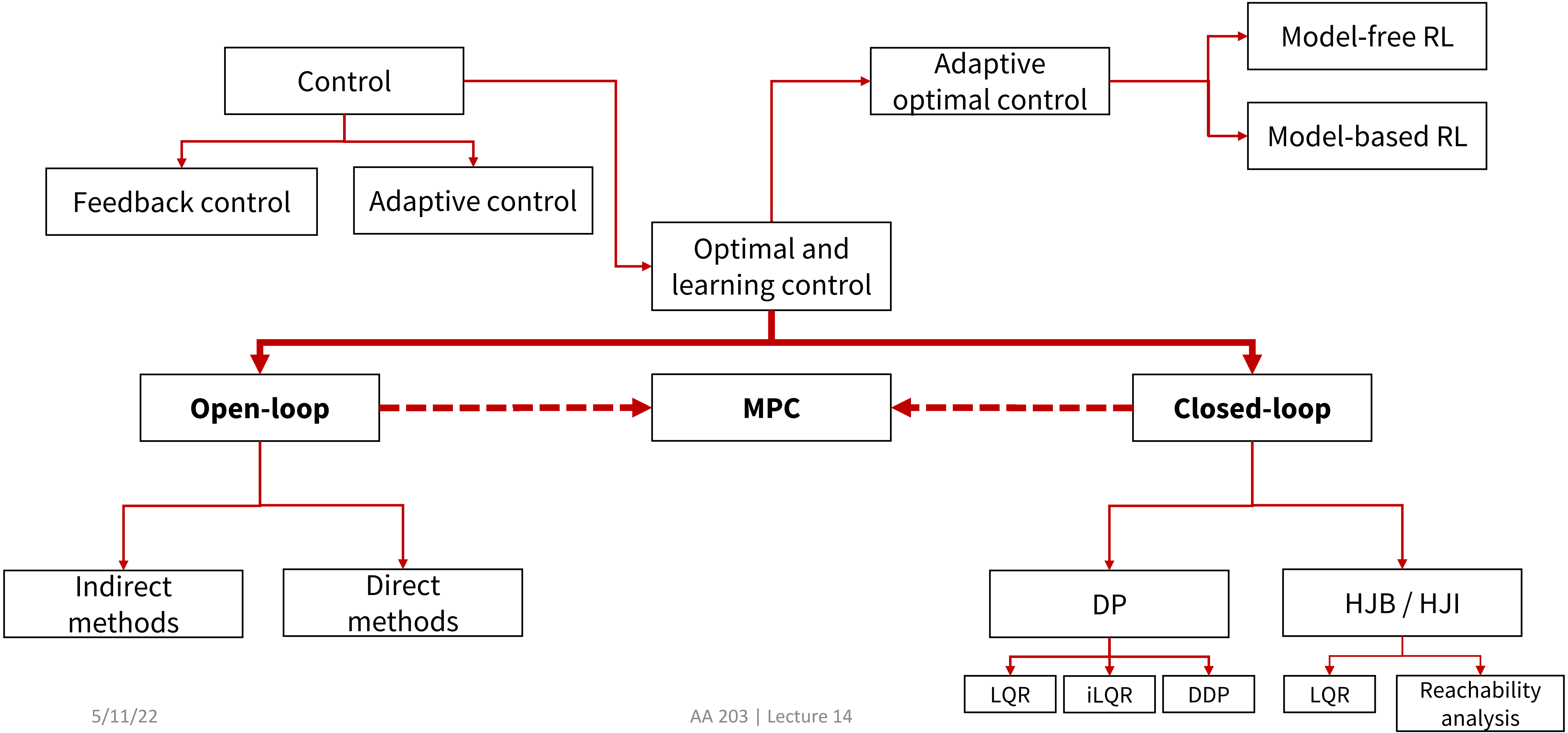
AA203
Optimal and Learning-based Control

