ELE6214 - Commande de Systèmes Incertains

Lecture 1: Introduction to Uncertainty and Adaptive Control

Bowen Yi
bowen.yi@polymtl.ca

Polytechnique Montréal

Hiver 2025

- La section de Automation et systèmes
- Département de Génie Électrique
- Office: M-5412
- Email: bowen.yi@polymtl.ca
- Phone: +1 (514) 340 4711 ext. 3908

- Research interests:
    - Nonlinear systems (estimation, learning and control)
    - Robotics

Outline

Course Objectives

- Understand uncertainties in control systems
- Teach the theory and practice of the mainstream techniques, to design control systems with uncertainties, in
  - Adaptive control
  - Online parameter estimation/learning
  - Robust control
- Prerequisites:
  - ELE6202 Multivariable Systems (or equivalents)

Evaluation

- Project: 2 graded project reports (35% each)
- Oral Exam/Presentation: 30%
- Grading:
  - For late submission: -10% if late for each day
  - Submit individually;
  - Not allowed to share the final reports or detailed methods
  - Generative AI tools are allowed.

Schedule

Adaptive Control (Sastry & Bodson, Ch. 1,2,3,5)

- Lecture 1: Introduction to Uncertainty & Adaptive Control (3h)
- Lecture 2: Real-Time Parameter Estimation (3h)
- Lecture 3: Online System Identification (2h)
- Lecture 4: Model Reference Adaptive Control (6h)
- Lecture 5: Robustness of Adaptive Systems (2h)

Robust Control (Scherer's notes)

- Lecture 6: Robustness for SISO Systems
- Lecture 7: Stabilizing Controllers, Generalized Plant Concept
- Lecture 8: Robust Stability Analysis
- Lecture 9: Nominal and Robust Performance Analysis
- Lecture 10: Synthesis of H_∞ Controllers

Advanced Topics (TBD, if time allows) & Presentation

References

1. S. Sastry and M. Bodson, *Adaptive Control: Stability, Convergence, and Robustness*, 1989[a]
2. K.J. Åström and B. Wittenmark, *Adaptive Control*, 2nd Ed., 2008.
3. G.C. Goodwin and K.S. Sin, *Adaptive Filtering Prediction and Control*, 2014. (discrete-time systems)
4. C. Scherer, *Theory of Robust Control*, University of Stuttgart.[b]
5. J.C. Doyle, B.A. Francis, and A. Tannenbaum, *Feedback Control Theory*, 1992.
6. K. Zhou and J.C. Doyle, *Essentials of Robust Control*, 1998.[c]

[a] https://my.ece.utah.edu/~bodson/acscr/acscr.pdf
[b] https://www.imng.uni-stuttgart.de/mst/files/RC.pdf
[c] https://www.ece.lsu.edu/kemin/essentials.htm

Outline

Systems and Signals

Linear Time-Invariant, Finite-Dimensional Systems

$$\begin{aligned} \dot{x} &= Ax + Bu, \ x(0) = x_0 \\ y &= Cx + Du \end{aligned} \tag{1}$$

with state $u \in \mathbb{R}^n$, input $u \in \mathbb{R}^m$, and output $y \in \mathbb{R}^q$.

- x, y and u are signals: functions of time $t \in [0, \infty)$ that are piece-wise continuous.
- Notations: $x(\cdot)$ denotes the signal as a whole, and $x(t)$ is the value of the signal at t.
- Solution:

$$y(t) = Ce^{At}x_0 + \int_0^t Ce^{A(t-s)}B(s)u(s)ds + Du(t), \ \forall t \geq 0.$$

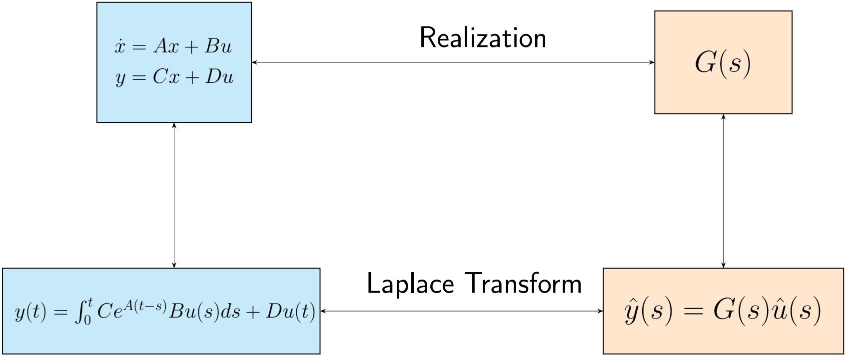
- The system has the transfer matrix/function $G(s)$ defined as

$$G(s) = C(sI - A)^{-1}B + D, \tag{2}$$

which is a matrix whose elements consist of real-rational and proper functions in s.
- The fundamental relation between the state-space model and frequency domain representation is studied in realization theory.

$$G(s) = C_G(sI - A_G)^{-1}B_G + D_G,$$

A feasible realization (A_G, B_G, C_G, D_G) and minimal realization (controllable & observable).
- View the system as a device that processes signals from input $u(\cdot)$ to output $y(\cdot)$.



We use the symbol (NOT a partitioned matrix!)

$$G = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right], \qquad \begin{bmatrix} \dot{x} \\ y \end{bmatrix} = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \begin{bmatrix} x \\ u \end{bmatrix}$$

both for the mapping $u \to y$ as defined via the differential equation with initial condition 0, and for the corresponding transfer matrix G.

Operations for Realization

Suppose we are given

$$G = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right], \quad G_1 = \left[\begin{array}{c|c} A_1 & B_1 \\ \hline C_1 & D_1 \end{array}\right], \quad G_2 = \left[\begin{array}{c|c} A_2 & B_2 \\ \hline C_2 & D_2 \end{array}\right].$$

- If G_1 and G_2 have the same dimension, their sum has a realization

$$G_1 + G_2 = \left[\begin{array}{cc|c} A_1 & 0 & B_1 \\ 0 & A_2 & B_2 \\ \hline C_1 & C_2 & D_1 + D_2 \end{array}\right].$$

- If the number of columns of G_1 is equal to the number of rows G_2, their product has a realization

$$G_1 G_2 = \left[\begin{array}{cc|c} A_1 & B_1 C_2 & B_1 D_2 \\ 0 & A_2 & B_2 \\ \hline C_1 & D_1 C_2 & D_1 D_2 \end{array}\right].$$

Operations for Realization (cont'd)

- If D is invertible, then G^{-1} exists, is proper and has a realization

$$G^{-1} = \left[\begin{array}{c|c} A - BD^{-1}C & BD^{-1} \\ \hline -D^{-1}C & D^{-1} \end{array}\right]$$

- Suppose that the *square* transfer function G has a proper inverse. Then, $G(\infty)$ is invertible.

A square matrix has a proper inverse $\iff$ $G(\infty)$ is invertible.

