Data-driven control design

linear and nonlinear systems

Lecture 3

C. De Persis°, P. Tesi◇

° Institute of Engineering and Technology
University of Groningen
◇ Department of Information Engineering
University of Firenze

Recap from previous lectures

- We saw a method to design controllers directly from (noisy) data
- Controller design is based on low-complexity experiments and data-dependent semi-definite programs (SDP)
- The method provides stability guarantees
- The method is interpretable
- So far, we have considered linear dynamics

The remaining part of the course will be devoted to nonlinear systems

Problem overview

The framework (general)

Consider a nonlinear system

$$\begin{cases} x(k+1) = f(x(k), u(k), d(k)) \\ y(k) = h(x(k), u(k), n(k)) \end{cases}$$

- $x \in \mathbb{R}^n$ state, $u \in \mathbb{R}^m$ control, $y \in \mathbb{R}^p$ output
- $d \in \mathbb{R}^s$ unmeasured disturbance
- $n \in \mathbb{R}^q$ unmeasured noise
- f, h unknown functions

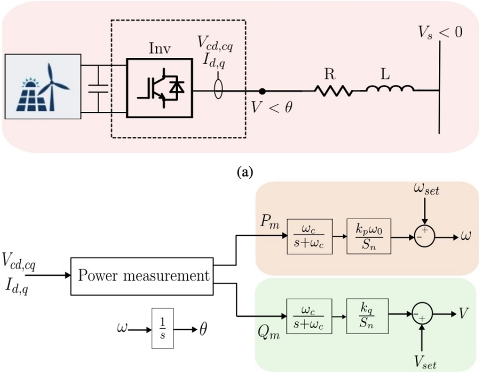
We want to design a stabilizing controller (stabilizing in some sense) based on a dataset of input-output samples collected from the system.

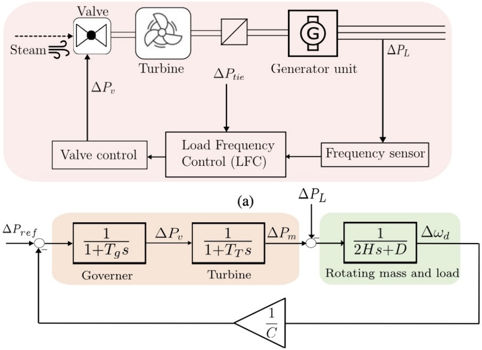
Motivations for direct design

System identification of nonlinear systems is difficult and time consuming, even famous methods (NARMAX, Volterra series) can fail

---

Example Distributed energy resource (DER) systems



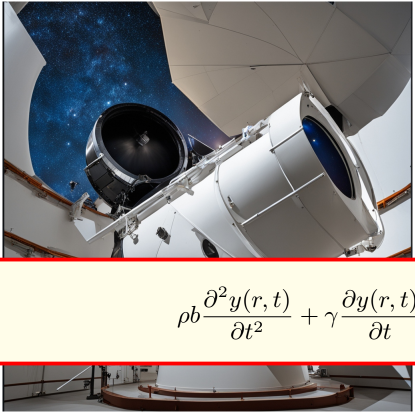


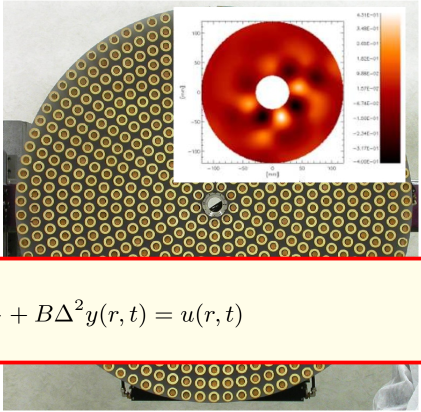
Courtesy of Dr. Javad Khazaei, Lehigh University, PA, USA

Best predictive models are not necessarily optimal for control design. Actually, sometimes models can hardly be used for control design

---

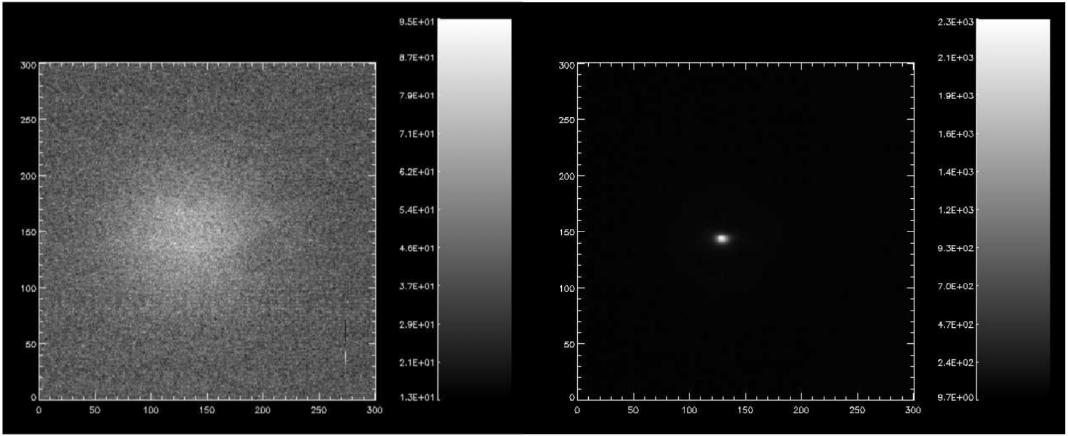
Example Adaptive optics systems





$$\rho b \frac{\partial^2 y(r,t)}{\partial t^2} + \gamma \frac{\partial y(r,t)}{\partial t} + B\Delta^2 y(r,t) = u(r,t)$$

Image without and with AO control system



Challenges of direct design

Some of the challenges are those inherent to nonlinear control

- control design is not systematic
- stability is more tricky (e.g., GAS $\not\Rightarrow$ ISS)
- stability properties are generally local

Some of the challenges are peculiar of data-driven control

- dynamics are unknown (at best, uncertain)
- noise corrupts our information on the dynamics

Providing theoretical guarantees is the main bottleneck (PID design, Unfalsified Control, VRFT, Reinforcement Learning, Neural Nets)

Roadmap

We will study direct data-driven control design for nonlinear systems.

We will study stabilization via:

1. Lyapunov's linearization method
2. Approximate feedback linearization (this afternoon)
3. Contractive design (tomorrow)

Recap of basics facts from control theory

Equilibria and asymptotic stability

Consider a smooth nonlinear system $x^+ = F(x)$ where $F : \mathbb{R}^n \to \mathbb{R}^n$ is a vector field.

An **equilibrium** $x_e \in \mathbb{R}^n$ is a vector such that $x_e = F(x_e)$.