Stability of MPC, implementation aspects

Stanford University

Autonomous Systems Lab

Roadmap


Control
Feedback control
Adaptive control
Optimal and learning control
Adaptive optimal control
Model-free RL
Model-based RL
Open-loop
MPC
Closed-loop
Indirect methods
Direct methods
DP
HJB / HJI
LQR
iLQR
DDP
LQR
Reachability analysis

Model predictive control

- Stability of MPC
- Implementation aspects of MPC
- Robust MPC

- Reading:
  - F. Borrelli, A. Bemporad, M. Morari. *Predictive Control for Linear and Hybrid Systems*, 2017.
  - J. B. Rawlings, D. Q. Mayne, M. M. Diehl. *Model Predictive Control: Theory, Computation, and Design*, 2017.

Stability of MPC

- Persistent feasibility does not guarantee that the closed-loop trajectories converge towards the desired equilibrium point

- One of the most popular approaches to guarantee persistent feasibility and stability of the MPC law makes use of a control invariant terminal set X_f for feasibility, and of a terminal function $p(\cdot)$ for stability

- To prove stability, we leverage the tool of Lyapunov stability theory

Lyapunov stability theory

- Lyapunov theorem: Consider the equilibrium point $\mathbf{x} = 0$ for the autonomous system $\{\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k)\}$ (with $\mathbf{f}(\mathbf{0}) = \mathbf{0}$). Let $\Omega \subset \mathbb{R}^n$ be a closed, bounded, positively invariant set containing the origin. Let $V: \mathbb{R}^n \to \mathbb{R}$ be a function, continuous at the origin, such that

$$V(\mathbf{0}) = 0 \text{ and } V(\mathbf{x}) > 0 \quad \forall \mathbf{x} \in \Omega \setminus \{\mathbf{0}\}$$

$$V(\mathbf{x}_{k+1}) - V(\mathbf{x}_k) < 0 \quad \forall \mathbf{x}_k \in \Omega \setminus \{\mathbf{0}\}$$

  Then $\mathbf{x} = 0$ is asymptotically stable in Ω.

- The idea is to show that with appropriate choices of X_f and $p(\cdot)$, J_0^* is a Lyapunov function for the closed-loop system

MPC stability theorem

- MPC stability theorem (for quadratic cost): Assume

A0: $Q = Q^T \succ 0, R = R^T \succ 0, P \succ 0$

A1: Sets X, X_f, and U contain the origin in their interior and are closed

A2: $X_f \subseteq X$ is control invariant and bounded

A3: $\min_{\mathbf{u}\in U,\, A\mathbf{x}+B\mathbf{u} \in X_f} \big(-p(\mathbf{x}) + c(\mathbf{x},\mathbf{u}) + p(A\mathbf{x} + B\mathbf{u})\big) \leq 0, \forall \mathbf{x} \in X_f$

Then, the origin of the closed-loop system is asymptotically stable with domain of attraction X_0.

MPC stability theorem

- Proof:

1. Note that, by assumption A2, persistent feasibility is guaranteed for *any* P, Q, R
2. We want to show that J_0^* is a Lyapunov function for the closed-loop system $\mathbf{x}(t+1) = \mathbf{f}_{\text{cl}}(\mathbf{x}(t))$, with respect to the equilibrium $\mathbf{f}_{\text{cl}}(\mathbf{0}) = \mathbf{0}$ (the origin is indeed an equilibrium as $\mathbf{0} \in X, \mathbf{0} \in U$, and the cost is positive for any non-zero control sequence)
3. X_0 is bounded and closed (follows from assumption on X_f)
4. $J_0^*(\mathbf{0}) = 0$ (value is nonnegative by construction, and 0 is achievable)

MPC stability theorem

- Proof:

5. $J_0^*(\mathbf{x}) > 0$ for all $\mathbf{x} \in X_0 \setminus \{\mathbf{0}\}$
6. Next we show the decay property. Since the setup is time-invariant, we can study the decay property between $t = 0$ and $t = 1$
   - Let $\mathbf{x}(0) \in X_0$, let $U_0^{[0]} = [\mathbf{u}_0^{[0]}, \mathbf{u}_1^{[0]}, \dots, \mathbf{u}_{N-1}^{[0]}]$ be the optimal control sequence, and let $[\mathbf{x}(0), \mathbf{x}_1^{[0]}, \dots, \mathbf{x}_N^{[0]}]$ be the corresponding trajectory
   - After applying $\mathbf{u}_0^{[0]}$, one obtains $\mathbf{x}(1) = A\mathbf{x}(0) + B\mathbf{u}_0^{[0]}$
   - Consider the sequence of controls $[\mathbf{u}_1^{[0]}, \mathbf{u}_2^{[0]}, \dots, \mathbf{u}_{N-1}^{[0]}, \mathbf{v}]$, where $\mathbf{v} \in U$, and the corresponding state trajectory is $[\mathbf{x}(1), \mathbf{x}_2^{[0]}, \dots, \mathbf{x}_N^{[0]}, A\mathbf{x}_N^{[0]} + B\mathbf{v}]$

MPC stability theorem

- Since $\mathbf{x}_N^{[0]} \in X_f$ (by terminal constraint), and since X_f is control invariant,
$$\exists \bar{\mathbf{v}} \in U \text{ such that } A\mathbf{x}_N^{[0]} + B\bar{\mathbf{v}} \in X_f$$
- With such a choice of $\bar{\mathbf{v}}$, the sequence $[\mathbf{u}_1^{[0]}, \mathbf{u}_2^{[0]}, \ldots, \mathbf{u}_{N-1}^{[0]}, \bar{\mathbf{v}}]$ is feasible for the MPC optimization problem at time $t = 1$
- Since this sequence is not necessarily optimal

$$J_0^*\big(\mathbf{x}(1)\big) \le p\left(A\mathbf{x}_N^{[0]} + B\bar{\mathbf{v}}\right) + \sum_{k=1}^{N-1} c\left(\mathbf{x}_k^{[0]}, \mathbf{u}_k^{[0]}\right) + c\left(\mathbf{x}_N^{[0]}, \bar{\mathbf{v}}\right)$$

MPC stability theorem

- Since $\mathbf{x}_N^{[0]} \in X_f$ (by terminal constraint), and since X_f is control invariant,

$$\exists \bar{\mathbf{v}} \in U \text{ such that } A\mathbf{x}_N^{[0]} + B\bar{\mathbf{v}} \in X_f$$

- With such a choice of $\bar{\mathbf{v}}$, the sequence $[\mathbf{u}_1^{[0]}, \mathbf{u}_2^{[0]}, \ldots, \mathbf{u}_{N-1}^{[0]}, \bar{\mathbf{v}}]$ is feasible for the MPC optimization problem at time $t = 1$
- Since this sequence is not necessarily optimal

$$J_0^*(\mathbf{x}(1)) \leq p\left(A\mathbf{x}_N^{[0]} + B\bar{\mathbf{v}}\right) + \sum_{k=1}^{N-1} c\left(\mathbf{x}_k^{[0]}, \mathbf{u}_k^{[0]}\right) + c\left(\mathbf{x}_N^{[0]}, \bar{\mathbf{v}}\right)$$
$$+ p\left(\mathbf{x}_N^{[0]}\right) - p\left(\mathbf{x}_N^{[0]}\right) + c\left(\mathbf{x}(0), \mathbf{u}_0^{[0]}\right) - c\left(\mathbf{x}(0), \mathbf{u}_0^{[0]}\right)$$

MPC stability theorem

- Equivalently

$$J_0^*(\mathbf{x}(1)) \leq p\left(A\mathbf{x}_N^{[0]} + B\bar{\mathbf{v}}\right) + J_0^*(\mathbf{x}(0)) - p\left(\mathbf{x}_N^{[0]}\right) - c\left(\mathbf{x}(0), \mathbf{u}_0^{[0]}\right) + c(\mathbf{x}_N^{[0]}, \bar{\mathbf{v}})$$

