Model Predictive Control

Francesco Borrelli, Manfred Morari

UC Berkley

Institut für Automatik
ETH Zürich

Fall Semester 2014 (revised August 2014)

Table of Contents

Table of Contents

Table of Contents

Models of Dynamic Systems

- **Goal:** Introduce mathematical models to be used in Model Predictive Control (MPC) describing the behavior of dynamic systems
- Model classification: state space/transfer function, linear/nonlinear, time-varying/time-invariant, continuous-time/discrete-time, deterministic/stochastic
- If not stated differently, we use deterministic models
- Models of physical systems derived from first principles are mainly: nonlinear, time-invariant, continuous-time, state space models (*)
- Target models for standard MPC are mainly:
  linear, time-invariant, discrete-time, state space models (†)
- Focus of this section is on how to 'transform' (*) to (†)

Nonlinear, Time-Invariant, Continuous-Time, State Space Models (1/3)

$$\dot{x} = g(x, u)$$
$$y = h(x, u)$$

$x \in \mathbb{R}^n$	state vector	$g(x,u) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$	system dynamics
$u \in \mathbb{R}^m$	input vector	$h(x,u) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$	output function
$y \in \mathbb{R}^p$	output vector		

- Very general class of models
- Higher order ODEs can be easily brought to this form (next slide)
- Analysis and control synthesis generally hard $\rightarrow$ *linearization* to bring it to linear, time-invariant (LTI), continuous-time, state space form

Nonlinear, Time-Invariant, Continuous-Time, State Space Models (2/3)

Equivalence of one n-th order ODE and n 1-st order ODEs

$$x^{(n)} + g_n(x, \dot{x}, \ddot{x}, \ldots, x^{(n-1)}) = 0$$

Define

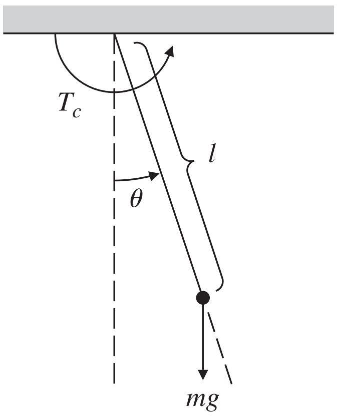
$$x_{i+1} = x^{(i)}, \quad i = 0, \ldots, n-1$$

Transformed system

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_3 \\ \vdots \quad & \quad \vdots \\ \dot{x}_{n-1} &= x_n \\ \dot{x}_n &= -g_n(x_1, x_2, \ldots, x_n) \end{aligned}$$

Nonlinear, Time-Invariant, Continuous-Time, State Space Models (3/3)

Example: Pendulum



Moment of inertia wrt. rotational axis	$m\ l^2$
Torque caused by external force	T_c
Torque caused by gravity	$m\,g\,l\sin(\theta)$

System equation $\quad m\ l^2\ \ddot{\theta} = T_c - m\,g\,l\sin(\theta)$

Using $x_1 \triangleq \theta, x_2 \triangleq \dot{\theta} = \dot{x}_1$ and $u \triangleq T_c/m\,l^2$ the system can be brought to standard form

$$\dot{x} = \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ -\frac{g}{l}\sin(x_1) + u \end{bmatrix} = g(x, u)$$

Output equation depends on the measurement configuration, i.e. if θ is measured then $y = h(x, u) = x_1$.

LTI Continuous-Time State Space Models (1/6)

$$\dot{x} = A^c x + B^c u$$
$$y = Cx + Du$$

$x \in \mathbb{R}^n$	state vector	$A^c \in \mathbb{R}^{n \times n}$	system matrix
$u \in \mathbb{R}^m$	input vector	$B^c \in \mathbb{R}^{n \times m}$	input matrix
$y \in \mathbb{R}^p$	output vector	$C \in \mathbb{R}^{p \times n}$	output matrix
		$D \in \mathbb{R}^{p \times m}$	throughput matrix

- Vast theory exists for the analysis and control synthesis of linear systems
- Exact solution (next slide)

LTI Continuous-Time State Space Models (2/6)