Suppose F has an equilibrium at x_e. Then

- The equilibrium is called **Lyapunov stable** if for every $\epsilon > 0$ there exists $\delta > 0$ such that $\|x(0) - x_e\| < \delta$ implies $\|x(t) - x_e\| < \epsilon$ for all $t \geq 0$
- The equilibrium is called **asymptotically stable** if it is Lyapunov stable and there exists $\delta > 0$ such that $\|x(0) - x_e\| < \delta$ implies $\lim_{t\to\infty} \|x(t) - x_e\| = 0$

Linear systems $x^+ = \Phi x$:

▷ All equilibria have the same stability properties

▷ We use the term "system stability"

▷ Asymptotic stability is a "global" property

Lyapunov's direct method

Definition (Definite function) A function $V : \mathbb{R}^n \to \mathbb{R}$ is locally positive definite if there exists $r > 0$ such that $V(0) = 0$ and $V(x) > 0$ for all $\|x\| \leq r$, $x \neq 0$. It is positive definite if $r = +\infty$.

Theorem (Lyapunov direct method) Consider a smooth nonlinear system $x^+ = F(x)$ and suppose that the system has an equilibrium at $x_e = 0$, so that $F(0) = 0$. If there exists a continuous function $V(x)$ such that

- $V(x)$ is locally positive definite
- $\mathrm{d}V(x) := V(F(x)) - V(x)$ is locally negative definite

then $x_e = 0$ is an asymptotically stable equilibrium. It is globally asymptotically stable if $V(x)$ and $\mathrm{d}V(x)$ are positive and negative definite, respectively, and $\lim_{\|x\| \to \infty} V(x) = \infty$.

Linear systems $x^+ = \Phi x$:

- ▷ $\exists V(x) \iff P \succ 0$ such that $\Phi^\top P \Phi - P \prec 0$
- ▷ In this case, $V(x)$ can be chosen quadratic: $V(x) = x^\top P x$

Sketch of proof (LAS)

- Let $\mathcal{S}$ be a set where V and $\mathrm{d}V$ are locally pd and nd, and F is locally Lipschitz.
- Given $\epsilon > 0$, choose $r \in (0, \epsilon]$ such that $\mathcal{B}_r := \{x : \|x\| \leq r\} \subseteq \mathcal{S}$.
- Pick $\alpha > 0$ such that $\Omega := \{x : V(x) \leq \alpha\} \subseteq \mathcal{B}_r$.
- Note that if $x(k) \in \Omega$ then $x(k+1) \in \Omega$ because $\mathrm{d}V$ is nd inside Ω.
- Thus $x(0) \in \Omega$ implies $x(k) \in \Omega$ for all $k \geq 0$.
- Since V is continuous there exists $\delta \in (0, r]$ such that $\|x\| < \delta$ implies $V(x) < \alpha$.
- Thus, $\|x(0)\| < \delta \implies x(0) \in \Omega \implies x(k) \in \Omega\ \forall k \implies x(k) \in \mathcal{B}_r\ \forall k \implies x(0)$ is stable.

- As $\mathrm{d}V$ is nd inside $\mathcal{B}_r$, there exists $c \geq 0$ such that $V(x(k)) \to c$ as $k \to \infty$.
- Suppose by contradiction that $c > 0$.
- Since $c > 0$ there exists $d > 0$ such that $x(k)$ lies outside $\mathcal{B}_d := \{x : \|x\| \leq d\}$.
- Let $\omega := \min_{\|x\| \in [d,r]} V(x) - V(F(x))$ which is positive because $d > 0$.
- We also know that for any $\underline{\omega}$ there exists $\underline{k}$ such that $V(x(k)) \leq c + \underline{\omega}$ for all $k \geq \underline{k}$.
- Pick $\underline{\omega} < \omega$. Since $x(k)$ is such that $\|x(k)\| \in [c, r]$ we have $V(x) - V(F(x)) \geq \omega$.
- After $\underline{k}$, $V(F(x(k))) \leq V(x(k)) - \omega \leq c + \underline{\omega} - \omega < c$. A contradiction.

Stability inequalities for linear systems revisited

Fact 1 (Lyapunov stability inequality) System $x^+ = \Phi\, x$ is asymptotically stable if and only if there exists $P \succ 0$ such that

$$\Phi^\top P \Phi - P \prec 0 \tag{1}$$

Fact 2 (Another Lyapunov stability inequality) System $x^+ = \Phi x$ is asymptotically stable if and only if there exist $\Omega \succ 0, P \succ 0$ such that

$$\Phi^\top P \Phi - P + \Omega \prec 0 \tag{2}$$

(Fact 2 $\Longrightarrow$ Fact 1) Obvious.
(Fact 1 $\Longrightarrow$ Fact 2) Suppose there exists $P \succ 0$ such that $\Phi^\top P \Phi - P \prec 0$.
For any $\Theta \succ 0$ the inequality

$$\Phi^\top P \Phi - P + \epsilon \Theta \prec 0 \tag{3}$$

is satisfied for $\epsilon > 0$ sufficiently small, which implies (2) with $\Omega = \epsilon \Theta$.

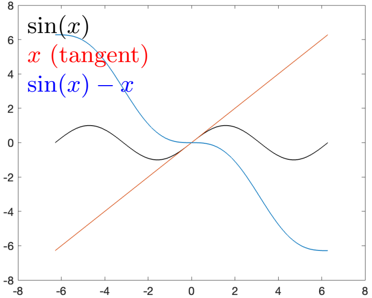
Linear approximation

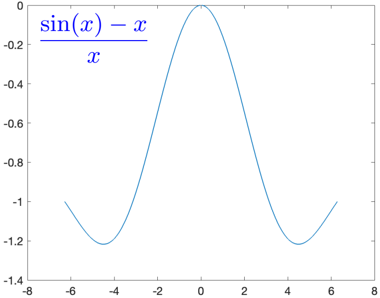
At the core of another famous Lyapunov's method is the linear approximation.
Given a smooth function $y = F(x)$ of one real variable and an arbitrary real a, Taylor's theorem says that

$$F(x) = F(a) + \left.\frac{dF}{dx}\right|_{x=a} (x-a) + \rho(x)$$

where $\rho(x)$ is the remainder, which is a function that satisfies $\lim_{x\to a} \frac{\rho(x)}{x-a} = 0$

Example $F(x) = \sin(x)$, $a = 0 \implies F(x) = x + \sin(x) - x$





Lyapunov's linearization method

Building on linear approximation is Lyapunov's linearization method or Taylor's method.

Consider a smooth system $x(k+1) = F(x(k))$, $x \in \mathbb{R}^n$, with an equilibrium point $x_e = 0$. The **first-order expansion** of the dynamics around the equilibrium gives [1]

$$\begin{aligned} x(k+1) &= \left(\left. \frac{\partial F}{\partial x} \right|_{x=0} \right) x(k) + \rho(x(k)) \\ &=: \Phi x(k) + \rho(x(k)) \end{aligned}$$

where $\rho(x)$ is the remainder, which is a function that satisfies $\lim_{x \to 0} \frac{\|\rho(x)\|}{\|x\|} = 0$.

---

[1] **Jacobian matrix** of partial derivatives: $\frac{\partial F}{\partial x} = \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_n}{\partial x_1} & \cdots & \frac{\partial F_n}{\partial x_n} \end{pmatrix} \in \mathbb{R}^{n \times n}$

Theorem (Lyapunov's linearization method) Consider a nonlinear system $x^+ = F(x)$ with equilibrium $x_e = 0$ and consider the linear approximation $x^+ = \Phi x$ with $\Phi = \frac{\partial F}{\partial x}\Big|_{x=0}$. If $x^+ = \Phi x$ is asymptotically stable then $x_e = 0$ is an asymptotically stable equilibrium for $x^+ = F(x)$.