- Since $\mathbf{x}_N^{[0]} \in X_f$, by assumption A3, we can select $\bar{\mathbf{v}}$ such that

$$J_0^*(\mathbf{x}(1)) \leq J_0^*(\mathbf{x}(0)) - c\left(\mathbf{x}(0), \mathbf{u}_0^{[0]}\right)$$

- Since $c\left(\mathbf{x}(0), \mathbf{u}_0^{[0]}\right) > 0$ for all $\mathbf{x}(0) \in X_0 \setminus \{0\}$,

$$J_0^*(\mathbf{x}(1)) - J_0^*(\mathbf{x}(0)) < 0$$

- The last step is to prove continuity; details are omitted and can be found in Borrelli, Bemporad, Morari, 2017
- Note: A2 is used to guarantee persistent feasibility; this assumption can be replaced with an assumption on the horizon N

How to choose X_f and P?

- Case 1: assume A is asymptotically stable
  - Set X_f as the maximally positive invariant set O_∞ for system $\mathbf{x}(t+1) = A\mathbf{x}(t), \quad \mathbf{x}(t) \in X$
  - X_f is a control invariant set for system $\mathbf{x}(t+1) = A\mathbf{x}(t) + B\mathbf{u}(t)$, as $\mathbf{u} = 0$ is a feasible control
  - As for stability, $\mathbf{u} = 0$ is feasible and $A\mathbf{x} \in X_f$ if $\mathbf{x} \in X_f$, thus assumption A3 becomes

$$-\mathbf{x}^T P\mathbf{x} + \mathbf{x}^T Q\mathbf{x} + \mathbf{x}^T A^T PA\mathbf{x} \leq 0, \text{ for all } \mathbf{x} \in X_f,$$

which is true since, due to the fact that A is asymptotically stable,

$$\exists P \succ 0 \mid -P + Q + A^T PA = 0 \quad \text{(Lyapunov Equation)}$$

Cost-to-go/value function

How to choose X_f and P?

- Case 2: general case (e.g., if A is open-loop unstable)
  - Let F_∞ be the optimal gain for the infinite-horizon LQR controller
  - Set X_f as the maximal positive invariant set for system

$$\mathbf{x}(t+1) = (A + BF_\infty)\mathbf{x}(t)$$

  (with constraints $\mathbf{x}(t) \in X$, and $F_\infty \mathbf{x}(t) \in U$)
  - Set P as the solution P_∞ to the discrete-time Riccati equation, i.e., the value function via LQR

$$-P + Q + A^T PA - (A^T PB)(R + B^T PB)^{-1}(B^T PA) = 0$$

How to choose X_f and P?

- Case 2: general case (e.g., if A is open-loop unstable)
  - Let F_∞ be the optimal gain for the infinite-horizon LQR controller
  - Set X_f as the maximal positive invariant set for system

$$\mathbf{x}(t+1) = (A + BF_\infty)\mathbf{x}(t)$$

  (with constraints $\mathbf{x}(t) \in X$, and $F_\infty \mathbf{x}(t) \in U$)
  - Set P as the solution P_∞ to the discrete-time Riccati equation, i.e., the value function via LQR

$$-P + Q + A^T PA - (A^T PB)(R + B^T PB)^{-1}(B^T PA) = 0$$

- Note: both cases as presented are just (suboptimal) choices!