Solution to linear ODEs

- Consider the ODE (written with explicit time dependence) $\dot{x}(t) = A^c x(t) + B^c u(t)$ with initial condition $x_0 \triangleq x(t_0)$, then its solution is given by

$$x(t) = e^{A^c(t-t_0)} x_0 + \int_{t_0}^{t} e^{A^c(t-\tau)} Bu(\tau) \mathrm{d}\tau$$

where $e^{A^c t} \triangleq \sum_{n=0}^{\infty} \frac{(A^c t)^n}{n!}$

LTI Continuous-Time State Space Models (3/6)

- **Problem:** Most physical systems are nonlinear but linear systems are much better understood
- Nonlinear systems can be well approximated by a linear system in a 'small' neighborhood around a point in state space
- **Idea:** Control keeps the system around some operating point $\rightarrow$ replace nonlinear by a linearized system around operating point

First order Taylor expansion of $f(\cdot)$ around $\bar{x}$

$$f(x) \approx f(\bar{x}) + \left.\frac{\partial f}{\partial x'}\right|_{x=\bar{x}} (x-\bar{x}), \text{ with } \frac{\partial f}{\partial x'} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}$$

LTI Continuous-Time State Space Models (4/6)

Linearization

u_s keeps the system around stationary operating point x_s
$\rightarrow \dot{x}_s = g(x_s, u_s) = 0, y_s = h(x_s, u_s)$

$$\begin{aligned}
\dot{x} &= \underbrace{g(x_s,u_s)}_{=0} + \underbrace{\left.\frac{\partial g}{\partial x'}\right|_{\substack{x=x_s\\u=u_s}}}_{=A^c} \underbrace{(x-x_s)}_{=\Delta x} + \underbrace{\left.\frac{\partial g}{\partial u'}\right|_{\substack{x=x_s\\u=u_s}}}_{=B^c} \underbrace{(u-u_s)}_{=\Delta u} \\
\Rightarrow \dot{x} - \underbrace{\dot{x}_s}_{=0} &= \Delta\dot{x} = A^c \Delta x + B^c \Delta u \\
y &= \underbrace{h(x_s,u_s)}_{y_s} + \underbrace{\left.\frac{\partial h}{\partial x'}\right|_{\substack{x=x_s\\u=u_s}}}_{=C} (x-x_s) + \underbrace{\left.\frac{\partial h}{\partial u'}\right|_{\substack{x=x_s\\u=u_s}}}_{=D} (u-u_s) \\
\Rightarrow \underbrace{\Delta y}_{y-y_s} &= C\Delta x + D\Delta u
\end{aligned}$$

LTI Continuous-Time State Space Models (5/6)

Linearization

- The linearized system is written in terms of *deviation* variables $\Delta x, \Delta u$, Δy
- Linearized system is only a good approximation for 'small' $\Delta x, \Delta u$
- Subsequently, instead of Δx, Δu and Δy, x, u and y are used for brevity

LTI Continuous-Time State Space Models (6/6)

Example: Linearization of pendulum equations

$$\dot{x} = \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ -\frac{g}{l}\sin(x_1) + u \end{bmatrix} = g(x,u)$$

$$y = x_1 = h(x,u)$$

Want to keep the pendulum around $x_s = (\pi/4, 0)' \rightarrow u_s = \frac{g}{l}\sin(\pi/4)$

$$A^c = \left.\frac{\partial g}{\partial x'}\right|_{\substack{x=x_s\\u=u_s}} = \begin{bmatrix} 0 & 1 \\ -\frac{g}{l}\cos(\pi/4) & 0 \end{bmatrix}, \quad B^c = \left.\frac{\partial g}{\partial u'}\right|_{\substack{x=x_s\\u=u_s}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$C = \left.\frac{\partial h}{\partial x'}\right|_{\substack{x=x_s\\u=u_s}} = [1 \;\; 0], \quad D = \left.\frac{\partial h}{\partial u'}\right|_{\substack{x=x_s\\u=u_s}} = 0$$

Nonlinear, Time-Invariant, Discrete-Time, State Space Models

- Nonlinear discrete-time systems are described by difference equations

$$\begin{aligned} x(k+1) &= g(x(k), u(k)) \\ y(k) &= h(x(k), u(k)) \end{aligned}$$