Taylor's expansion: $x^+ = \Phi x + \rho(x)$

Φ stable $\implies$ there exist $\Omega \succ 0$, $P \succ 0$ such that $\Phi^\top P \Phi - P + \Omega \prec 0$

Let $V(x) = x^\top P x$. Then,

$$\begin{aligned} \mathrm{d}V(x) := V(x^+) - V(x) &= (\Phi x + \rho(x))^\top P(\Phi x + \rho(x)) - x^\top P x \\ &= x^\top(\Phi^\top P \Phi - P)x + (2\Phi x + \rho(x))^\top P \rho(x) \\ &\leq -x^\top \Omega x + (2\Phi x + \rho(x))^\top P \rho(x) \end{aligned}$$

showing that $\mathrm{d}V(x)$ is negative definite around the origin.

Formally, as $\rho(x)$ converges to zero faster than linearly, for any $m > 0$ there exists $r > 0$ such that

$$\|\rho(x)\| \leq m\|x\| \quad \forall x \in \mathcal{B} = \{x : \|x\| \leq r\}$$

Thus

$$\begin{aligned} \mathrm{d}V(x) &\leq -x^\top \Omega x + (2\Phi x + \rho(x))^\top P\rho(x) \\ &\leq -\lambda_{\min}(\Omega)\|x\|^2 + 2m\|P\Phi\|\|x\|^2 + m^2\|P\|\|x\|^2 \end{aligned}$$

Picking m sufficiently small shows that $\mathrm{d}V(x)$ is negative definite around the origin.

Stabilization via Lyapunov's linearization

Consider a smooth nonlinear system $x^+ = f(x,u)$ with an equilibrium point $(x_e, u_e) = 0$. The Taylor's expansion now gives

$$\begin{aligned} x^+ &= \begin{bmatrix} \frac{\partial f}{\partial x}\Big|_{(x,u)=0} & \frac{\partial f}{\partial u}\Big|_{(x,u)=0} \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} + r(x,u) \\ &=: Ax + Bu + r(x,u) \end{aligned}$$

where $r(x,u)$ is the remainder. A linear control law $\underline{u = Kx}$ gives

$$x^+ = f(x, Kx) = (A + BK)x + r(x, Kx)$$

Corollary (Stabilization via Lyapunov's linearization) Consider system $x^+ = f(x,u)$ with an equilibrium $(x_e, u_e) = 0$, and let $A := \frac{\partial f}{\partial x}|_{(x,u)=0}$ and $B := \frac{\partial f}{\partial x}|_{(x,u)=0}$. Let $u = Kx$. If $x^+ = (A + BK)x$ is asymptotically stable $x_e = 0$ is an asymptotically stable equilibrium for $x^+ = f(x, Kx)$.

The linearization method can be extended to **nonzero equilibria**

Consider a smooth nonlinear system $x^+ = f(x,u)$ with a nonzero equilibrium (x_e, u_e). Let $\tilde{x} = x - x_e$, $\tilde{u} = u - u_e$. Taylor's expansion now returns

$$\begin{aligned} x^+ &= x_e + \left[\left.\frac{\partial f}{\partial x}\right|_{(x,u)=(x_e,u_e)} \quad \left.\frac{\partial f}{\partial u}\right|_{(x,u)=(x_e,u_e)} \right] \begin{bmatrix} x - x_e \\ u - u_e \end{bmatrix} + r(x,u) \\ \implies \qquad \tilde{x}^+ &= A\tilde{x} + B\tilde{u} + r(x,u) \end{aligned}$$

Hence, the problem reduces to designing a stabilizing control law $\tilde{u} = K\tilde{x}$.

The control law $\tilde{u} = K\tilde{x}$ has the explicit form

$$u = u_e + K(x - x_e)$$

which corresponds to feedback–feedforward architecture

Data-driven design via Lyapunov's linearization

(noise-free case)

The framework

Consider a nonlinear system

$$x(k+1) = f(x(k), u(k))$$

where:

- $x \in \mathbb{R}^n$ state, $u \in \mathbb{R}^m$ control
- f is an unknown smooth function
- $(x_e, u_e) = 0$ is a known equilibrium

Problem Suppose we perform an experiment on the system and collect the dataset

$$\mathbb{D} := \{\bar{x}(k), \bar{u}(k)\}_{k=0}^{T}$$

where the samples satisfy

$$\bar{x}(k+1) = f(\bar{x}(k), \bar{u}(k)), \quad k = 0, 1, \ldots, T$$

Using $\mathbb{D}$, design a control law $u(k) = Kx(k)$ that renders $x_e = 0$ an asymptotically stable equilibrium for the closed-loop system $x(k+1) = f(x(k), Kx(k))$.

Data-based representation of the closed-loop system

Taylor's form: $x(k+1) = Ax(k) + Bu(k) + r(x(k), u(k))$, with A, B, r unknown

---

Data satisfy $\bar{x}(k+1) = A\bar{x}(k) + B\bar{u}(k) + r(\bar{x}(k), \bar{u}(k))$, $k = 0, 1, \ldots, T$

The data-based relation of the system gives

$$\underbrace{\begin{bmatrix} \bar{x}(1) & \bar{x}(2) & \ldots & \bar{x}(T) \end{bmatrix}}_{X_1} =$$

$$A \cdot \underbrace{\begin{bmatrix} \bar{x}(0) & \bar{x}(1) & \ldots & \bar{x}(T-1) \end{bmatrix}}_{X_0} + B \cdot \underbrace{\begin{bmatrix} \bar{u}(0) & \bar{u}(1) & \ldots & \bar{u}(T-1) \end{bmatrix}}_{U_0}$$

$$+ \underbrace{\begin{bmatrix} r(\bar{x}(0), \bar{u}(0)) & r(\bar{x}(1), \bar{u}(1)) & \ldots & r(\bar{x}(T-1), \bar{u}(T-1)) \end{bmatrix}}_{R_0}$$

In compact form:

$$X_1 = AX_0 + BU_0 + R_0$$

Data-based relation: $X_1 = \begin{bmatrix} B & A \end{bmatrix} \begin{bmatrix} U_0 \\ X_0 \end{bmatrix} + R_0$.

Consider a controller $u = Kx$ and the resulting closed-loop system

$$\begin{aligned} x^+ &= Ax + Bu + r(x,u) \\ &= (A+BK)x + r(x,Kx) \end{aligned}$$

For any K, G that solve

$$\begin{bmatrix} K \\ I \end{bmatrix} = \begin{bmatrix} U_0 \\ X_0 \end{bmatrix} G$$

we have:

$$A + BK = \begin{bmatrix} B & A \end{bmatrix} \begin{bmatrix} K \\ I \end{bmatrix} = \begin{bmatrix} B & A \end{bmatrix} \begin{bmatrix} U_0 \\ X_0 \end{bmatrix} G = (X_1 - R_0)G$$

and the closed-loop system has the **data-based representation**

$$x^+ = (X_1 - R_0)Gx + r(x, U_0 G x)$$

Robust approach to Lyapunov's stability

The closed-loop system has the data-based representation

$$x^+ = (X_1 - R_0)Gx + r(x, U_0 Gx)$$

with R_0, r unknown.

