f(x_k, u_k)$$



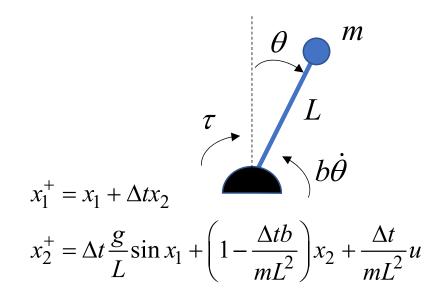
• Linearization about the origin:

$$\frac{\partial f}{\partial x} = \begin{bmatrix} 1 & \Delta t \\ \Delta t \frac{g}{L} \cos x_1 & 1 - \frac{\Delta t b}{mL^2} \end{bmatrix} \qquad \frac{\partial f}{\partial u} = \begin{bmatrix} 0 \\ \Delta t \\ mL^2 \end{bmatrix}$$

$$\frac{\partial f}{\partial u} = \begin{bmatrix} 0 \\ \Delta t \\ mL^2 \end{bmatrix}$$

$$A = \frac{\partial f}{\partial x}\Big|_{x,u=0} = \begin{bmatrix} 1 & \Delta t \\ \Delta t \frac{g}{L} & 1 - \frac{\Delta tb}{mL^2} \end{bmatrix}$$

$$B = \frac{\partial f}{\partial u}\Big|_{x,u=0} = \begin{bmatrix} 0 \\ \Delta t \\ mL^2 \end{bmatrix}$$





• Parameters:

```
%% System Definition
% System dimensions
nx = 2; % Number of states
nu = 1; % Number of inputs (controllable)
% Model parameters
g = 9.81; % Gravity [m/s^2]
L = 1; % Length of pendulum
m = 2; % Mass of pendulum
b = 3; % Friction coeefficient
dt = 0.1; % Discrete time step size
thetaMaxSS = pi/2; % Maximum angle that torque can hold at steady-state [rad]
tauMax = abs(m*g*L*sin(thetaMaxSS));
% Nonlinear discrete time model
f = (x,u) [x(1) + dt*x(2); (1-dt*b/(m*L^2))*x(2) + dt*g/L*sin(x(1)) + dt/(m*L^2)*u(1)];
% Linear discrete time model
A = [1 dt; dt*g/L 1-dt*b/(m*L^2)];
B = [0; dt/(m*L^2)];
f lin = @(x,u) [A*x + B*u];
```

Constraints

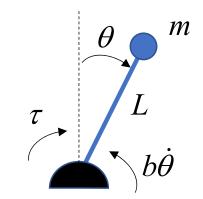
$$-5 \ rad \le x_1 \le 5 \ rad$$
$$-10 \ rad \le x_2 \le 10 \ rad \ / \ s$$
$$-\tau_{\text{max}} \le u \le \tau_{\text{max}}$$



• Control design:

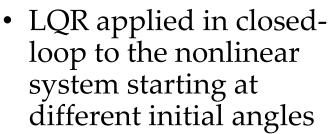
• Stage cost:
$$q(x_k, u_k) = \frac{1}{2} \left(x_k^T Q x_k + u_k^T R u_k \right)$$

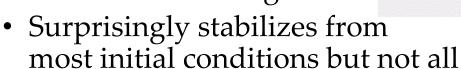
 $Q = I_2, R = 1$

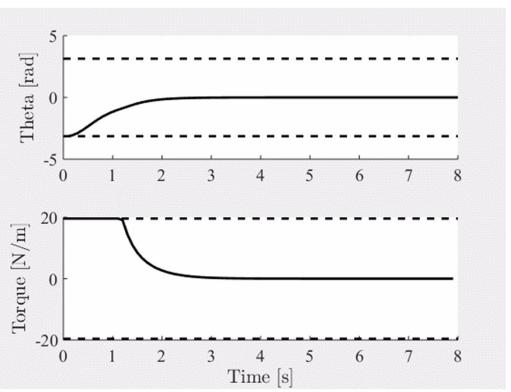


• LQR Design for candidate static feedback control law $u_k = Kx_k$

```
%% LQR design
Q = eye(nx);
R = eye(nu);
[K,~,~] = dlqr(A,B,Q,R);
K = -K; % For positive u = Kx
% Closed-loop linear dynamics
AK = A + B * K;
% Closed-loop stage cost
QK = A + K'*R*K;
```