$x \in \mathbb{R}^n$	state vector	$g(x,u) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$	system dynamics
$u \in \mathbb{R}^m$	input vector	$h(x,u) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$	output function
$y \in \mathbb{R}^p$	output vector		

LTI Discrete-Time State Space Models (1/2)

- Linear discrete-time systems are described by linear difference equations

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k) \end{aligned}$$

- Inputs and outputs of a discrete-time system are defined only at discrete time points, i.e. its inputs and outputs are sequences defined for $k \in \mathbb{Z}^+$
- Discrete time systems describe either
  1. Inherently discrete systems, eg. bank savings account balance at the k-th month
  $$x(k+1) = (1+\alpha)x(k) + u(k)$$
  2. 'Transformed' continuous-time system

LTI Discrete-Time State Space Models (2/2)

- Vast majority of controlled systems not inherently discrete-time systems
- Controllers almost always implemented using microprocessors
- Finite computation time must be considered in the control system design $\rightarrow$ *discretize* the continuous-time system
- Discretization is the procedure of obtaining an 'equivalent' discrete-time system from a continuous-time system
- The discrete-time model describes the state of the continuous-time system only at particular instances t_k, $k \in \mathbb{Z}^+$ in time, where $t_{k+1} = t_k + T_s$ and T_s is called the sampling time
- Usually $u(t) = u(t_k) \;\; \forall t \in [t_k, t_{k+1})$ is assumed (and implemented)

In Summary: We Work With Discrete Time Models

We will use:

- Nonlinear Discrete Time

$$\begin{aligned} x(k+1) &= g(x(k), u(k)) \\ y(k) &= h(x(k), u(k)) \end{aligned}$$

- or LTI Discrete Time

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k) \end{aligned}$$

Discretization

We call **discretization** the procedure of obtaining an "equivalent" DT model from a CT one.

Euler Discretization of Nonlinear Models

1. Given CT model
$$\begin{aligned}\dot{x}^c(t) &= g^c(x^c(t), u^c(t))\\ y^c(t) &= h^c(x(t), u^c(t))\end{aligned}$$
2. Approximate $\frac{d}{dt}x^c(t) \simeq \frac{x^c(t+T_s)-x^c(t)}{T_s}$
3. T_s is the **sampling time**
4. Notation: $x(k) \triangleq x^c(t_0 + kT_s)$, $u(k) \triangleq u^c(t_0 + kT_s)$
5. Then DT model is

$$\begin{aligned}x(k+1) &= x(k) + T_s g^c(x(k), u(k)) = g(x(k), u(k))\\ y(k) &= h^c(x(k), u(k)) = h(x(k), u(k))\end{aligned}$$

Under regularity assumptions, if T_s is small and CT and DT have "same" initial conditions and inputs, then outputs of CT and DT systems "will be close"

Euler Discretization of Linear Models

1. Given CT model
$$\begin{aligned}\dot{x}^c(t) &= A^c\, x(t) + B^c\, u(t)\\ y^c(t) &= C^c\, x(t) + D^c\, u(t)\end{aligned}$$
2. the DT model obtained with Euler discretization is

$$\begin{aligned}x(k+1) &= Ax(k) + Bu(k)\\ y(k) &= Cx(k) + Du(k)\end{aligned}$$

with $A = I + T_s A^c$, $B = T_s B^c$, $C = C^c$, $D = D^c$.

- There are a variety of discretization approaches (matlab: help c2d)

ZOH Discretization (1/2)

Discretization of LTI continuous-time state space models

- Recall the solution of the ODE $x(t) = e^{A^c(t-t_0)}x_0 + \int_{t_0}^{t} e^{A^c(t-\tau)}Bu(\tau)\mathrm{d}\tau$
- Choose $t_0 = t_k$ (hence $x_0 = x(t_0) = x(t_k)$), $t = t_{k+1}$ and use $t_{k+1} - t_k = T_s$ and $u(t) = u(t_k) \quad \forall t \in [t_k, t_{k+1})$

$$
\begin{aligned}
x(t_{k+1}) &= e^{A^c T_s}x(t_k) + \int_{t_k}^{t_{k+1}} e^{A^c(t_{k+1}-\tau)}B^c\mathrm{d}\tau u(t_k) \\
&= \underbrace{e^{A^c T_s}}_{\triangleq A} x(t_k) + \underbrace{\int_0^{T_s} e^{A^c(T_s-\tau')}B^c\mathrm{d}\tau'}_{\triangleq B} u(t_k) \\
&= Ax(t_k) + Bu(t_k)
\end{aligned}
$$