Controllability & Observability

Recall the PBH test:

- (A, B) is controllable $\iff$ the full rank of

$$\begin{bmatrix} A - \lambda I & B \end{bmatrix}, \quad \forall \lambda \in \mathbb{C}$$

- (A, C) is observable $\iff$ the full rank of

$$\begin{bmatrix} A - \lambda I \\ C \end{bmatrix}, \quad \forall \lambda \in \mathbb{C}$$

- (A, B) is stabilizable $\iff$ the full rank of

$$\begin{bmatrix} A - \lambda I & B \end{bmatrix}, \quad \forall \lambda \in \mathbb{C}_{\geq 0}$$

- (A, C) is detectable $\iff$ the full rank of

$$\begin{bmatrix} A - \lambda I \\ C \end{bmatrix}, \quad \forall \lambda \in \mathbb{C}_{\geq 0}$$

Stability of LTI Systems (Frequency Domain)

A transfer function matrix $H(s)$, whose elements are real rational functions, are stable if

1) $H(s)$ is proper (i.e. no pole at ∞)[a]; and
2) $H(s)$ has only poles in $\mathbb{C}_{<0}$.

[a] Equivalently, the degree of the numerator $\leq$ the degree of the denominator.

Remarks

- Strictly proper: replace $\leq$ by $<$, equivalently, $\lim_{|s|\to\infty} H(s) = 0$.
- Engineering meaning: pure differentiator $H(s) = s$ (not proper)

RH_∞ Space

For the set of real rational proper and stable matrices of dimension $k \times m$ we use the symobol

$$RH_\infty^{k\times m} \quad \text{or} \quad RH_\infty,$$

latter when the dimensionality is clear.

Close under 3 operations

- A scalar multiple of one stable transfer matrix
- Sum of two stable transfer function matrices
- Product of two stable transfer function matrices

Quizz

1. Which of the transfer functions are proper?

$$g_1(s) = \frac{s+1}{s-1}, \quad g_2(s) = \frac{s}{s+1} - \frac{s^5}{s^4-1}$$

2. What is the value of $g_1(s)$ at infinity?
3. What is the value of $C(sI-A)^{-1}B+D$ at infinity?
4. Which of the following rational matrices have a proper inverse:

$$\begin{bmatrix} \frac{1}{s} & \frac{1}{s^2} \\ s^2 & s \end{bmatrix}, \quad \begin{bmatrix} \frac{1}{s} & \frac{1}{s^2} \\ \frac{1}{s^2} & \frac{1}{s} \end{bmatrix}, \quad \begin{bmatrix} 1 & \frac{1}{s} \\ \frac{1}{s} & 1 \end{bmatrix}$$

Stability of LTI Systems (Time Domain)

A state space model (A, B, C, D) is said to be stable if

$$\forall \lambda_i\{A\} \in \mathbb{C}_{<0}.$$

Relation (Frequency and Time Domains)

The state space model $\dot{x} = Ax + Bu,\ y = Cx + Du$ and the corresponding transfer function matrix $G(s) = C(sI - A)^{-1}B + D$ have the following relations:

- If the state space model is stable, then $G(s)$ is stable.
- Conversely, if $G(s)$ is stable, (A, B) is stablizable and (A, C) is detectable, then the matrix A is stable.

Summary of System Descriptions in this Course

- A quadruple (A, B, C, D) of matrices defines the state-space system [time domain]

$$\dot{x} = Ax + Bu,\ y = Cx + Du, \quad x(0) = x_0$$

which is considered as a map from $u(\cdot)$ to $y(\cdot)$.

- The quadruple (A, B, C, D) is also expressed as [time domain]

$$y = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] u$$

- In frequency domain

$$\hat{y}(s) = G(s)u(s), \quad G(s) := C(sI - A)^{-1}B + D$$

- An operator $\hat{G}$ with any realization (A, B, C, D) s.t.

$$y(t) = \hat{G}[u(t)].$$

Useful Norms for Signals & Transfer Functions

- Bounded vector-valued signal $u(\cdot)$ (maximal amplitude/peak):

$$\|u\|_\infty = \sup_{t \geq 0} \|u(t)\| < \infty$$

- Integral signal $u(\cdot)$ (energy):[1]

$$\|x\|_2 = \sqrt{\int_0^\infty \|x(t)\|^2 dt}$$

BIBO stability is related to $\|u\|_2 < \infty \implies \|y\|_2 < \infty$. Intuitively, we consider the energy-to-energy gain

$$\gamma_{\texttt{energy}} = \sup_{0<\|u\|_2<\infty} \frac{\|y\|_2}{\|u\|_2} \underset{\text{can prove}}{=} \sigma_{\texttt{max}}(G(j\omega)) := \|G(j\omega)\|$$

- For a stable transfer function matrix

$$\|G\|_\infty = \sup_{\omega \in \mathbb{R}} \|G(j\omega)\|.$$

[1] A signal with a large energy can have a small peak and vice versa.

Outline

Uncertainty

Differences always between the actual system and the model

- Unknown parameters and parameter variations

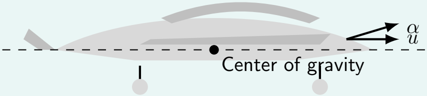
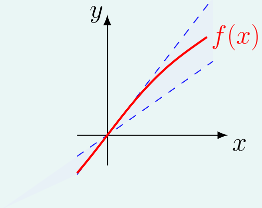


Figure: Aircraft: Inertial change due to fuel consumption

Takeoff and landing weights for a Boeing 777-300ER from Montreal (YUL) to Paris (CDG) [Generated by ChatGPT]

Parameter	Value
Distance	5,550 km
Takeoff Weight (TOW)	300,000 kg
Landing Weight (LW)	240,000 kg
Fuel Burned	60,000–70,000 kg

- Unmodeled nonlinearities in the linear model



- Synthesized controller may be different from implemented controller (Simplification for control implementation)

- Detailed model → low-order simpler model for control synthesis

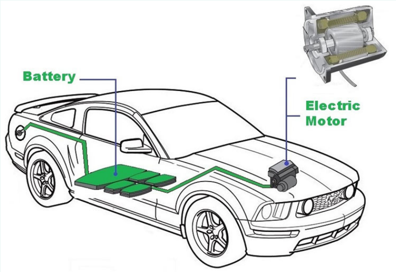


Figure: Motor driven Electric Vehicles (Input in model: Torque via ignoring motor dynamics; Actual input: Voltage)

- Unmodeled dynamics: at high frequency both structure and order of model are unknown.

"A control engineer calls this mismatch uncertainty. Note that this is an abuse of notation since neither the system nor the model are uncertain; it is rather our knowledge about the actual physical system that we could call uncertain."

— Carsten Scherer

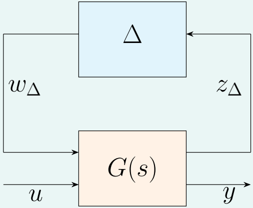
Classes of Uncertainty

- *Parametric (real) uncertainty.*[a] The structure of the model (including the order) is known, but some of of parameters $\theta = \left[\theta_1, \ldots, \theta_\ell\right]^\top$ are unknown or uncertain.

$$\dot{x} = A(\theta)x + B(\theta)u$$
$$y = C(\theta)x + D(\theta)u$$

The vector θ is unknown and possibly time-varying.