Lyapunov's linearization: $(X_1 - R_0)G$ stable $\Longrightarrow$ origin is stable

---

As in Lecture 2

Since R_0 is unknown we solve

$$((X_1 - R)G)^\top P(X_1 - R)G - P + \Omega \prec 0 \quad \underline{\text{for all}}\ R \in \mathcal{R}$$

where $\mathcal{R}$ is a set we choose.

If $R_0 \in \mathcal{R}$ then $(X_1 - R_0)G$ will be stable.

Main result

Theorem Consider the nonlinear system $x^+ = f(x,u)$ whit equilibrium $(x_e, u_e) = (0,0)$. Let U_0, X_0, X_1 be data matrices and let $\mathcal{R} = \{R : RR^\top \preceq \Delta\}$, where Δ is chosen by the designer. Finally, let $\Theta \succ 0$ be arbitrary. Suppose there exist $S \in \mathbb{S}^{n\times n}, Y \in \mathbb{R}^{T\times n}$, $\varepsilon > 0$ such that

$$S = X_0 Y, \qquad \begin{bmatrix} S - \Theta & (X_1 Y)^\top & Y^\top \\ X_1 Y & S - \varepsilon\Delta & 0 \\ Y & 0 & \varepsilon I \end{bmatrix} \succ 0$$

If $R_0 \in \mathcal{R}$ then $K = U_0 Y S^{-1}$ stabilizes the equilibrium $x_e = 0$.

Let $G = YS^{-1}$. The two identities $K = U_0 G$ and $I = X_0 G$ imply $\begin{bmatrix} K \\ I \end{bmatrix} = \begin{bmatrix} U_0 \\ X_0 \end{bmatrix} G$

This implies $A + BK = (X_1 - R_0)G$

Since $R_0 \in \mathcal{R}$ by assumption, it is enough that $(X_1 - R)G$ is stable for all $R \in \mathcal{R}$

The LMI implies stability of $(X_1 - R)G$ for all $R \in \mathcal{R}$ (Lecture 2)

Suppose the LMI holds, i.e.,

$$\left[\begin{array}{cc|c} S-\Theta & (X_1Y)^\top & Y^\top \\ X_1Y & S-\varepsilon\Delta & 0 \\ \hline Y & 0 & \varepsilon I \end{array}\right] \succ 0$$

A Schur complement gives

$$\begin{bmatrix} S-\Theta & (X_1Y)^\top \\ X_1Y & S-\varepsilon\Delta \end{bmatrix} - \varepsilon^{-1}\begin{bmatrix} Y^\top \\ 0 \end{bmatrix}\begin{bmatrix} Y & 0 \end{bmatrix} \succ 0 \iff$$

$$\underbrace{\begin{bmatrix} S-\Theta & (X_1Y)^\top \\ X_1Y & S \end{bmatrix}}_{-V} - \varepsilon\begin{bmatrix} 0 \\ I \end{bmatrix}\Delta\underbrace{\begin{bmatrix} 0 & I \end{bmatrix}}_{N} - \varepsilon^{-1}\underbrace{\begin{bmatrix} Y^\top \\ 0 \end{bmatrix}}_{M}\begin{bmatrix} Y & 0 \end{bmatrix} \succ 0$$

By **Petersen's Lemma**, $V + \varepsilon^{-1}MM^\top + \varepsilon N^\top \Delta N \prec 0$ implies $V + MR^\top N + N^\top RM^\top \prec 0$ for all $R \in \mathcal{R} = \{RR^\top \preceq \Delta\}$

$$\begin{bmatrix} S-\Theta & (X_1Y)^\top \\ X_1Y & S \end{bmatrix} - \begin{bmatrix} Y^\top \\ 0 \end{bmatrix} R^\top \begin{bmatrix} 0 & I \end{bmatrix} - \begin{bmatrix} 0 \\ I \end{bmatrix} R \begin{bmatrix} Y & 0 \end{bmatrix} \succ 0 \quad \forall R \in \mathcal{R}$$

Condition

$$\begin{bmatrix} S-\Theta & (X_1Y)^\top \\ X_1Y & S \end{bmatrix} - \begin{bmatrix} Y^\top \\ 0 \end{bmatrix} R^\top \begin{bmatrix} 0 & I \end{bmatrix} - \begin{bmatrix} 0 \\ I \end{bmatrix} R \begin{bmatrix} Y & 0 \end{bmatrix} \succ 0 \quad \forall R \in \mathcal{R}$$

is equivalent to

$$\begin{bmatrix} S-\Theta & ((X_1-R)Y)^\top \\ (X_1-R)Y & S \end{bmatrix} \succ 0 \quad \forall R \in \mathcal{R}$$

Another Schur complement gives

$$((X_1-R)Y)^\top S^{-1}(X_1-R)Y - S + \Theta \prec 0 \quad \forall R \in \mathcal{R}$$

Pre- and post-multiplying for S^{-1} gives

$$S^{-1}((X_1-R)Y)^\top S^{-1}(X_1-R)YS^{-1} - S^{-1} + \underbrace{S^{-1}\Theta S^{-1}}_{\Omega} \prec 0 \quad \forall R \in \mathcal{R}$$

The change of variable $G = YS^{-1}$ gives

$$((X_1-R)G)^\top S^{-1}(X_1-R)G - S^{-1} + \Omega \prec 0 \quad \forall R \in \mathcal{R}$$

which is the Lyapunov condition with $P = S^{-1}$,
and $V(x) = x^\top Px$ is a valid Lyapunov function.

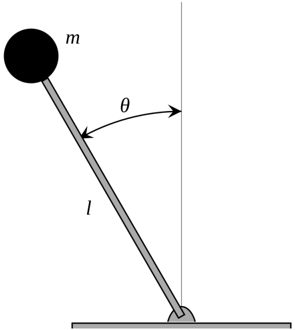
Remarks

- Design program as in the linear case
- The matrix R_0 plays a similar role as the disturbance matrix D_0 in the linear case, but is dynamics-dependent
- The choice of $\mathcal{R} = \{R : RR^\top \preceq \Delta\}$ requires bounds on the dynamics
- Since R_0 is dynamics-dependent, R_0 can be made as small as desired by running the experiment sufficiently close to the equilibrium
- If the linearized system is reachable the problem is feasible when R_0 is sufficiently small and u is suitably chosen (persistence of excitation is not enough because when we excite the dynamics we also excite R_0) [2]

[2] C. De Persis, P. Tesi, "Designing experiments for data-driven control of nonlinear systems", IFAC-PapersOnLine 2021
M. Alsalti, V.G. Lopez, M.A. Müller, "On the design of persistently exciting inputs for data-driven control of linear and nonlinear systems", IEEE L-CSS 2023