- We found the *exact* discrete-time model predicting the state of the continuous-time system at time t_{k+1} given $x(t_k)$, $k \in \mathbb{Z}_+$ under the assumption of a constant $u(t)$ during a sampling interval
- $B = (A^c)^{-1}(A - I)B^c$, if A^c invertible

ZOH Discretization (2/2)

Example: Discretization of the linearized pendulum equations
Using $g/l = 10[s^{-2}]$ the pendulum equations linearized about $x_s = (\pi/4, 0)$ are given by

$$
\dot{x}(t) = \begin{pmatrix} 0 & 1 \\ -10/\sqrt{2} & 0 \end{pmatrix} x(t) + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u(t)
$$

Discretizing the continuous-time system using the definitions of A and B, and $T_s = 0.1$ s, we get the following discrete-time system

$$
x(k+1) = \begin{pmatrix} 0.965 & 0.099 \\ -0.699 & 0.965 \end{pmatrix} x(k) + \begin{pmatrix} 0.005 \\ 0.100 \end{pmatrix} u(k)
$$

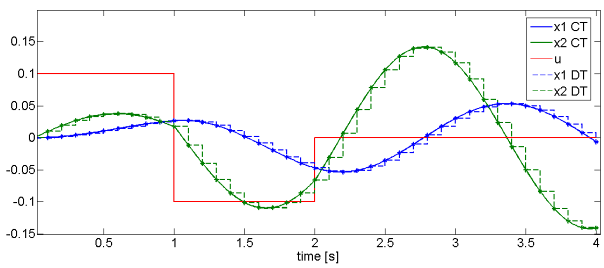


Table of Contents

Analysis of LTI Discrete-Time Systems

- **Goal:** Introduce the concepts of stability, controllability and observability
- From this point on we consider only discrete-time LTI systems for the rest of the lecture

Coordinate Transformations (1/2)

- Consider again the system

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k) \end{aligned}$$

- Input-output behavior, i.e. the sequence $\{y(k)\}_{k=0,1,2\ldots}$ entirely defined by $x(0)$ and $\{u(k)\}_{k=0,1,2\ldots}$
- Infinitely many choices of the state that yield the same input-output behavior
- Certain choices facilitate system analysis

Coordinate Transformations (2/2)

- Consider the linear transformation $\tilde{x} = Tx$ with $\det(T) \neq 0$ (invertible)

$$\begin{aligned} T^{-1}\tilde{x}(k+1) &= AT^{-1}\tilde{x}(k) + Bu(k) \\ y(k) &= CT^{-1}\tilde{x}(k) + Du(k) \end{aligned}$$

or

$$\begin{aligned} \tilde{x}(k+1) &= \underbrace{TAT^{-1}}_{\tilde{A}}\tilde{x}(k) + \underbrace{TB}_{\tilde{B}}u(k) \\ y(k) &= \underbrace{CT^{-1}}_{\tilde{C}}\tilde{x}(k) + \underbrace{D}_{\tilde{D}}u(k) \end{aligned}$$

- Note: $u(k)$ and $y(k)$ are unchanged

Stability of Linear Systems (1/3)

Theorem: Asymptotic Stability of Linear Systems

The LTI system

$$x(k+1) = Ax(k)$$

is globally asymptotically stable

$$\lim_{k\to\infty} x(k) = 0, \forall x(0) \in \mathbb{R}^n$$

if and only if $|\lambda_i| < 1$, $\forall i = 1, \cdots, n$ where λ_i are the eigenvalues of A. [1]

[1] for cont., time LTI systems $\dot{x} = Ax$, the conditions is $Re(\lambda_i) < 0$

Stability of Linear Systems (2/3)

"Proof" of asymptotic stability condition

- Assume that A has n linearly independent eigenvectors $e_1, \cdots, e_n$ then the coordinate transformation $\tilde{x} = [e_1, \cdots, e_n]^{-1}x = Tx$ transforms an LTI discrete-time system to

$$\tilde{x}(k+1) = TAT^{-1}\tilde{x}(k) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} \tilde{x}(k) = \Lambda \tilde{x}(k)$$