- *Dynamic (complex/frequency-dependent) uncertainty.* The model is in error due to missing dynamics, usually at high frequencies.



$$\dot{x} = Ax + B_1 w_\Delta + B_2 u$$
$$z_\Delta = C_1 x + D_{11} w_\Delta + D_{12} u$$
$$y = C_2 x + D_{21} w_\Delta$$
$$w_\Delta = \Delta z_\Delta$$

[a] Petersen & Tempo, *Automatica*, 2014.

Classes of Uncertainty (cont'd)

Dynamic uncertainty can be further classified into:

D1 *Unstructured uncertainty*. Roughly, a single constraint on

$$w_\Delta = \Delta(x,t)z_\Delta$$

D1.1 Norm bounded uncertainty:

$$\|\Delta(x(t),t))\|_\infty \leq 1$$

D1.2 Bounded real uncertainty: the transfer function matrix $\Delta(s)$ satisfies the bounded real condition $\|\Delta(s)\|_\infty \leq 1$, e.g.

- Addictive, multiplicative, and normalized coprime factor uncertainty

D1.3 Positive real uncertainty: $\Delta(s)$ s.t.

$$\Delta(j\omega) + \Delta(j\omega)^* \succeq 0,\ \forall\omega$$

D1.4 Negative imaginary uncertainty:

$$j(\Delta(j\omega) - \Delta(j\omega)^\star) \succeq 0,\ \forall\omega \geq 0$$

Classes of Uncertainty (cont'd)

D2 *Structured uncertainty*. Roughly, multiple constraints on uncertainty.

D2.1 Structured singular values uncertainty

$$\Delta = \begin{bmatrix} \Delta_1 & & \\ & \ddots & \\ & & \Delta_b \end{bmatrix}$$

D2.2 Integral quadratic constraint (IQC) constraint in time-domain:

$$\int_0^T |w_\Delta(s)|^2 dt \leq \int_0^T |z_\Delta(t)|^2 dt + d$$

D2.3 IQC constraint in frequency-domain:

$$\int_{-\infty}^{\infty} \begin{bmatrix} W_\Delta(j\omega) \\ Z_\Delta(j\omega) \end{bmatrix}^* \Pi(j\omega) \begin{bmatrix} W_\Delta(j\omega) \\ Z_\Delta(j\omega) \end{bmatrix} d\omega \geq 0$$

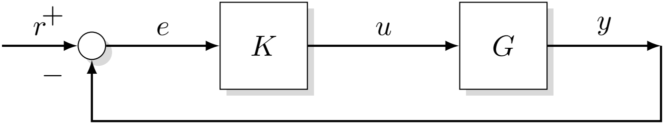
More details for dynamic uncertainty in the second part of the course.

Standard Regulation & Tracking Problems

For the given plant G (possibly with uncertainty), consider the feedback interconnection

$$y = Gu, \quad u = K(r - y)$$

with the controller K.



Standard goals in designing K

- Stabilize the interconnection
- Output y tracks r well, i.e. the norm $\|y - r\|$ is small enough
- Control action u should not be too large

Constant r: regulation; Time-varying $r(t)$: tracking.

Our approaches

This course covers both parametric and dynamic (possibly structured or not) uncertainty:

- Parametric uncertainty: Adaptive Control ("Adapt on the fly")
- Dynamic uncertainty: Robust Control ("One design fits all")

Outline

Introduction

Tune to adjust for proper response[2]

Adapt to adjust to a specific use or situation

Autonomous independence, self-governing

Learn to acquire knowledge or skill by study, instruction or experience

Reason the intellectual process of seeking truth or knowledge by inferring from either fact of logic

Intelligence the capacity to acquire and apply knowledge

In Automatic Control

- Gain scheduling - adjust controller parameter based on direct measurement of system and environmental parameters
- Automatic tuning - tuning on demand
- Adaptation - continuous adjustment of controller parameters based on regular measured signals

[2] The introduction part is mainly from Karl J. Åström's lecture.

Brief History of Adaptive Control

- Adaptive control: learn enough about a plant/process and its environment for control – restricted domain, prior info
- Development similar to neural networks
  Many ups and downs, lots of strong egos
- Early work driven adaptive flight control 1950–1970.
  The brave era: Develop an idea, hack a system, simulate and fly!
  Several adaptive schemes emerged no analysis
  Disasters in flight tests – the X-15 crash Nov 15 1967
  Gregory ed, Proc. Self Adaptive Flight Control Systems, 1959.
- Emergence of adaptive theory 1970–1980
  Model reference adaptive control emerged from flight control stability theory – a tracking problem
  The self tuning regulator emerged from process control and stochastic control theory – a regulation problem
- Microprocessor based products 1980
- Robust adaptive control 1990
- Machine Learning and Adaptation 2020

Publications in Scopus

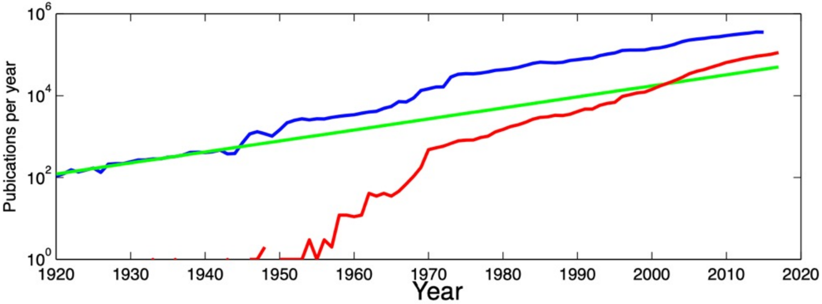
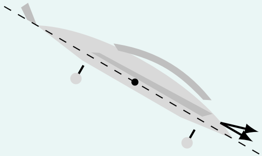
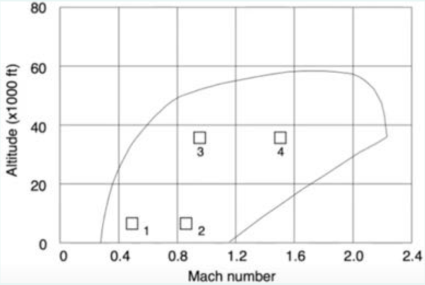


Figure: Control vs Adaptive Control

Pitch Control of Aircraft





Eigenvalues of dynamics matrix

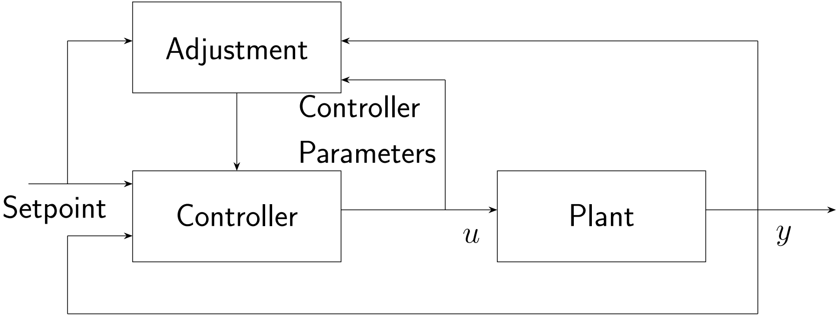
FC1: -14, -3.07, 1.23

FC2: -14, -4.90, 1.78

FC3: -14, -1.87, 0.56

FC4: -14, $-0.83 \pm 4.3i$

Block Diagram of Adaptive Systems



Intuition

- To obtain a progressively better understanding of the plant (for control), we need an identification technique.
- Intuitive to aggregate system identification and control
- If system identification is recursive – models are periodically updated using previous estimates and new data – identification and control may be performed concurrently.