Nonlinear systems	Linear systems with disturbances
$x^+ = Ax + Bu + r(x,u)$	$x^+ = Ax + Bu + d$
type of program convex program	type of program convex program
choice of $\mathcal{R}$ requires bounds on dynamics	choice of $\mathcal{D}$ requires bounds on noise
feasibility of the program linearized system is reachable R_0 sufficiently small and u suitably chosen can be guaranteed by design with experiment close to the equilibrium	feasibility of the program system is reachable D_0 sufficiently small and u persistently exciting can be guaranteed by design (averaging) if noise has certain characteristics

Inverted pendulum



Dynamics:

$$\begin{cases} x_1^+ = x_1 + T_s x_2 \\ x_2^+ = \dfrac{T_s g}{l} \sin x_1 + \left(1 - \dfrac{T_s \mu}{ml^2}\right) x_2 + \dfrac{T_s}{ml^2} u \end{cases}$$

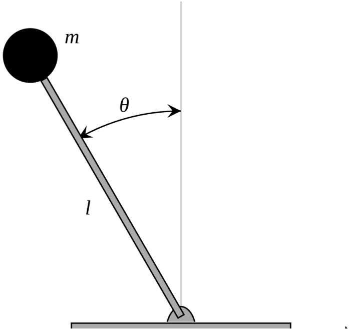
- x_1 = angular position θ
- x_2 = angular velocity
- m = mass to be balanced • l = length of the pendulum
- μ = friction • g = acceleration due to gravity • T_s = sampling time

---

Dynamics in Taylor's expansion around the equilibrium:

$$\begin{cases} x_1^+ = x_1 + T_s x_2 \\ x_2^+ = \dfrac{T_s g}{l} x_1 + \left(1 - \dfrac{T_s \mu}{ml^2}\right) x_2 + \dfrac{T_s}{ml^2} u + \dfrac{T_s g}{l} (\sin x_1 - x_1) \end{cases}$$

$$\begin{bmatrix} x_1^+ \\ x_2^+ \end{bmatrix} = \begin{bmatrix} 1 & T_s \\ \frac{T_s g}{l} & 1 - \frac{T_s \mu}{ml^2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{T_s}{ml^2} \end{bmatrix} u + \begin{bmatrix} 0 \\ \frac{T_s g}{l} \end{bmatrix} (\sin x_1 - x_1)$$



Dynamics:

$$\begin{bmatrix} x_1^+ \\ x_2^+ \end{bmatrix} = \begin{bmatrix} 1 & T_s \\ \frac{T_s g}{l} & 1 - \frac{T_s \mu}{ml^2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{T_s}{ml^2} \end{bmatrix} u + \begin{bmatrix} 0 \\ \frac{T_s g}{l} \end{bmatrix} (\sin x_1 - x_1)$$

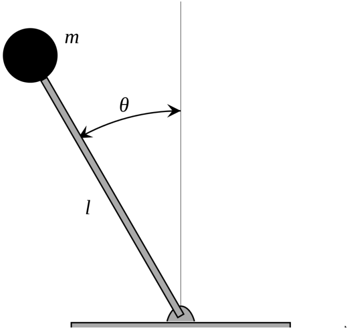
Suppose $m = 1, l = 1, g = 9.8, \mu = 0.01$ (unknown) and $T_s = 0.01$.

We run an experiment with $|u| \leq 0.5$ and solve the design program with the first $T = 100$ samples.

The program is solved with $\Delta = \begin{bmatrix} 0 & 0 \\ 0 & 1e\text{–}4 \end{bmatrix}$ and $R_0 R_0^\top \approx \begin{bmatrix} 0 & 0 \\ 0 & 1e\text{–}6 \end{bmatrix} \preceq \Delta$.

The program is feasible and returns $K = \begin{bmatrix} -18.2150 & -10.7779 \end{bmatrix}$.

Closed-loop behavior



Region of Attraction and Invariant sets

Stability properties with this method are local

Definition (Region of Attraction) Consider a dynamical system $x(k+1) = F(x(k))$ with an asymptotically stable equilibrium $x_e = 0$. The region of attraction (RoA) to x_e is the largest set $\mathcal{X}$ for which $x(0) \in \mathcal{X}$ implies $\lim_{t\to\infty} x(t) = 0$.

Definition (Invariant set) The set $\mathcal{X}$ is an invariant for system $x(k+1) = F(x(k))$ if for all $x(0) \in \mathcal{X}$ the solution $x(t) \in \mathcal{X}$ for $t > 0$.

Invariance:

- Milder property (does not require convergence)
- Useful for estimating RoA
- Useful for including safety constraints

Consider a nonlinear system in Taylor's form $x^+ = \Phi x + \rho(x)$ and suppose Φ stable.

Let $\Omega \succ 0$, $P \succ 0$ solve the Lyapunov inequality

$$\Phi^\top P \Phi - P + \Omega \prec 0$$

Letting $V(x) = x^\top P x$ we have

$$\begin{aligned} \mathrm{d}V(x) &:= V(x^+) - V(x) \\ &= (\Phi x + \rho(x))^\top P (\Phi x + \rho(x)) - x^\top P x \\ &= x^\top (\Phi^\top P \Phi - P) x + (2\Phi x + \rho(x))^\top P \rho(x) \\ &\le -x^\top \Omega x + (2\Phi x + \rho(x))^\top P \rho(x) \end{aligned}$$

We have

$$\mathrm{d}V(x) \leq \underbrace{-x^\top \Omega x + (2\Phi x + \rho(x))^\top P \rho(x)}_{\ell(x)}$$

Define $\boxed{\mathcal{X} := \{x : \ell(x) < 0\}}$ All we need is to find an invariant set contained in $\mathcal{X}$

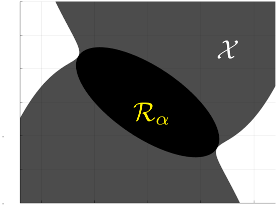
We call a **Lyapunov sub-level set** a set of the form

$$\mathcal{R}_\alpha := \{x : V(x) \leq \alpha\}, \ \alpha > 0$$

Consider any Lyapunov sub-level set $\mathcal{R}_\alpha$. If

$$\boxed{\mathcal{R}_\alpha \subseteq (\mathcal{X} \cup \{0\})}$$

then $\mathcal{R}_\alpha$ is an invariant set and provides estimate of the RoA



Data-based estimate of RoA

Recall the data-based representation:

$$x^+ = \underbrace{(X_1 - R_0)G}_{\Phi}\, x + \underbrace{r(x, U_0 G x)}_{\rho(x)}$$

and <u>we know</u> $\Omega \succ 0$ and $P \succ 0$ such that $\Phi^\top P \Phi - P + \Omega \prec 0$.

Letting $V(x) = x^\top P x$ we have

$$\begin{aligned}
\mathrm{d}V(x) &\leq -x^\top \Omega x + (2\Phi x + \rho(x))^\top P \rho(x) \\
&\leq -x^\top \Omega x + 2\|P\Phi\|\|x\|\|\rho(x)\| + \|P\|\|\rho(x)\|^2 \\
&\leq -x^\top \Omega x + \underbrace{2(\|PX_1G\| + \|P\|\sqrt{\|\Delta\|}\|G\|)\|x\|\delta(x) + \|P\|\delta(x)^2}_{\text{known over-approximation}}
\end{aligned}$$

where $\delta(x)$ is a known function such that $\|\rho(x)\| \leq \delta(x)$ for all x (globally for simplicity).