- The state $\tilde{x}(k)$ can be explicitly formulated as a function of $\tilde{x}(0) = Tx(0)$

$$\tilde{x}(k) = \Lambda^k \tilde{x}(0) = \begin{pmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n^k \end{pmatrix} \tilde{x}(0)$$

Stability of Linear Systems (3/3)

"Proof" of asymptotic stability condition

- We thus have that

$$\begin{aligned}\tilde{x}(k) = \Lambda^k \tilde{x}(0) \quad &\Rightarrow \quad |\tilde{x}(k)| = |\Lambda^k \tilde{x}(0)| \quad \text{(component-wise)} \\ &\Rightarrow \quad |\tilde{x}(k)| = |\Lambda^k| \cdot |\tilde{x}(0)| \\ &\Rightarrow \quad |\tilde{x}_i(k)| = |\lambda_i^k| \cdot |\tilde{x}_i(0)| = |\lambda_i|^k \cdot |\tilde{x}_i(0)|\end{aligned}$$

- If any $|\lambda_i| \geq 1$ then $\lim_{k\to\infty} \tilde{x}(k) \neq 0$ for $\tilde{x}(0) \neq 0$. On the other hand if $|\lambda_i| < 1 \ \forall i \in 1, \cdots, n$ then $\lim_{k\to\infty} \tilde{x}(k) = 0$ and we have asymptotic stability
- If the system does not have n linearly independent eigenvectors it can not be brought into diagonal form and Jordan matrices have to be used for the proof but the assertions still hold

Stability of Nonlinear Systems (1/5)

- For nonlinear systems there are many definitions of stability.
- Informally, we define a system to be stable in the sense of Lyapunov, if it stays in any arbitrarily small neighborhood of the origin when it is disturbed slightly.
- In the following we always mean "stability" in the sense of Lyapunov.
- We consider first the stability of a nonlinear, time-invariant, discrete-time system

$$x_{k+1} = g(x_k) \tag{1}$$

  with an *equilibrium point* at 0, i.e. $g(0) = 0$.
- Note that system (1) encompasses any open- or closed-loop autonomous system.
- We will then derive simpler stability conditions for the specific case of LTI systems.
- Note that always stability is a property of an equilibrium point of a system.

Stability of Nonlinear Systems (2/5)

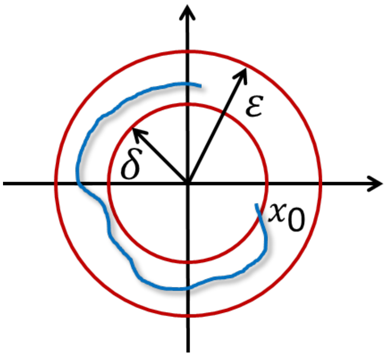
Definitions

Formally, the equilibrium point $x = 0$ of a system (1) is

- *stable* if for every $\epsilon > 0$ there exists a $\delta(\epsilon)$ such that

$$||x_0|| < \delta(\epsilon) \to ||x_k|| < \epsilon, \forall k \geq 0$$

- *unstable* otherwise.



An equilibrium point $x = 0$ of system (1) is

- *asymptotically stable* in $\Omega \subseteq \mathbb{R}^n$ if it is Lyapunov stable and

$$\lim_{k \to \infty} x_k = 0, \ \forall x_0 \in \Omega$$

- *globally asymptotically stable* if it is asymptotically stable and $\Omega = \mathbb{R}^n$

Stability of Nonlinear Systems (3/5)

Lyapunov functions

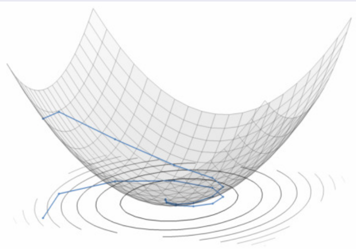
- We can show stability by constructing a *Lyapunov function*
- Idea: A mechanical system is asymptotically stable when the total mechanical energy is decreasing over time (friction losses). A Lyapunov function is a system theoretic generalization of energy

Definition: Lyapunov function

Consider the equilibrium point $x = 0$ of system (1). Let $\Omega \subset \mathbb{R}^n$ be a closed and bounded set containing the origin. A function $V : \mathbb{R}^n \to \mathbb{R}$, continuous at the origin, finite for every $x \in \Omega$, and such that

$$V(0) = 0 \text{ and } V(x) > 0, \ \forall x \in \Omega \setminus \{0\}$$
$$V(g(x_k)) - V(x_k) \leq -\alpha(x_k) \ \forall x_k \in \Omega \setminus \{0\}$$

where $\alpha : \mathbb{R}^n \to \mathbb{R}$ is continuous positive definite,



is called a *Lyapunov function*.