Roughly speaking,

$$\text{Adaptive Control} = \\ \text{(non-adaptive) Control Scheme} + \text{Recursive System Identification}$$

Some Landmarks

- Early flight control systems 1955
- Dynamic programming Bellman 1957
- Kalman's self-optimizing regulator 1958
- Dual control Feldbaum 1960
- System identification 1965
- Self-optimizing control Draper Li 1966
- Learning control Tsypkin 1971
- Algorithms MRAS STR 1970
- Stability analysis (Lyapunov, passivity) 1980
- Industrial product 1980
- PID auto-tuning 1982
- Robustness 1985
- Autonomous control 1995
- Adaptation and machine learning – a renaissance 2015

A Simple Example (MIT Rule)

Consider a first-order LTI system

$$y = G(s)[u] = \frac{k}{s+a}[u]$$

with known $a > 0$ and *unknown* $k > 0$. Our target is a design a feedback to make the closed-loop satisfy the model

$$y_m = M(s)[u] = \frac{1}{s+a}[u]$$

If k *was* known, we use the proportional control with gain $\theta_\star = \frac{1}{k}$.

In MIT rule, we design the output error $e(\theta) = y_m - y(\theta)$ and optimize the cost function $J(\theta) = \frac{1}{2}e^2$. It yields $\frac{\partial y(\theta)}{\partial \theta} = \partial(\frac{k}{s+a})\theta[u]/\partial\theta = k y_m$ and the gradient of J is $\frac{\partial J}{\partial \theta} = -key_m$. Hence, we select the gradient dynamics

$$\dot{\theta} = -\gamma e y_m, \quad \gamma > 0 \text{ (Adaptation gain)}$$

How can we get proven properties (stability, convergence, and beyond)?

Outline

Barbalat's Lemma

If $f(t)$ is a uniformly continuous function, s.t.

$$\lim_{t\to\infty}\int_0^t f(s)ds < \infty,$$

then $f(t) \to 0$ as $t \to \infty$.

Corollary

If

$$g, \dot{g} \in L_\infty, \text{ and } g \in L_p$$

form some $p \in [1, \infty)$, then $g(t) \to 0$ as $t \to \infty$.

- Norm: $\|u\|_p = (\int_0^\infty |u(s)|^p ds)^{\frac{1}{p}}$ and $\|u\|_\infty = \sup_{t\geq 0}|u(t)|$.
- Space $L_p := \{f(t) | \|f\|_p < \infty\}$.

Differential Equations

The system

$$\dot{x} = f(x,t), \quad x(t_0) = x_0 \in \mathbb{R}^n$$

is said to be

- autonomous or time-invariant, if f does not depend on t;
- time-varying, otherwise.
- linear, if $f(x,t) = A(t)x$;
- nonlinear, otherwise.
- has an *equilibrium* point $x_\star$, if $f(x_\star, t) \equiv 0$.

Lipschitz Condition

The function f is *Lipschitz* in x, if for some $h > 0$, $\exists \ell \geq 0$ s.t.

$$|f(x_1,t) - f(x_2,t)| \leq \ell |x_1 - x_2|, \; \forall x_1, x_2 \in B_h, \; t \geq 0.$$

The Lipschitz constant $\ell \implies$ existence and uniqueness.

Lemma (Bellman-Gronwall)

Let $x(\cdot), a(\cdot), u(\cdot) : \mathbb{R}_+ \to \mathbb{R}_+$, and $T \geq 0$. If

$$x(t) \leq \int_0^t a(\tau)x(\tau)d\tau + u(t),\ \forall t \in [0, T], \tag{3}$$

then

$$x(t) \leq \int_0^t a(\tau)u(\tau)\exp\left(\int_\tau^t a(\sigma)d\sigma\right)d\tau + u(t),\ \forall t \in [0, T]. \tag{4}$$

When $u(\cdot) \in C^1$,

$$x(t) \leq u(0)\exp\left(\int_0^t a(\sigma)d\sigma\right) + \int_0^t \dot{u}(\tau)\exp\left(\int_\tau^t a(\sigma)d\sigma\right)d\tau$$

Stability

The equilibrium $x = 0$ for $\dot{x} = f(x,t)$ is

- stable, if $\forall t_0 \geq 0$ and $\epsilon > 0$, $\exists \delta(t_0, \epsilon)$ s.t.

$$|x_0| < \delta(t_0, \epsilon) \implies |x(t)| < \epsilon, \; \forall t \geq t_0.$$

- uniformly stable, if $x = 0$ is stable and δ is independent of t_0.
- asymptotically stable, if $x = 0$ is stable and attractive, i.e. $\forall t_0$, $\exists \delta(t_0)$ s.t.

$$|x_0| < \delta \implies \lim_{t \to \infty} |x(t)| = 0.$$

- uniformly asymptotically stable (UAS), if $x = 0$ is uniformly stable and $x(t)$ converges to 0 uniformly in t_0. I.e., $\exists \delta > 0$ and a function $\gamma : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}_+$, s.t. $\lim_{\tau \to \infty}(\tau, x_0) = 0$ for all x_0 &

$$|x_0| < \delta \implies |x(t)| \leq \gamma(t - t_0, x_0), \; \forall t_0 \geq 0.$$

The equilibrium $x = 0$ for $\dot{x} = f(x, t)$ is

- Globally asymptotically stable (GAS), if $x = 0$ is asymptotically stable and $\lim_{t\to\infty} |x(t)| = 0$, for all $x_0 \in \mathbb{R}^n$
- Uniformly globally asymptotically stable (UGAS) ...
- Exponentially stable, if $\exists m, \alpha > 0$ s.t.

$$|x(t)| \leq m e^{-\alpha(t-t_0)} |x_0|, \quad \forall x_0 \in B_h,\ t \geq t_0 \geq 0$$

  and the constant α is called as the rate of convergence.
- Globally exponentially stable (GES) ...

Comparison Functions

- Class $\mathcal{K}$ function: A function $\alpha : \mathbb{R}_+ \to \mathbb{R}_+$ belongs to class $\mathcal{K}$, denoted as $\alpha \in \mathcal{K}$, if it is continuous, *strictly increasing*, and $\alpha(0) = 0$.
- Class $\mathcal{K}_\infty$ function: It is said to belong to class $\mathcal{K}_\infty$ if $\lim_{r\to\infty} \alpha(r) = \infty$.
- Class $\mathcal{KL}$ function: A continuous function $\beta : [0, a) \times [0, \infty)$ is said to belong to class $\mathcal{KL}$ if
  1. for each fixed s, $\beta(r, s) \in \mathcal{K}$ w.r.t. r;
  2. for each fixed r, $\beta(r, s)$ is *decreasing* w.r.t. s and $\lim_{s\to\infty}(r, s) = 0$.