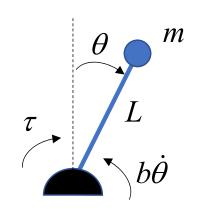
- Control design:
 - Lyapunov equation for terminal cost

$$A_K^T P A_K - P + \mu Q_K = 0 \qquad \longrightarrow \qquad p(x_N) = \frac{1}{2} x_N^T P x_N$$

```
%% Lyapunov Equation

mu = 10; % Must be greater than 1

P = dlyap(AK',mu*Q); % Remember A' since Matlab uses A*X*A' - X + Q = 0
```



Finding a sublevel set for terminal constraint

$$x_N \in \mathcal{X}_f = \text{lev}_a \ p(x_N) = \left\{ x_N \in \mathbb{R}^n \mid \frac{1}{2} x^T P x \le a \right\}$$

Analyze linearization error

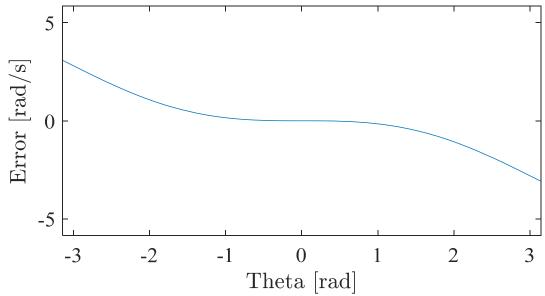
$$e(x) = f(x, Kx) - A_K x$$

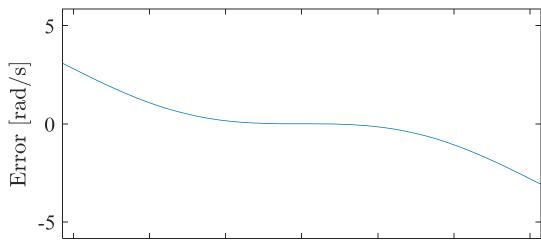
• By comparing nonlinear and linear model, we should get

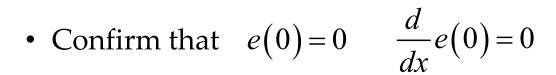
$$e(x) = \begin{bmatrix} 0 \\ \Delta t \frac{g}{L} \sin x_1 - \Delta t \frac{g}{L} x_1 \end{bmatrix}$$



- Control design:
 - Plot the linearization error $e(x) = \begin{bmatrix} 0 \\ \Delta t \frac{g}{L} \sin x_1 \Delta t \frac{g}{L} x_1 \end{bmatrix}$





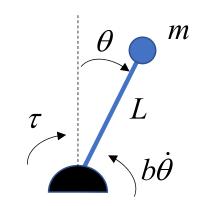




- Control design:
 - Plot the second derivative of the linearization error

$$e(x) = \begin{bmatrix} 0 \\ \Delta t \frac{g}{L} \sin x_1 - \Delta t \frac{g}{L} x_1 \end{bmatrix} \qquad \frac{d^2}{dx^2} e(x) = \begin{bmatrix} 0 \\ -\Delta t \frac{g}{L} \sin x_1 \end{bmatrix}$$

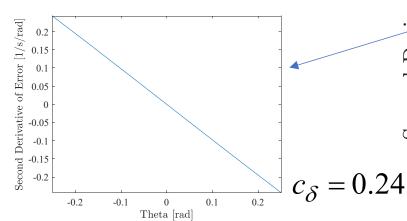
$$\frac{d^2}{dx^2}e(x) = \begin{vmatrix} 0 \\ -\Delta t \frac{g}{L}\sin x_1 \end{vmatrix}$$