Stability of Nonlinear Systems (4/5)

Lyapunov theorem

Theorem: Lyapunov stability (asymptotic stability)

If a system (1) admits a Lyapunov function $V(x)$, then $x = 0$ is *asymptotically stable* in Ω.

Theorem: Lyapunov stability (global asymptotic stability)

If a system (1) admits a Lyapunov function $V(x)$ that additionally satisfies

$$\|x\| \to \infty \Rightarrow V(x) \to \infty,$$

then $x = 0$ is *globally asymptotically stable*.

Stability of Nonlinear Systems (5/5)

Remarks

- Note that the Lyapunov theorems only provide sufficient conditions
- Lyapunov theory is a powerful concept for proving stability of a control system, but for general nonlinear systems it is usually difficult to find a Lyapunov function
- Lyapunov functions can sometimes be derived from physical considerations
- One common approach:
  - Decide on *form* of Lyapunov function (e.g., quadratic)
  - Search for parameter values e.g. via optimization so that the required properties hold
- For linear systems there exist constructive theoretical results on the existence of a quadratic Lyapunov function

Global Lyapunov Stability of Linear Systems (1/3)

- Consider the linear system

$$x(k+1) = Ax(k) \tag{2}$$

- Take $V(x) = x'Px$ with $P > 0$ (positive definite) as a candidate Lyapunov function. It satisfies $V(0) = 0$, $V(x) > 0$ and $\|x\| \to \infty \Rightarrow V(x) \to \infty$.
- Check 'energy decrease' condition

$$\begin{aligned} V(Ax(k)) - V(x(k)) &= x'(k)A'PAx(k) - x'(k)Px(k) \\ &= x'(k)(A'PA - P)x(k) \leq -\alpha(x(k)) \end{aligned}$$

- We can choose $\alpha(x(k)) = x'(k)Qx(k)$, $Q > 0$. Hence, the condition can be satisfied if a $P > 0$ can be found that solves the *discrete-time Lyapunov equation*

$$A'PA - P = -Q, \;\; Q > 0. \tag{3}$$

Global Lyapunov Stability of Linear Systems (2/3)

Theorem: Existence of solution to the DT Lyapunov equation

The discrete-time Lyapunov equation (3) has a unique solution $P > 0$ if and only if A has all eigenvalues inside the unit circle, i.e. if the system $x(k+1) = Ax(k)$ is stable.

- Therefore, for LTI systems global asymptotic Lyapunov stability is not only sufficient but also necessary, and it agrees with the notion of stability based on eigenvalue location.
- Note that stability is always “global” for linear systems.

Global Lyapunov Stability of Linear Systems (3/3)

Property of P

- The matrix P can also be used to determine the infinite horizon cost-to-go for an asymptotically stable autonomous system $x(k+1) = Ax(k)$ with a quadratic cost function determined by Q.
- More precisely, defining $\Psi(x(0))$ as

$$\Psi(x(0)) = \sum_{k=0}^{\infty} x(k)'Qx(k) = \sum_{k=0}^{\infty} x(0)'(A^k)'QA^k x(0) \tag{4}$$

we have that

$$\Psi(x(0)) = x(0)'Px(0). \tag{5}$$

"Proof"

- Define $H_k \triangleq (A^k)'QA^k$ and $P \triangleq \sum_{k=0}^{\infty} H_k$ (limit of the sum exists because the system is assumed asymptotically stable).
- We have that $A'H_kA = (A^{k+1})'QA^{k+1} = H_{k+1}$.
- Thus $A'PA = \sum_{k=0}^{\infty} A'H_kA = \sum_{k=0}^{\infty} H_{k+1} = \sum_{k=1}^{\infty} H_k = P - H_0 = P - Q$.