Positive Definite Function, Decrescent Function

A continuous function $V(x,t) : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}$ is called

- locally positive definite function (l.p.d.f)[a] if, for some $h > 0$ and some $\alpha(\cdot) \in \mathcal{K}$

$$V(0,t) = 0 \text{ and } V(x,t) \geq \alpha(|x|), \quad \forall x \in B_h,\ t \geq 0.$$

- positive definite function (p.d.f), by replacing the above "$\forall x \in B_h$" by "$\forall x \in \mathbb{R}^n$".
- decrescent, if $\exists \beta(\cdot) \in \mathcal{K}$ s.t.

$$V(x,t) \leq \beta(|x|),\ \forall x \in B_h,\ t \geq 0.$$

[a] Imagine like an "energy function".

Example

Consider the functions below.

- $V(x,t) = |x|^2$: p.d.f., decrescent
- $V(x,t) = x^\top P x$ with $P \succ 0$: p.d.f., decrescent
- $V(x,t) = (t+1)|x|^2$: p.d.f.
- $V(x,t) = e^{-t}|x|^2$: decrescent
- $V(x,t) = \sin^2(|x|^2)$: l.p.d.f., decrescent

Lyapunov Stability Theorems

Consider the system $\dot{x} = f(x,t)$ and a candidate Lyapunov function $V(x,t) \in C^1$ with

$$\dot{V}(x,t) = \frac{\partial V(x,t)}{\partial t} + \frac{\partial V(x,t)}{\partial x} f(x,t).$$

Then,

Conditions on $V(x,t)$	Conditions on $-\dot{V}(x,t)$	Conclusions
l.p.d.f.	≥ 0 locally	stable
l.p.d.f., decrescent	≥ 0 locally	US
l.p.d.f.	l.p.d.f.	AS
l.p.d.f., decrescent	l.p.d.f.	UAS
p.d.f., decrescent	p.d.f.	UGAS
$a_1 \lvert x \rvert^2 \leq V(x,t) \leq a_2 \lvert x \rvert^2$ $\lvert \partial V / \partial x \rvert \leq \alpha_4 \lvert x \rvert$	$\leq -a_3 \lvert x \rvert^2$	GES

Linear Time-Varying (LTV) Systems

Consider the LTV system

$$\dot{x} = A(t)x_0, \quad x(t_0) = x_0,$$

whose solution satisfies

$$x(t) = \Phi(t, t_0)x_0.$$

The state transition matrix $\Phi(t, t_0) \in \mathbb{R}^{n \times n}$ is the unique solution to

$$\frac{d}{dt}\Phi(t, t_0) = A(t)\Phi(t, t_0), \quad \Phi(t_0, t_0) = I,$$

and satisfies the semigroup property

$$\Phi(t, t_0) = \Phi(t, \tau)\Phi(\tau, t_0), \quad \forall t \geq \tau \geq t_0,$$

thus $\Phi(t, t_0)^{-1} = \Phi(t_0, t)$.

LTV Stability

For the LTV system $\dot{x} = A(t)x$,

- $x = 0$ is UAS $\iff$ $x = 0$ is exponentially stable.
- $x = 0$ is E.S. $\iff$ $\exists$ some $m, \alpha > 0$ s.t.

$$\|\Phi(t, t_0)\| = m \exp(-\alpha(t - t_0))$$

  for all $t \geq t_0 \geq 0$.

Uniformly Complete Observability (UCO)

The system

$$\dot{x} = A(t)x, \quad y = C(t)x$$

is called UCO if $\exists$ strictly positive constants β_1, β_2, δ s.t., $\forall t \geq 0$

$$\beta_2 I \geq W(t, t + \delta) \geq \beta_1 I$$

with observability gramian $W = \int_t^{t+\delta} \Phi^\top(\tau, t) C^\top(\tau) C(\tau) \Phi(\tau, t) d\tau$.

What have we learned today?

- Review the representation of LTI systems (state-space & transfer matrix)
- Sources and classification of uncertainty
- Introduction to adaptive control
- Review the elements of stability theory

Homework

1. Read (Sastry & Bodson, Chapter 1) - Preliminaries.
2. If you have time, also read (Sastry & Bodson, Chapter 0).

ELE6214 - Commande de Systèmes Incertains

Lecture 2: Real-Time Parameter Estimation

Bowen Yi

bowen.yi@polymtl.ca

Polytechnique Montréal

Hiver 2025

POLYTECHNIQUE MONTRÉAL

UNIVERSITÉ D'INGÉNIERIE

Outline

Wide Applications

- Mathematics
- Statistics – Estimation theory
- Biology – Medical statistics
- Economics – Econometrics
- **Control – System identification**
- Signal processing
- Numerical analysis
- Physics
- . . .

System Identification: The Control View[1]

- How to get the models
  - Physics - white boxes
  - Experiments - black boxes
  - Combination - grey boxes
- How to do the experiments (data)
  - Experimental conditions - excitation
  - Fit and validation sets
- Model structure
  - Transfer functions
  - Impulse responses
  - State models
- Parameter estimation
  - Statistics
  - Loss function - likelihood
  - Validation
- Adaptive control
  - Estimate and control *simultaneously*

[1] This part is from K.J. Astrom's slides.

System Identification

Non-parametric methods

- Bode or Nyquist diagrams
- Step or impulse responses

Parametric methods

- Transfer functions
- Sampled models

  Output error OE: $A(q)[x] = B(q)[u], \; y = x + e$ with $q := \frac{d}{dt}$
  Transfer function: $H(s) = \frac{N(s)}{M(s)}$
  Moving average
  Autoregressive: $A(q)[y] = e$
  Autoregressive with external input: $A(q)[y] = B(q)[u] + e$
  Autoregressive moving average with external input (ARMAX):

$$A(q)y(k) = B(q)u(k) + C(q)e(k)$$

- Methods suitable for adaptive control

Estimation Theory

- Special branch of statistics
- Unknowns parameters θ, observations y
- The likelihood function $p(\theta|Y)$ is the probability density of the observations y given the parameters θ
- Log likelihood function $L(\theta|Y) = \log p(\theta|Y)$
- Consistency - parameters converge in probability as sample size goes to infinity
- Cramér-Rao lower bound[2]
- Efficiency - estimate achieves the Cramér-Rao bound when sample size goes to infinity[3]

---

[2] It relates to estimation of deterministic but unknown parameter θ, i.e. the precision of any unbiased estimator is $\leq$ the Fisher information $I(\theta)$.

[3] Harald Cramér, *Mathematical Methods of Statistics*, 1946.

System Identification

Classic 1955 –

- Step or impulse responses
- Frequency response - transfer functions
- Spectrum analyzers - measure transfer function directly

Adaptive control 1959 –

- Estimate parameters in real time
- IFAC Symposium on Adaptive Control Teddington 1965

Identification 1965 –

- State space models
- Sampled models
- Strongly influenced by statistics

Maximum Likelihood Estimation

Let θ be the unknown parameters and Y all observed measurements. The likelihood function $L(\theta|Y)$ is the probability density function of the observations Y given the parameters

$$L(\theta|Y) = p(Y|\theta).$$

It is useful to deal with the *loglikelihood* function $L(\theta|Y) = \log p(Y|\theta)$. The *maximum likelihood* estimate is

$$\hat{\theta} = \operatorname*{argmin}_{\theta} L(\theta|Y).$$

The Fisher information matrix I has the i, j-th element

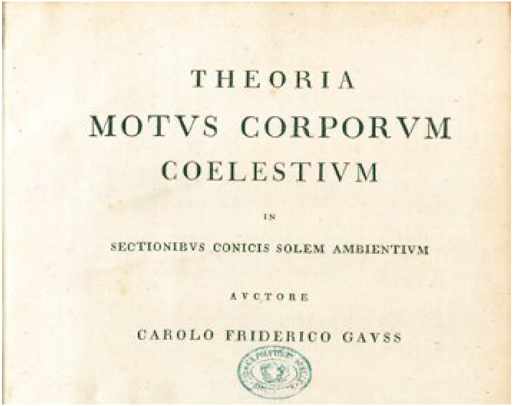
$$I_{i,j} = -E\frac{\partial^2 L(\theta|Y)}{\partial\theta_i \partial\theta_j},$$

which is a lower bound of the covariance of the estimate.