Explicit MPC

- In some cases, the MPC law can be *pre-computed* → no need for online optimization
- Important case: constrained LQR

$$J_0^*(\mathbf{x}) = \min_{\mathbf{u}_0,\ldots,\mathbf{u}_{N-1}} \mathbf{x}_N^T P \mathbf{x}_N + \sum_{k=0}^{N-1} \mathbf{x}_k^T Q \mathbf{x}_k + \mathbf{u}_k^T R \mathbf{u}_k$$

subject to $\mathbf{x}_{k+1} = A\mathbf{x}_k + B\mathbf{u}_k, \quad k = 0, \ldots, N-1$

$\mathbf{x}_k \in X, \quad \mathbf{u}_k \in U, \quad k = 0, \ldots, N-1$

$\mathbf{x}_N \in X_f$

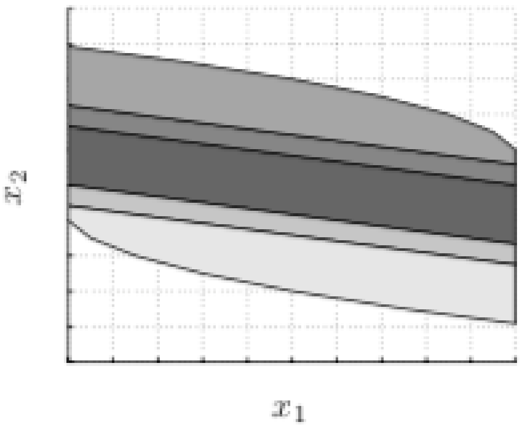
$\mathbf{x}_0 = \mathbf{x}$

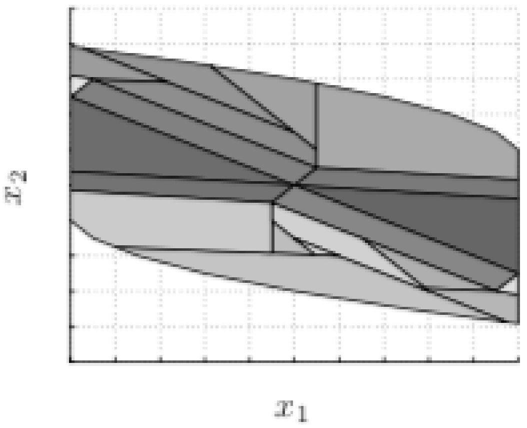
Explicit MPC

- The solution to the constrained LQR problem is a control which is a continuous piecewise affine function on polyhedral partition of the state space X, that is $\mathbf{u}_k^* = \pi_k(\mathbf{x}_k)$ where

$$\pi_k(\mathbf{x}) = F_k^j \mathbf{x} + g_k^j \;\text{ if }\; H_k^j \mathbf{x} \le K_k^j, \;\; j = 1, \dots, N_k^r$$

- Thus, online, one has to locate in which cell of the polyhedral partition the state $\mathbf{x}$ lies, and then one obtains the optimal control via a look-up table query





Tuning and practical use

- At present there is no other technique than MPC to design controllers for general large linear multivariable systems with input and output constraints with a stability guarantee
- Design approach (for squared 2-norm cost):
  - Choose horizon length N and the control invariant target set X_f
  - Control invariant target set X_f should be as large as possible for performance
  - Choose the parameters Q and R freely to affect the control performance
  - Adjust P as per the stability theorem
  - Useful toolbox (MATLAB): https://www.mpt3.org/
- In practice, sometimes choosing a good terminal cost is enough (i.e., don't need to enforce a terminal control invariant condition), though you may be sacrificing guarantees

MPC for reference tracking

- Usual cost

$$\sum_{k=0}^{N-1} \mathbf{x}_k^T Q \mathbf{x}_k + \mathbf{u}_k^T R \mathbf{u}_k$$

does not work, as in steady state control does not need to be zero

- $\delta\mathbf{u}$- formulation: reason in terms of *control changes*

$$\mathbf{u}_k = \mathbf{u}_{k-1} + \delta\mathbf{u}_k$$

MPC for reference tracking

- The MPC problem is readily modified to

$$J_0^*(\mathbf{x}(t)) = \min_{\delta\mathbf{u}_0,\dots,\delta\mathbf{u}_{N-1}} \sum_k \|\mathbf{y}_k - \mathbf{r}_k\|_Q^2 + \|\delta\mathbf{u}_k\|_R^2$$