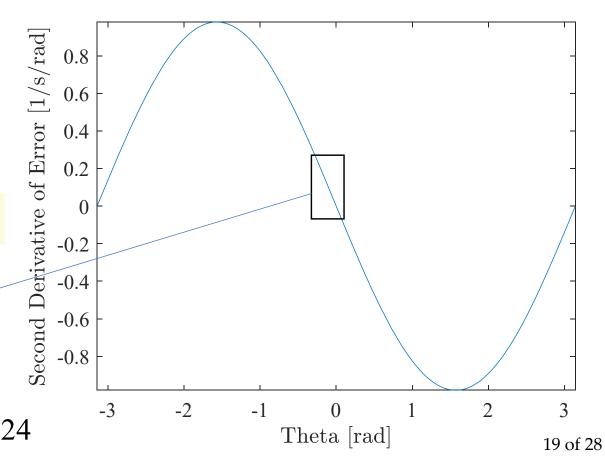


Choose

 $\delta > 0$ and $c_{\delta} > 0$ such that

$$\left\| \frac{\partial^2}{\partial x^2} e(x) \right\|_2 \le c_{\delta}, \quad \forall x \in \delta \mathcal{B}$$







20 of 28

• Control design:

a = fzero(f aRoot,le0);

= sqrt(2*a/lambdaMinP);

• Find a > 0 that satisfies

$$\frac{1}{8}c_{\delta}^{2}\lambda_{\max}(P)\frac{2a}{\lambda_{\min}(P)} + \frac{1}{2}c_{\delta}\|PA_{K}\|_{2}\sqrt{\frac{2a}{\lambda_{\min}(P)}} \leq \frac{\mu-1}{2}\lambda_{\min}(Q_{K})$$

and $\sqrt{\frac{2a}{\lambda_{\min}(P)}} \le \delta$

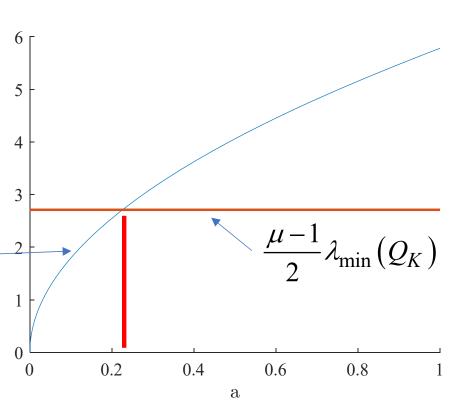
```
lambdaMinP = min(eig(P));
lambdaMaxP = max(eig(P));
lambdaMinQK = min(eig(QK));
normPAK = norm(P*AK);

cl = 1/4*c_delta^2*lambdaMaxP/lambdaMinP;
c2 = sqrt(2)/2*c_delta*normPAK/sqrt(lambdaMinP);

f_a = @(a) cl*a(l) + c2*sqrt(a(l));

f_aRoot = @(a) f_a(a) - (mu-l)/2*lambdaMinQK;
```

$$0.1838 = r = \sqrt{\frac{2a}{\lambda_{\min}(P)}} \le \delta = 0.25$$





• Control design:

% Prediction horizon

constraints = [];
objective = 0;
for k = 1:N

% Optimization variables

u = sdpvar(repmat(nu,1,N),repmat(1,1,N));

x = sdpvar(repmat(nx, 1, N+1), repmat(1, 1, N+1));

N = 80;

end

```
J_{0}^{*}(x_{0}) = \min_{U_{0}} \sum_{k=0}^{N-1} q(x_{k}, u_{k}) + p(x_{N}) \qquad p(x_{N}) = \frac{1}{2} x_{N}^{T} P x_{N}
s.t.
x_{k+1} = f(x_{k}, u_{k}), k \in \{0, 1, ..., N-1\}
h(x_{k}, u_{k}) \leq 0, k \in \{0, 1, ..., N-1\}
x_{N} = \mathcal{X}_{f} \qquad \mathcal{X}_{f} = \text{lev}_{a} p(x_{N}) = = \left\{x_{N} \in \mathbb{R}^{n} \mid \frac{1}{2} x_{N}^{T} P x_{N} \leq a\right\}
x_{0} = x(0)
```

objective = objective + $1/2*x_{k}'*Q*x_{k} + 1/2*u_{k}'*R*u_{k};$

constraints = [constraints, x_{k+1} == f(x_{k},u_{k})];
constraints = [constraints, -tauMax <= u {k} <= tauMax];</pre>

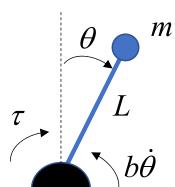
constraints = [constraints, $-5 \le x_{k}(1) \le 5$]; constraints = [constraints, $-10 \le x_{k}(2) \le 10$];

constraints = [constraints, $x_{N+1}'*P*x_{N+1} \le 2*a$];

objective = objective + 1/2*x {N+1}'*P*x {N+1};