In this course, we focus on *deterministic* approaches.

Outline

Date Back to Gauss Original Work 1809



Priority dispute: in the history of statistics is that between Gauss and Legendre, over the discovery of the method of least squares.

Regressor Model

$$y(t) = H(\phi(t), \theta) + e(t)$$

$y(t) \in \mathbb{R}^m$ - observed data

$\theta \in \mathbb{R}^n$ - unknown *constant* parameters

ϕ (proper dim) - known function

$e \in \mathbb{R}^n$ - residuals (small)

Linear Regressor

$$\begin{aligned} y(t) &= \phi^\top(t)\theta + e(t) \\ &= \phi_1(t)\theta_1 + \ldots + \phi_n(t)\theta_n + e(t). \end{aligned}$$

with $\phi(t) \in \mathbb{R}^{n \times m}$.[a]

[a] It can be viewed as the LTV system

$$\dot{\theta} = 0, \quad y(t) = \phi^\top(t)\theta.$$

Offline Solution

Consider the linear regressor

$$y(t) = \phi^\top(t)\theta,$$

and we defined the collection vectors

$$Y(t) = \text{col}(y(1), \ldots, y(t)),\ \Psi(t) = \text{col}(\phi^\top(1), \ldots, \phi^\top(t))$$
$$E(t) = \text{col}(e(1), \ldots, e(t)).$$

Using the historical information, we have $Y(t) = \Psi(t)\theta + E(t)$. Our target is to minimize, with respect to θ, the cost function

$$V(\theta, t) = \frac{1}{2}\sum_{i=1}^{t} e(i)^2 = \frac{1}{2}\sum_{i=1}^{t}[y(i) - \phi^\top(i)\theta]^2 = \frac{1}{2}\|E\|^2,$$

where

$$E = Y - \hat{Y} = Y - \Psi\theta.$$

Theorem (Least-Squares Theorem)

The parameter $\hat{\theta}_\star$ that minimizes the cost function are given by the *normal equations*

$$\Psi^\top \Psi \hat{\theta}_\star = \Psi^\top Y. \tag{1}$$

If $\Phi^\top \Phi$ is nonsingular, the miniumum is unique and given by

$$\hat{\theta}_\star = (\Psi^\top \Psi)^{-1} \Psi^\top Y.$$

This is a well-known result in any textbooks on linear algebra or matrix theory.

Proof.

The cost function can be written as

$$\begin{aligned} 2V(\theta, t) &= E^\top E = (Y - \Psi\theta)^\top (Y - \Psi\theta) \\ &= Y^\top Y - Y^\top \Psi\theta - \theta^\top \Psi^\top Y + \theta^\top \Psi^\top \Psi\theta. \end{aligned}$$

Complete the square

$$\begin{aligned} 2V(\theta, t) &= Y^\top Y - Y^\top \Psi\theta - \theta^\top \Psi^\top Y + \theta^\top \Psi^\top \Psi\theta \\ &\quad + Y^\top \Psi(\Psi^\top \Psi)^{-1}\Psi^\top Y - Y^\top \Psi(\Psi^\top \Psi)^{-1}\Psi^\top Y \\ &= Y^\top [I - \Psi(\Psi^\top \Psi)^{-1}\Psi^\top] Y \\ &\quad + (\theta - (\Psi^\top \Psi)^{-1}\Psi^\top Y)^\top \Psi^\top \Psi(\theta - (\Psi^\top \Psi)^{-1}\Psi^\top Y). \end{aligned}$$

Therefore, the minimum

$$\min\ 2V(\theta, t) = Y^\top (I - \Psi(\Psi^\top \Psi)^{-1}\Psi^\top) Y$$

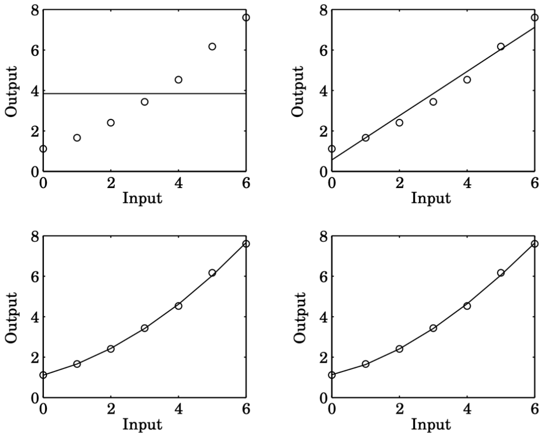
is assumed for $\hat{\theta}_\star = (\Psi^\top \Psi)^{-1}\Psi^\top Y$. ∎

Example: Least Squares Estimation

$$
\begin{aligned}
y(t) &= b_0 + b_1 u(t) + b_2 u^2(t) + e(t) \\
\sigma_e &= 0.1 \\
\varphi^\top(t) &= \begin{bmatrix} 1 & u(t) & u^2(t) \end{bmatrix} \\
\theta^\top &= \begin{bmatrix} b_0 & b_1 & b_2 \end{bmatrix}
\end{aligned}
$$

- Estimated models $b_0 = 1, b_1 = 0.5, b_2 = 0.1, \sigma = 0.1$
- Model 1: $y(t) = b_0$
- Model 2: $y(t) = b_0 + b_1 u$
- Model 3: $y(t) = b_0 + b_1 u + b_2 u^2$
- Model 4: $y(t) = b_0 + b_1 u + b_2 u^2 + b_3 u^3$

Model	$\hat{b}_0$	$\hat{b}_1$	$\hat{b}_2$	$\hat{b}_3$	$2V\sigma^{-2}$
1	3.85	-	-	-	3446
2	0.57	1.09	-	-	101
3	1.11	0.45	0.11 -	-	3.1
4	1.13	0.37	0.14	-0.003	2.7



Cost function smaller with more parameters, when to stop?

Outline

Recursive Estimators

Consider the linear regression model:

$$y(t) = \phi^\top(t)\theta. \tag{2}$$

- Idea: find a formula that expresses
  - in discrete time, $\hat{\theta}(t)$ in terms of $\hat{\theta}(t-1)$; or
  - in continuous time, $\hat{\theta}(t)$ is a solution of the dynamical system

$$\dot{\hat{\theta}} = \beta(\hat{\theta}, y(t), \phi(t)), \quad \hat{\theta}(t_0) = \hat{\theta}_0.$$

- Purposes:
  - Recursive computation of estimate as data is obtained is very useful for adaptive control
  - Track (slowly-varying) parameter variations

In this course, we study 1) Gradient Algorithms and 2) Recursive Least Squares Algorithms

Gradient Algorithms

Consider the linear regressor

$$y(t) = \phi^\top(t)\theta \tag{3}$$

with constant, unknown $\theta \in \mathbb{R}^n$, and available $\phi \in \mathbb{R}^n$ and $y \in \mathbb{R}$.