$$\begin{aligned}
\text{subject to} \quad & \mathbf{x}_{k+1} = A\mathbf{x}_k + B\mathbf{u}_k, && k = 0,\dots,N-1 \\
& \mathbf{y}_k = C\mathbf{x}_k, && k = 0,\dots,N-1 \\
& \mathbf{x}_k \in X, \quad \mathbf{u}_k \in U, && k = 0,\dots,N-1 \\
& \mathbf{x}_N \in X_f \\
& \mathbf{u}_k = \mathbf{u}_{k-1} + \delta\mathbf{u}_k, && k = 0,\dots,N-1 \\
& \mathbf{x}_0 = \mathbf{x}(t), \quad \mathbf{u}_{-1} = \mathbf{u}(t-1)
\end{aligned}$$

- The control input is then $\mathbf{u}(t) = \delta\mathbf{u}_0^* + \mathbf{u}(t-1)$

Robust MPC

- We have so far not explicitly considered disturbances in constraint satisfaction
- Consider system of the form

$$\mathbf{x}_{k+1} = A\mathbf{x}_k + B\mathbf{u}_k + \mathbf{w}_k$$
$$\mathbf{w}_k \in W \quad \forall k$$

  with constraints $\mathbf{x} \in X, \mathbf{u} \in U$, and W is bounded.

- Can we guarantee stability and persistent feasibility for this system?

Robust optimal control problem

$$J_0^*(\mathbf{x}(t)) = \max_{\mathbf{w}_0,\dots,\mathbf{w}_{N-1}} \min_{\mathbf{u}_0,\dots,\mathbf{u}_{N-1}} p(\mathbf{x}_N) + \sum_{k=0}^{N-1} c(\mathbf{x}_k, \mathbf{u}_k)$$

$$\begin{aligned}
\text{subject to} \quad & \mathbf{x}_{k+1} = A\mathbf{x}_k + B\mathbf{u}_k + \mathbf{w}_k, \quad && k = 0, \dots, N-1 \\
& \mathbf{x}_k \in X,\ \mathbf{u}_k \in U,\ \mathbf{w}_k \in W \quad && k = 0, \dots, N-1 \\
& \mathbf{x}_N \in X_f \\
& \mathbf{x}_0 = \mathbf{x}(t)
\end{aligned}$$

Robust MPC

- Key idea: consider forward reachable sets at each time

$$S_0(\mathbf{x}_0) = \{\mathbf{x}(0)\}$$
$$S_k(\mathbf{x}_0, \mathbf{u}_{0:k-1}) = AS_{k-1}(\mathbf{x}_0, \mathbf{u}_{0:k-2}) + B\mathbf{u}_{k-1} + W$$

All trajectories in these “tubes” must satisfy constraints.

Robust MPC

$$J_0^*(\mathbf{x}(t)) = \max_{\mathbf{w}_0,\ldots,\mathbf{w}_{N-1}} \quad \min_{\mathbf{u}_0,\ldots,\mathbf{u}_{N-1}} p(\mathbf{x}_N) + \sum_{k=0}^{N-1} c(\mathbf{x}_k, \mathbf{u}_k)$$

$$\begin{aligned} \text{subject to} \quad & \mathbf{x}_{k+1} = A\mathbf{x}_k + B\mathbf{u}_k + \mathbf{w}_k, \qquad k = 0, \ldots, N-1 \\ & S_k \in X,\ \mathbf{u}_k \in U,\ \mathbf{w}_k \in W \qquad k = 0, \ldots, N-1 \\ & S_N \in X_f \\ & \mathbf{x}_0 = \mathbf{x}(t) \end{aligned}$$

Where $p(\mathbf{x}_N)$ is *robustly stable* and X_f is *robust control invariant.*