- Nonlinear MPC implementation:
 - Used IPOPT (could use fmincon)
 - Nonlinear programs are much harder to solve
 - If you type opts.ipopt, you will see 100+ settings you can configure (most you should not change)

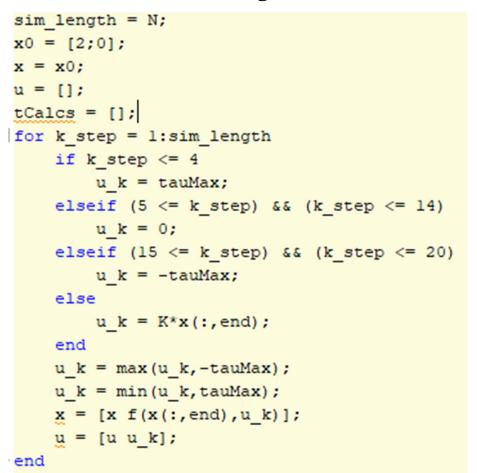


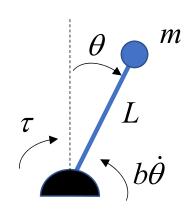
- No guarantee that you will find the globally optimal solution
 - No guarantee that you will find a feasible local solution
- Warm-starting is extremely valuable
 - Give a good initial guess of the optimal solution to start the solver close to a local minimum
 - Change your initial guess, could get completely different solution (different local minimum)

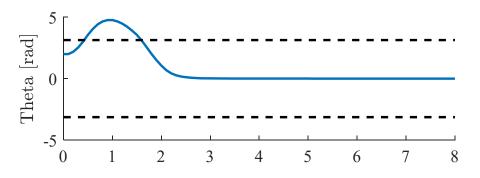
```
opts = sdpsettings('solver','ipopt','ipopt.max_iter',10000,'usex0',1,'debug',0,'verbose',2);
controller = optimizer(constraints,objective,opts,{x_{1}},[u_,x_]);
```

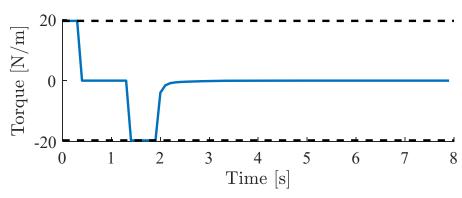


- Starting from an initial condition in the range where LQR did not work $x_0 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$
- Warm-start
 - Designed a simple control strategy to use as an initial guess of a solution





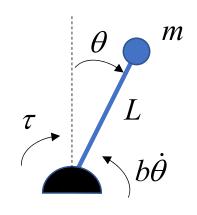


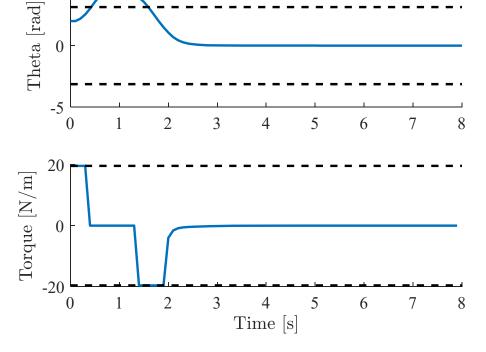




- Starting from an initial condition in the range where LQR did not work $x_0 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$
- Warm-start
 - Assign this solution to the yalmip variables
 - And check that this solution satisfies the constraints of the MPC formulation

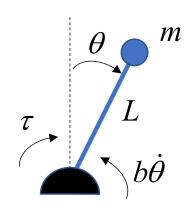
```
for k = 1:N
    assign(x_{k},x(:,k));
    assign(u_{k},u(k));
end
assign(x_{k+1},x(:,k+1));
check(constraints)
```