Standard Gradient Algorithm [Important!]

The update law

$$\dot{\hat{\theta}} = -\gamma\phi(t)(\phi^\top(t)\hat{\theta} - y(t)), \quad \hat{\theta}(t_0) = \hat{\theta}_0. \tag{4}$$

- $\gamma > 0$ - Adaptation gain (constant)
- $\hat{\theta}_0$ - Initial condition is arbitrary.

Adaptation gain $\gamma > 0$ allows us to vary the rate of adaptation.

$$y(t) = \phi^\top(t)\theta$$

Define

Estimation error $\tilde{\theta}(t) := \hat{\theta}(t) - \theta$

Output error $e := \phi^\top(t)\hat{\theta} - y(t)$.

We want to minimize the cost function

$$J(\hat{\theta}, t) = |e(\theta, t)|^2,$$

whose gradient is

$$\nabla_{\hat{\theta}} J(\hat{\theta}, t) = 2\phi e = 2\phi(\phi^\top\hat{\theta} - y) \implies \dot{\hat{\theta}} = -\frac{\gamma}{2}\nabla_{\hat{\theta}} J$$

It can be viewed as steepest descent.

If the regression $\phi \notin L_\infty$, we have

$$y(t) = \phi^\top(t)\theta \implies \underbrace{\frac{y}{\sqrt{1+\alpha|\phi|^2}}}_{\tilde{y}(t)} = \underbrace{\frac{\phi^\top(t)}{\sqrt{1+\alpha|\phi|^2}}}_{\tilde{\phi}^\top(t)}\theta$$

Alternative 1: *Normalized* Gradient Algorithm

$$\dot{\hat{\theta}} = -\frac{\gamma}{1+\alpha\phi^\top\phi}\phi(\phi^\top\hat{\theta} - y)$$

with constant parameters $\gamma > 0$ (adaptation gain) and $\alpha > 0$.

1. It is equivalent to the standard, with ϕ replaced by $\frac{\phi}{\sqrt{1+\alpha|\phi|^2}}$.
2. For the normalized estimator, RHS is *globally Lipschitz* in ϕ, even when ϕ is unbounded.

Alternative 2: Normalized Gradient Algorithm with Projection

Sometime the parameter θ is known a priori to lie in a set $\Theta \subset \mathbb{R}^n$ (closed, convex and delimited by a smooth boundary). Modify as

$$\dot{\hat{\theta}} = \begin{cases} -\dfrac{\gamma}{1+\alpha|\phi|^2}\phi(\phi^\top\hat{\theta} - y) & \text{if } \theta \in \text{int}(\Theta) \\ \text{Proj}\left[-\dfrac{\gamma}{1+\alpha|\phi|^2}\phi(\phi^\top\hat{\theta} - y)\right] & \text{if } \theta \in \partial\Theta \text{ and } e\phi^\top\theta_{\texttt{perp}} < 0 \end{cases}$$

with

- intΘ, $\partial\Theta$ - Interior and boundary of Θ
- Proj$[z]$ - projecting z onto the hyperplant tangent to $\partial\Theta$ at θ
- $\theta_{\texttt{prep}}$ - unit vector perpendicular to hyperplane, pointing outward

4

[4]Summary of projectors design: E. Lavretsky, T.E. Gibson, and A.M. Annaswamy, ArXiv 2012. (https://arxiv.org/pdf/1112.4232)

Example (Simple Projection)

A priori bounds p_i^- and p_i^+ are known, i.e.

$$\theta_i^\star \in [p_i^-, p_i^+].$$

The update law is then modified to

$$\begin{aligned} \dot{\theta}_i = 0 \quad & \text{if } \theta_i = p_i^- \text{ and } \dot{\theta}_i < 0 \\ & \text{or } \theta_i = p_i^+ \text{ and } \dot{\theta}_i > 0. \end{aligned}$$

Properties of Gradient Algorithms

- Linear regressor: $y(t) = \phi^\top(t)\theta$
- Gradient estimator: $\dot{\hat{\theta}} = -\gamma\phi(\phi\hat{\theta} - y)$
- Linear error equation: $e := \phi^\top\tilde{\theta} \quad (= \phi^\top\hat{\theta} - y)$

Theorem (Properties of Standard Gradient Estimator)

Consider the above gradient estimator with $\gamma > 0$ and the regression function $\phi : \mathbb{R}_+ \to \mathbb{R}^n$ piecewise continuous. Then,

1. Output error $e \in L_2$
2. Estimate error $\tilde{\theta} \in L_\infty$ (bounded estimate)
3. Monotonicity: $|\tilde{\theta}(t_a)| \geq |\tilde{\theta}(t_b)|$, for all $t_b \geq t_a \geq 0$.

Proof.

The dynamics of the estimate error $\tilde{\theta} := \hat{\theta} - \theta$ is the LTV system

$$\dot{\tilde{\theta}} = -\gamma \phi(t) \phi^\top(t) \tilde{\theta}. \tag{5}$$

Select the Lyapunov function $V(\tilde{\theta}) = \frac{1}{2}\tilde{\theta}^\top \tilde{\theta}$. Its derivative is

$$\dot{V} = -\gamma(\phi^\top \tilde{\theta})^2 = -\gamma e^2 \leq 0.$$

It means $\int_{t_a}^{t_b} \dot{V}(s) ds \leq 0$, thus $V(\tilde{\theta}(t_b)) - V(\tilde{\theta}(t_b)) \leq 0$, equivalently the third item. This also leads of the boundedness in 2).

Since V is positive and monotonically decreasing, the limit $V(\infty)$ is well define and

$$V(\infty) - V(0) = \int_0^\infty \dot{V}(s) = -\gamma \int_0^\infty e(s)^2 ds < \infty.$$

Therefore, $e \in L_2$ verifying 1). ∎

Theorem (Properties of Normalized Gradient Estimators)

Consider the linear regressor $y(t) = \phi^\top(t)\theta$ with $\phi \in PC[0,\infty)$ and the normalized gradient estimator

$$\dot{\hat{\theta}} = -\frac{\gamma}{1+\alpha|\phi|^2}\phi(\phi^\top\hat{\theta} - y)$$

Then,

1. $\frac{e}{\sqrt{1+\alpha|\phi|^2}} \in L_2 \cap L_\infty$
2. $\tilde{\theta} \in L_\infty$ and $\frac{d}{dt}\tilde{\theta} \in L_2 \cap L_\infty$
3. $\beta := \frac{\phi^\top\tilde{\theta}}{1+\|\phi(\cdot)\|_\infty} \in L_2 \cap L_\infty$.

The error dynamics is

$$\dot{\tilde{\theta}} = -\frac{\gamma}{1+\alpha|\phi|^2}\phi\phi^\top\tilde{\theta},\ \gamma > 0.$$

Its proof can be found in (Sastry & Bodson, page 64).