<u>We only need bounds on the dynamics</u>

Letting $V(x) = x^\top P x$ we have

$$\begin{aligned} \mathrm{d}V(x) &\leq -x^\top \Omega x + (2\Phi x + \rho(x))^\top P \rho(x) \\ &\leq -x^\top \Omega x + \underbrace{2(\|PX_1 G\| + \|P\|\sqrt{\|\Delta\|}\|G\|)\|x\|\delta(x) + \|P\|\delta(x)^2}_{\text{known over-approximation } h(x)} \end{aligned}$$

where $\delta(x)$ is a known function such that $\|\rho(x)\| \leq \delta(x)$ for all x.

Corollary (Invariant sets and RoA estimate based on data) Let $\delta(x)$ be a known function such that $\|\rho(x)\| \leq \delta(x)$ for all x, and define $\mathcal{X} := \{x : -x^\top \Omega x + h(x) < 0\}$. Any Lyapunov sub-level set

$$\mathcal{R}_\alpha := \{x : V(x) \leq \alpha\} \subseteq (\mathcal{X} \cup \{0\})$$

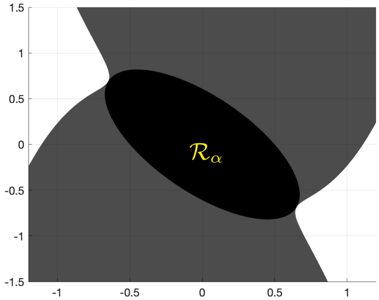
is an invariant set
and provides an estimate of the RoA,

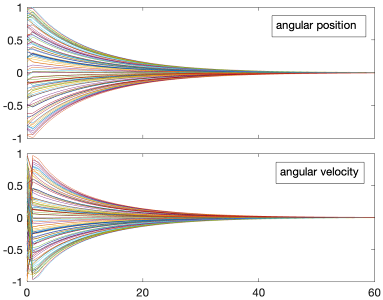
Inverted pendulum (cont'd)

Dynamics:

$$\begin{bmatrix} x_1^+ \\ x_2^+ \end{bmatrix} = \begin{bmatrix} 1 & T_s \\ \frac{T_s g}{l} & 1 - \frac{T_s \mu}{ml^2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{T_s}{ml^2} \end{bmatrix} u + \begin{bmatrix} 0 \\ \frac{T_s g}{l} \end{bmatrix} (\sin x_1 - x_1)$$

Estimate obtained using $\delta(x) = 2\|\sin x_1 - x_1\|$ (> 100% over-approximation)
The bounding function $\delta(x)$ can be obtained from physical considerations or from some estimation method, and the method can be just a data-fitting algorithm





Consider a scalar function $f : \mathcal{X} \to \mathbb{R}$, $\mathcal{X} \subseteq \mathbb{R}$
and a dataset $\mathbb{D} := \{(y_k, x_k),\ k = 0, 1, \ldots, T\}$ that carries information on f:

$$y_k = f(x_k), \quad k = 0, 1, \ldots, T$$

Consider a **(psd) kernel** $\kappa : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ and its RKHS $\mathcal{H}$.
Let

$$\mathbf{k}(x) := \begin{bmatrix} \kappa(x_0, x) & \kappa(x_1, x) & \cdots & \kappa(x_T, x) \end{bmatrix}^\top$$

and $K = \left[\kappa(x_r, x_k)\right]_{rk}$.

KERNEL RIDGE REGRESSION (KRR) (Schölkopf et al, 2001)

$$\underset{s_f \in \mathcal{H}}{\text{minimize}} \ \sum_{k=0}^{T} \left|y_k - s_f(x_k)\right|^2 + \lambda \cdot \|s_f\|_{\mathcal{H}}^2,$$

where $\lambda > 0$ controls the smoothness of the estimator, $\|\cdot\|_{\mathcal{H}}$ inner product on $\mathcal{H}$.
The minimizer is

$$s_f(x) = A\mathbf{k}(x)$$

where $A := [y_0 \ y_1 \ \cdots \ y_T]\,(\lambda I + K)^{-1}$.

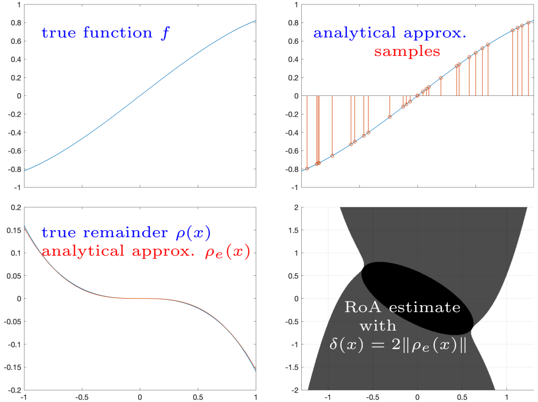
Estimate of the ROA obtained with Gaussian kernels $\kappa(s,x) = \exp(-(s-x)^2)$, $\lambda = 0.01$.

We collect $T = 30$ samples of the open-loop behavior from zero velocity and random position x_1, which gives the data-based relation

$$\underbrace{x_2(k+1)}_{y_k} = \underbrace{\frac{T_s g}{l}\sin x_1(k)}_{f(x_k)}, \quad k = 0, 1, \ldots, T$$

The minimizer is $s_f(x) = A\mathbf{k}(x)$. Since $\mathbf{k}(x)$ has analytic expression we can compute its Taylor's expansion to determine $\delta(x)$

$$A = \begin{bmatrix} 1.0798 \\ -2.7006 \\ -1.5346 \\ -0.2564 \\ 8.2279 \\ 0.8729 \\ -0.0697 \\ 0.5598 \\ 0.5986 \\ -0.1021 \\ 1.1864 \\ -3.6753 \\ -2.0488 \\ -0.2636 \\ -0.1826 \\ -2.9156 \\ \vdots \end{bmatrix}^\top, \quad \mathbf{k}(x) = \begin{bmatrix} \exp(-(-0.4237-x)^2) \\ \exp(-(0.2799-x)^2) \\ \exp(-(-0.0616-x)^2) \\ \exp(-(0.4780-x)^2) \\ \exp(-(0.0385-x)^2) \\ \exp(-(0.0011-x)^2) \\ \vdots \end{bmatrix}$$



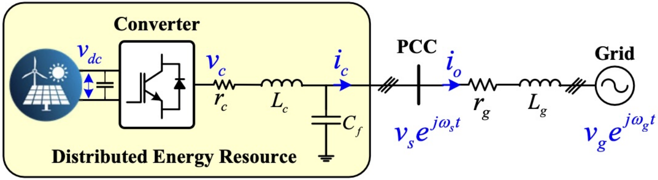
Case study: DER systems

Collaboration with College of Engineering at Lehigh University, PA, USA

A DER system involves power electronics interface (converter) and filtering component (RLC filter), which is connected to the main electricity grid at the point of common coupling (PCC). The power converter is mainly composed of a voltage source DC/AC converter fed by the energy resource (solar, wind, or battery) through a DC link.