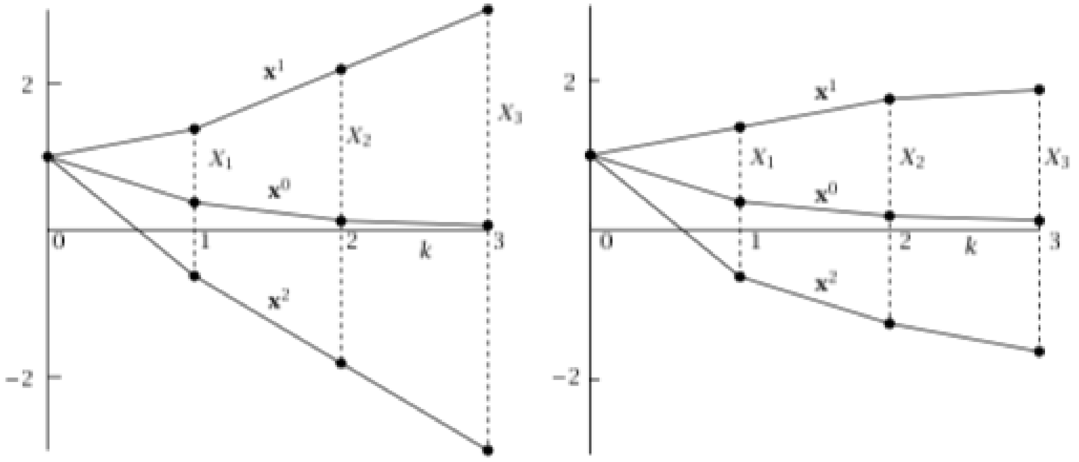
Tube MPC

- Forward tubes in robust MPC can be prohibitively large, motivating techniques to reduce their size
- Introduce nominal trajectory:

Nominal trajectory: $\bar{\mathbf{x}}_{k+1} = A\bar{\mathbf{x}}_k + B\mathbf{u}_k$

Error: $\mathbf{e}_k = \mathbf{x}_k - \bar{\mathbf{x}}_k$

Yields dynamics: $\mathbf{e}_{k+1} = A\mathbf{e}_k + \mathbf{w}_k$



(a) Open-loop trajectories. (b) Feedback trajectories.

- Consider feedback law: $\mathbf{u}_{\boldsymbol{k}} = \bar{\mathbf{u}}_k + F_\infty \mathbf{e}_k$

Tube MPC

- Adding error feedback gives dynamics

$$\bar{\mathbf{x}}_{k+1} = A\bar{\mathbf{x}}_k + B\bar{\mathbf{u}}_k$$
$$\mathbf{e}_{k+1} = (A + BF_\infty)\mathbf{e}_k + \mathbf{w}_k$$

Must choose $\bar{\mathbf{u}}_k$ to guarantee that $\bar{\mathbf{x}}_k + \mathbf{e}_k$ satisfy state, action, and terminal constraints for $k = 1, \ldots, N$.

What about nonlinearity?

- A very active field of research today!
- Control Barrier Functions (CBFs)
  - Analogous to Control Lyapunov Functions (CLFs) but for constraints
  - For general nonlinear dynamics $\dot{\mathbf{x}} = f(\mathbf{x}, \mathbf{u})$, if we can construct a function $B(\mathbf{x})$ satisfying
$$\max_{\mathbf{u} \in U} \nabla_{\mathbf{x}} \mathrm{B}(\mathbf{x})^{\mathrm{T}} f(\mathbf{x}, \mathbf{u}) \geq -\alpha\big(B(\mathbf{x})\big)$$
    then $C \coloneqq \{\mathbf{x} \in \mathbb{R}^{\mathrm{n}} \mid \mathrm{B}(\mathbf{x}) \geq 0\}$ is control invariant.
- Combining CBFs for persistent feasibility, CLFs for stability, horizon $N = 1$ results in a quadratic program for control-affine systems: CLF-CBF QPs
  - Ames, et al., “Control Barrier Function Based Quadratic Programs for Safety Critical Systems,” TAC, 2017.
- In practice, guarantees of persistent feasibility or stability are often sacrificed; heuristic choices of terminal constraint, cost are employed

Next time

- Back to learning!
  Learning and adaptive MPC