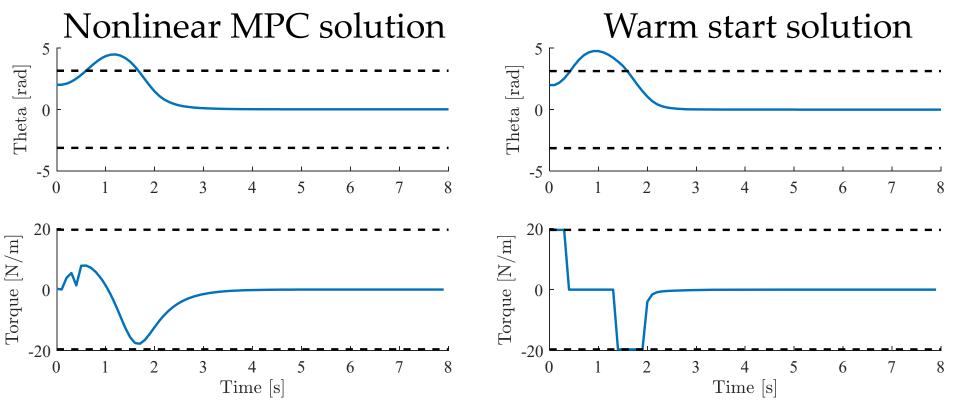






- Starting from an initial condition in the range where LQR did not work $x_0 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$
- MPC Solution

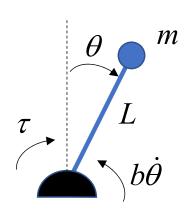


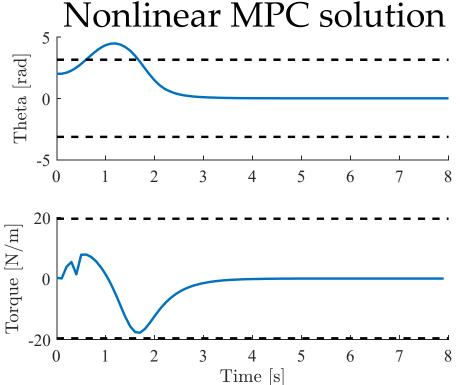


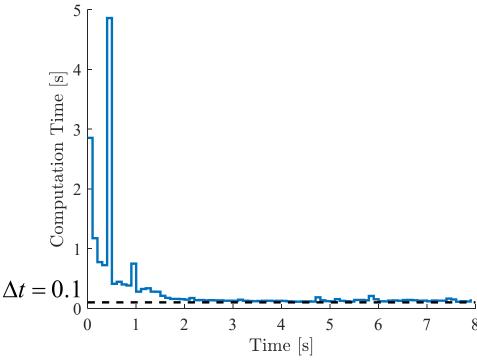
Significantly less use of the actuator



- Starting from an initial condition in the range where LQR did not work $x_0 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$
- MPC Solution



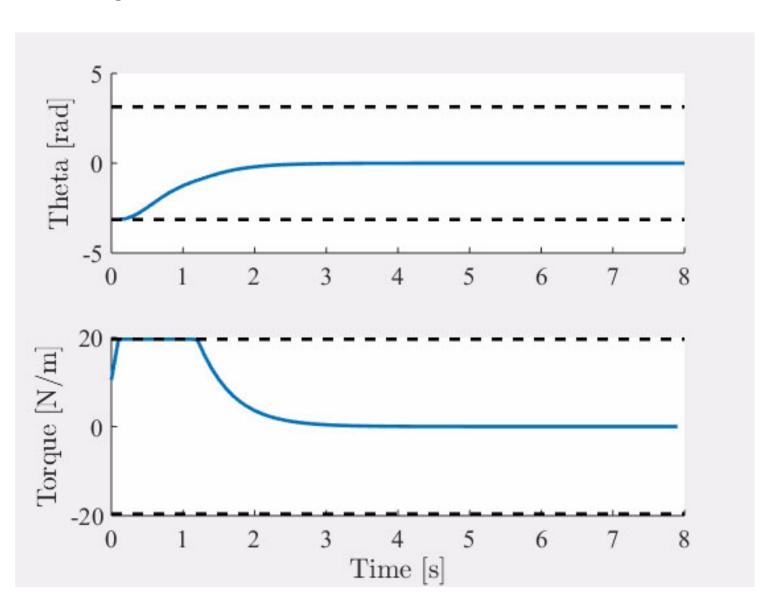


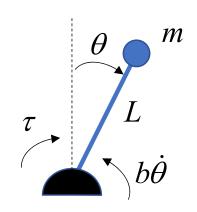


Could not be implemented in real-time!



• Range of initial conditions







• Range of initial conditions