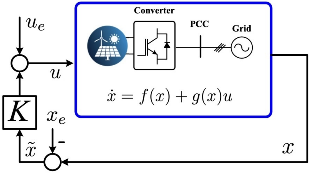
Each DER unit can be controlled as an AC voltage source through an adjustable voltage magnitude and angle at the PCC, $v_s e^{j\omega_s t}$. This is achieved through inner current regulators. Therefore, the control variables v_s and ω_s determine the forced response of the DER unit.

The variables to be controlled are converter current i_c, output current i_o, and the grid voltage frequency δ, which we want to regulate at $(i_c, i_o, \delta) = (15, 10, \omega_g)$.

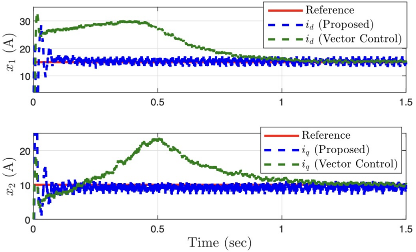


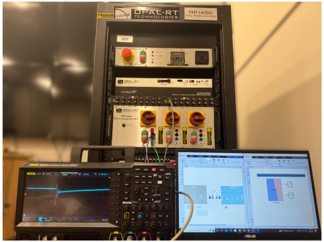
The model-based control scheme involves a feedback–feedforward architecture designed using a control-affine model $\dot{x} = f(x) + g(x)u$ where $x = (i_c, i_o, \delta)$.

$$f(x) = \begin{bmatrix} -\frac{r_g}{L_g}x_1 - \frac{v_g}{L_g}\cos x_3 \\ -\frac{r_g}{L_g}x_2 - \frac{v_g}{L_g}\sin x_3 \\ \omega_g \end{bmatrix}, \quad g(x) = \begin{bmatrix} \frac{1}{L_g} & x_2 \\ 0 & -x_1 \\ 0 & -1 \end{bmatrix}$$



Experimental results obtained with direct data-driven design (CT implementation)[3]





[3] J. Khazaei et al., "Direct data-driven control of grid-connected DERs using Taylor series expansion", IEEE TSE 2024

Summary

- We saw a method to design controllers for nonlinear systems using data
- Handles "functional" uncertainty
- Same positive features as in the linear case (stability guarantees, interpretability, low-complex design programs)
- Merits also depend on Lyapunov's linearization method, which requires no assumptions on the system structure
- Another advantage of Lyapunov's linearization method is that linearity enables robust control tools (also the extension to noise is straightforward)

Selected literature

This result appeared in

- C. De Persis, P. Tesi, “Formulas for data-driven control: Stabilization, optimality, and robustness”, IEEE TAC, 2019
- C. De Persis, P. Tesi, “Learning controllers for nonlinear systems from data”, Annual Reviews in Control, 2022

Many other results have been reported:

Special classes of systems

- A. Bisoffi, C. De Persis, P. Tesi, “Data-based stabilization of unknown bilinear systems with guaranteed basin of attraction”, Systems & Control Letters, 2020
- M. Guo, C. De Persis, P. Tesi, “Data-driven stabilization of nonlinear polynomial systems with noisy data, optimality, and robustness”, IEEE TAC, 2021
- R. Strässer, J. Berberich, F. Allgöwer, “Data-driven control of nonlinear systems: Beyond polynomial dynamics”, IEEE CDC, 2021
- Z. Yuan, J. Cortés, “Data-driven optimal control of bilinear systems”. IEEE L-CSS, 2022
- T. Dai, M. Sznaier,, “A convex optimization approach to synthesizing state feedback data-driven controllers for switched linear systems”, Automatica, 2023

General / control-affine systems

• M. Guo, C. De Persis, P. Tesi, “Data-driven stabilization of nonlinear systems via Taylor’s expansion” Lecture Notes in Control and Information Sciences, vol. 493, 2024
• — “Data-driven stabilizer design and closed-loop analysis of general nonlinear systems via Taylor’s expansion”, arXiv:2209.01071”, 2022
• R. Strässer, M. Schaller, K. Worthmann, J. Berberich, F. Allgöwer, “Koopman-based feedback design with stability guarantees”, arXiv 2312.01441, 2024
• Y. Min, S. Richards, N. Azizan, “Data-driven control with inherent Lyapunov stability”, arXiv 2303.03157, 2023

Stability properties remain local

Somewhere in the middle

Allow for general dynamics but assume that the dynamics belong to a known library of functions (popular also in new sys-ID methods such as SINDy by S. Brunton)

• C. De Persis, M. Rotulo, P. Tesi, “Learning controllers from data via approximate nonlinearity cancellation”, IEEE TAC, 2023

Towards Lecture 4

Consider a control-affine system

$$x(k+1) = f(x) + Bu$$

with $f : \mathbb{R}^n \to \mathbb{R}^n$ and B unknown.

Suppose that we know a library of functions that include f, i.e., a function $Z : \mathbb{R}^n \to \mathbb{R}^s$, $s \geq n$, such that

$$f(x) = AZ(x)$$

for some unknown matrix A.

The dynamics rewrites as

$$x(k+1) = AZ(x) + Bu$$

Knowledge of a library:

- ▷ reasonable in many cases (e.g., mechanical systems)
- ▷ reduces "functional" uncertainty to parametric uncertainty

Consider system

$$x(k+1) = AZ(x) + Bu$$

and arrange Z as

$$Z(x) = \begin{bmatrix} x \\ \hline Q(x) \end{bmatrix}$$

where $Q(x)$ collects all the nonlinearities. The system rewrites as

$$x(k+1) = A_1 x + A_2 Q(x) + Bu$$

A nonlinear control law of the form

$$u(k) = K_1 x(k) + K_2 Q(x)$$

returns

$$x(k+1) = \underbrace{(A_1 + BK_1)x(k)}_{\text{linear}} + \underbrace{(A_2 + BK_2)Q(x)}_{\text{nonlinear}}$$

Can we design K_1 and K_2 so as to stabilize the linear part and minimize the nonlinearity? Lecture 4 will discuss methods for **approximate nonlinearity cancellation**.

Additional material

Data-driven design via Lyapunov's linearization

(noisy case)

The framework

Consider now the system

$$x(k+1) = f(x(k), u(k)) + d(k), \quad k \in \mathbb{N}$$

where $d \in \mathbb{R}^n$ is an unmeasured disturbance.

Unlike the linear case, asymptotic stability does not guarantee any disturbance-to-state property (e.g., GAS $\not\Rightarrow$ ISS), and designing ISS controllers is difficult, even with exact knowledge of the system dynamics.

Useful results can be obtained with Lyapunov's linearization method.

Data-based relation

Taylor's form: $x(k+1) = Ax(k) + Bu(k) + r(x(k), u(k)) + d(k)$.