Controllability (1/3)

- **Definition:** A system $x(k+1) = Ax(k) + Bu(k)$ is *controllable*[2] if for any pair of states $x(0), x^*$ there exists a finite time N and a control sequence $\{u(0), \cdots, u(N-1)\}$ such that $x(N) = x^*$,i.e.

$$x^* = x(N) = A^N x(0) + \begin{pmatrix} B & AB & \cdots & A^{N-1}B \end{pmatrix} \begin{pmatrix} u(N-1) \\ u(N-2) \\ \vdots \\ u(0) \end{pmatrix}$$

- It follows from the *Cayley-Hamilton* theorem that A^k can be expressed as linear combinations of $A^i, i \in 0, 1, \cdots, n$ for $k \geq n$. Hence for all $N \geq n$

$$\text{range}\begin{pmatrix} B & AB & \cdots & A^{N-1}B \end{pmatrix} = \text{range}\begin{pmatrix} B & AB & \cdots & A^{n-1}B \end{pmatrix}$$

[2] often referred to as "reachable" for discrete time systems

Controllability (2/3)

- If the system cannot be controlled to x^* in n steps, then it cannot in an arbitrary number of steps
- Define the *controllability matrix* $\mathcal{C} = \begin{pmatrix} B & AB & \cdots & A^{n-1}B \end{pmatrix}$
- The system is controllable if

$$\mathcal{C} \begin{pmatrix} u(n-1) \\ u(n-2) \\ \vdots \\ u(0) \end{pmatrix} = x^* - A^n x(0)$$

  has a solution for all right-hand sides (RHS)
- From linear algebra: solution exists for all RHS iff n columns of $\mathcal{C}$ are linearly independent
- $\Rightarrow$ Necessary and and sufficient condition for controllability is

$$\text{rank}(\mathcal{C}) = n$$

Controllability (3/3)

Remarks

- Another related concept is stabilizability
- A system is called *stabilizable* if it there exists an input sequence that returns the state to the origin asymptotically, starting from an arbitrary initial state
- A system is stabilizable iff all of its uncontrollable modes are stable
- Stabilizability can be checked using the following condition

$$\text{if rank}\left([\lambda_i I - A \,|\, B]\right) = n \quad \forall \lambda_i \in \Lambda_A^+ \;\Rightarrow\; (A, B) \text{ is stabilizable}$$

  where Λ_A^+ is the set of all eigenvalues of A lying on or outside the unit circle.
- Controllability implies stabilizability

Observability (1/3)

- Consider the following system with zero input

$$\begin{aligned} x(k+1) &= Ax(k) \\ y(k) &= Cx(k) \end{aligned}$$

- **Definition:** A system is said to be *observable* if there exists a finite N such that for every $x(0)$ the measurements $y(0), y(1), \cdots y(N-1)$ uniquely distinguish the initial state $x(0)$

Observability (2/3)

- Question of uniqueness of the linear equations

$$\begin{pmatrix} y(0) \\ y(1) \\ \vdots \\ y(N-1) \end{pmatrix} = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{N-1} \end{pmatrix} x(0)$$

- As previously we can replace N by n. (*Cayley-Hamilton*)
- Define $\mathcal{O} = \begin{pmatrix} C' & (CA)' & \cdots & (CA^{n-1})' \end{pmatrix}'$
- From linear algebra: solution is unique iff the n columns of $\mathcal{O}$ are linearly independent
- $\Rightarrow$ Necessary and sufficient condition for observability of system (A, C) is

$$\text{rank}(\mathcal{O}) = n$$

Observability (3/3)

Remarks

- Another related concept is detectability
- A system is called *detectable* if it possible to construct from the measurement sequence a sequence of state estimates that converges to the true state asymptotically, starting from an arbitrary initial estimate
- A system is detectable iff all of its unobservable modes are stable
- Detectability can be checked using the following condition

$$\text{if rank}\left([A' - \lambda_i I \mid C']\right) = n \quad \forall \lambda_i \in \Lambda_A^+ \;\Rightarrow\; (A, C) \text{ is detectable}$$

  where Λ_A^+ is the set of all eigenvalues of A lying on or outside the unit circle.
- Observability implies detectability