The data-based relation for the system:

$$\underbrace{\begin{bmatrix} \bar{x}(1) & \bar{x}(2) & \dots & \bar{x}(T) \end{bmatrix}}_{X_1} =$$

$$A \cdot \underbrace{\begin{bmatrix} \bar{x}(0) & \bar{x}(1) & \dots & \bar{x}(T-1) \end{bmatrix}}_{X_0} + B \cdot \underbrace{\begin{bmatrix} \bar{u}(0) & \bar{u}(1) & \dots & \bar{u}(T-1) \end{bmatrix}}_{U_0}$$

$$+ \underbrace{\begin{bmatrix} r(\bar{x}(0), \bar{u}(0)) & r(\bar{x}(1), \bar{u}(1)) & \dots & r(\bar{x}(T-1), \bar{u}(T-1)) \end{bmatrix}}_{R_0}$$

$$+ \underbrace{\begin{bmatrix} \bar{d}(0) & \bar{d}(1) & \dots & \bar{d}(T-1) \end{bmatrix}}_{D_0}$$

In compact form:

$$X_1 = AX_0 + BU_0 + R_0 + D_0$$

Same as before with $R_0 + D_0$ acting as a disturbance matrix.

Main result

Letting $W_0 := R_0 + D_0$ we immediately have the following result.

Theorem Consider a nonlinear system $x^+ = f(x,u) + d$ with an equilibrium $(x_e, u_e) = 0$. Let U_0, X_0, X_1 be data matrices and let $\mathcal{W} = \{W : WW^\top \preceq \Delta\}$, where Δ is chosen by the designer. Finally, let $\Theta \succ 0$ be arbitrary. Suppose there exist $S \in \mathbb{S}^{n\times n}, Y \in \mathbb{R}^{T\times n}$, $\varepsilon > 0$ such that

$$S = X_0 Y, \quad \begin{bmatrix} S - \Theta & (X_1 Y)^\top & Y^\top \\ X_1 Y & S - \varepsilon\Delta & 0 \\ Y & 0 & \varepsilon I \end{bmatrix} \succ 0$$

If $W_0 \in \mathcal{W}$ then $K = U_0 Y S^{-1}$ stabilizes the equilibrium $x_e = 0$.

The closed-loop dynamics is:

$$x^+ = \underbrace{(X_1 - W_0)G}_{\Phi} x + \underbrace{r(x, U_0 G x) + d}_{\rho(x)+d}$$

and we have $\Omega \succ 0$, $P \succ 0$ solving $\Phi^\top P \Phi - P + \Omega \prec 0$.

Letting $V(x) = x^\top P x$ we have

$$\begin{aligned} V(x^+) - V(x) &\leq -x^\top \Omega x + (2\Phi x + \rho(x) + d)^\top P(\rho(x) + d) \\ &\leq -x^\top \Omega x + \underbrace{2(\|PX_1G\| + \|P\|\sqrt{\|\Delta\|}\|G\|)\|x\|(\delta(x) + \gamma) + \|P\|(\delta(x) + \gamma)^2}_{\text{known over-approximation } h(x)} \end{aligned}$$

where $\delta(x)$ is a continuous function such that $\|\rho(x)\| \leq \delta(x|$ for all x, and γ is a positive constant such that $\|d\| \leq \gamma$.

Unlike noise-free case, $h(x)$ can never converge to zero as x goes to zero

$$V(x^+) - V(x) \leq -x^\top \Omega x + \underbrace{2(\|X_1 GP\| + \sqrt{\|\Delta\|}\|GP\|)\|x\|(\delta(x)+\gamma) + \|P\|(\delta(x)+\gamma)^2}_{\text{known over-approximation } h(x)}$$

Definition (Robust invariant set) The set $\mathcal{X}$ is called robust invariant for the system $x(k+1) = F(x(k), d(k))$ if for all $x(0) \in \mathcal{X}$ and $d \in \mathcal{I}$, with $\mathcal{I}$ a compact set, the solution $x(t) \in \mathcal{X}$ for $t > 0$.

Proposition (Robust invariance from data) Let $\mathcal{X} := \{x : -x^\top \Omega x + h(x) \leq 0\}$ and consider any Lyapunov sub-level set $\mathcal{R}_\alpha := \{x : V(x) \leq \alpha\}$. Let $\mathcal{Z} := \mathcal{R}_\alpha \cap \mathcal{X}^c$, where $\mathcal{X}^c$ is the complement of $\mathcal{X}$ ($\mathcal{Z}$ is the subset of $\mathcal{R}_\alpha$ for which the Lyapunov difference can be positive). If

$$V(x) - x^\top \Omega x + h(x) \leq \alpha \quad \forall x \in \mathcal{Z}$$

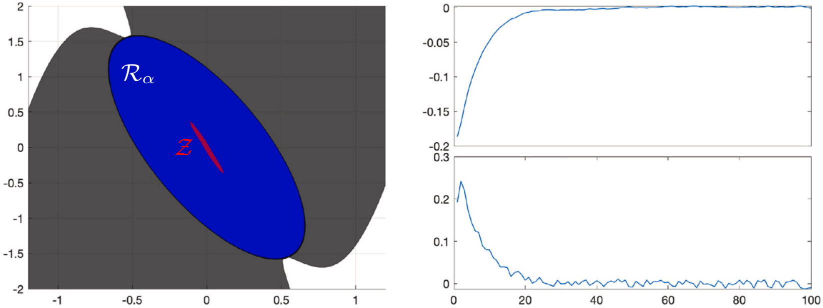
then $\mathcal{R}_\alpha$ is a robust invariant set for the closed-loop system.

Inverted pendulum (cont'd)

Dynamics:

$$\begin{bmatrix} x_1^+ \\ x_2^+ \end{bmatrix} = \begin{bmatrix} 1 & T_s \\ \frac{T_s g}{l} & 1 - \frac{T_s \mu}{ml^2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{T_s}{ml^2} \end{bmatrix} u + \begin{bmatrix} 0 \\ \frac{T_s g}{l} \end{bmatrix} (\sin x_1 - x_1) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} d$$

Same experimental setting as before, now with $|d| \leq 0.01$.
Controller obtained $K = \begin{bmatrix} -21.9778 & -9.6747 \end{bmatrix}$.
Estimate obtained using $\delta(x) = 2\|\sin x_1 - x_1\|$ (> 100% over-approximation)



Continuous-time results

Continuous-time results

Discrete-time results can be extended to continuous-time systems with X_1 matrix of <u>state derivatives</u> (adding noise if necessary).

Theorem Consider the nonlinear system $\dot{x} = f(x, u)$ with an equilibrium $(x_e, u_e) = 0$. Let U_0, X_0, X_1 be data matrices and let $\mathcal{R} = \left\{R : \; RR^\top \preceq \Delta\right\}$, where Δ is chosen by the designer. Finally, let $\Theta \succ 0$ be arbitrary. Suppose there exist $S \in \mathbb{S}^{n \times n}, Y \in \mathbb{R}^{T \times n}$, $\varepsilon > 0$ such that

$$S = X_0 Y, \quad \begin{bmatrix} (X_1 Y) + (X_1 Y)^\top + \Theta + \varepsilon \Delta & Y^\top \\ Y & -\varepsilon I \end{bmatrix} \prec 0$$

If $R_0 \in \mathcal{R}$ then $u = Kx$ with $K = U_0 Y S^{-1}$ stabilizes the equilibrium $x_e = 0$.