Gradient Estimator with Decaying Perturbation

Theorem (Effect of Exponentially Decaying Term)

Consider the *perturbed* linear regressor

$$y(t) = \phi^\top(t)\theta + \epsilon(t)$$

with $\epsilon(t)$ is an exponentially decaying term. Then, the theorems on standard and normalized gradient estimators still hold true.

Modify the Lyapunov function to

$$V(\tilde{\theta}) = \frac{1}{2}\tilde{\theta}^\top\tilde{\theta} + \frac{\gamma}{4}\int_t^\infty \epsilon^2(\tau)d\tau.$$

Then,

$$\dot{V} = -\gamma(\phi^\top\tilde{\theta})^2 - \gamma(\phi^\top\tilde{\theta})\epsilon - \frac{\gamma}{4}\epsilon^2 = -\gamma(\phi^\top\tilde{\theta} - \frac{1}{2}\epsilon)^2 \leq 0.$$

It follows a similar proof procedure.

Outline

Persistency of Excitation (PE)

- We derived the output error $e \in L_2$ and $\tilde{\theta} \in L_\infty$.
- How can we achieve $\tilde{\theta}(t) \to 0$ as $t \to \infty$?

Related to the LTV error dynamics

$$\dot{\tilde{\theta}} = -\gamma \phi(t) \phi^\top(t) \tilde{\theta}, \quad \gamma > 0$$

in the form of

$$\dot{\tilde{\theta}} = -A(t) \tilde{\theta}$$

with $A(t) \in \mathbb{R}^{n \times n}$ positive semidefinite (p.s.d.) for all t.

$A(t)$ uniformly p.d. with $\lambda_{\mathtt{min}}(A + A^\top) \geq 2\alpha \implies$ Exp. stability

Unfortunately, such is never the case, since

$$\text{rank}\ (\phi(t)\phi^\top(t)) = 1 < n \quad \forall t.$$

Definition (Persistency of Excitation)

A vector-valued function $\phi : \mathbb{R}_+ \to \mathbb{R}^n$ is *persistently exciting* (PE) if $\exists \alpha_1, \alpha_2, T > 0$ s.t.

$$\alpha_2 I \succeq \int_t^{t+T} \phi(\tau)\phi^\top(\tau) d\tau \succeq \alpha_1 I, \ \forall t \geq 0.$$

Interpretation

- Though $\phi\phi^\top$ is singular for all τ, the PE requires that ϕ rotates sufficiently in space that the integral of $\phi(\tau)\phi^\top(\tau)$ is uniformly p.d. over any interval of length $T > 0$.
- Re-expressing in scalar form

$$\alpha_2 \geq \int_t^{t+T} |\phi^\top(\tau)x|^2 d\tau, \quad \forall t \geq 0, \ |x| = 1.$$

Condition on energy of ϕ in all directions.

PE vs Uniform Complete Observability

The PE condition: $\exists \alpha_1, \alpha_2, T > 0$ s.t.

$$\alpha_2 I \succeq \int_t^{t+T} \phi(\tau)\phi^\top(\tau) d\tau \succeq \alpha_1 I, \ \forall t \geq 0.$$

$$\Updownarrow$$

The uniform complete observability (UCO) of the LTV system

$$\dot{\theta} = 0$$
$$y = \phi^\top(t)\theta$$

i.e. $A(t) = 0$ and $C(t) = \phi^\top(t)$

Theorem (PE and Exponential Stability)

Consider $\phi : \mathbb{R}_+ \to \mathbb{R}^n$ that is piecewise continuous. If ϕ is PE, then the zero equilibrium of

$$\dot{\tilde{\theta}} = -\gamma\phi(t)\phi^\top(t)\tilde{\theta},\ \gamma > 0 \tag{6}$$

is globally exponentially stable.

The converse claim is also true. It is necessary and sufficient.

Before proving the theorem, we need the following lemma.

Lemma

Assume $\forall \delta > 0$, $\exists k_\delta \geq 0$ s.t. $\forall t \geq 0$

$$\int_t^{t+T} \|K(\tau)\|^2 d\tau \leq k_\delta.$$

Then, the system (A, C) is UCO $\iff$ $(A + KC, C)$ is UCO.

Moreover, if the observability gramian of (A, C) satisfies

$$\beta_2 I \succeq W(t, t+\delta) \succeq \beta_1 I,$$

then the observability gramian of $(A + KC, C)$ satisfies these inequalities with identical δ and

$$\beta_1' = \frac{\beta_1}{1 + \sqrt{k_\delta \beta_2}}$$
$$\beta_2' = \beta_2 \exp(k_\delta \beta_2).$$

Proof of Theorem (PE & Exp Stability)

Consider the Lyapunov function $V = |\tilde{\theta}|^2$ s.t. $\dot{V} = -2\gamma|\phi^\top\tilde{\theta}|^2 \leq 0$,

$$\int_t^{t+T} \dot{V} d\tau = -2\gamma \int_t^{t+T} [\phi^\top(\tau)\tilde{\theta}(\tau)]^2 d\tau.$$

By the PE assumption, the system $(0, \phi^\top(t))$ is UCO. Under output injection with $K(t) = -\gamma\phi(t)$, the system becomes $(-\gamma\phi(t)\phi^\top(t), \phi^\top(t))$ with

$$k_\delta = \int_t^{t+T} |\gamma\phi(\tau)|^2 d\tau = \gamma^2 \mathrm{Tr}\left[\int_t^{t+T} \phi(\tau)\phi(\tau)^\top d\tau\right] \leq n\gamma^2\alpha_2.$$

By the lemma, the system $(A + KC, C)$ is UCO. Therefore, $\forall t$

$$\text{Gramian of } (A + KC, C) \leq -\frac{2\beta_1\gamma}{1 + \sqrt{n}\gamma\alpha_2}|\phi(t)|^2$$

Proof (cont'd)

By[a]

$$\int_t^{t+T} \dot{V} d\tau \leq -\frac{2\beta_1\gamma}{1+\sqrt{n}\gamma\alpha_2}|\phi(t)|^2$$

and the following (integral-type) Lyapunov stability theorem (see next slide), we complete the proof. ■

[a]For the system $(\tilde{A}, \tilde{C})$, the observability gramian is defined as

$$W_o(t_0, t_0+T) = \int_{t_0}^{t_0+T} \Phi^\top(s, t_0) C^\top(s) C(s) \Phi(s, t_0) ds.$$

The UCO condition

$$k_1 I \succeq W_o(t_0, t_0+T) \succeq k_2 I$$

is equivalent to

$$k_1|x(t_0)|^2 \geq \int_{t_0}^{t_0+T} |C(s)x(s)|^2 ds \geq k_2|x(t_0)|^2.$$

For the system $\dot{x} = f(x,t)$. If $\exists$ a function $V(x,t)$ and constants $k_1, k_2, k_3, \delta > 0$, s.t. $\forall x \in B_h,\ t \geq 0$

$$\begin{aligned} k_1|x|^2 \leq V(x,t) \leq k_2|x|^2 \\ \frac{\partial V}{\partial t}(x,t) + \frac{\partial V}{\partial x} f(x,t) \leq 0 \\ \int_t^{t+\delta} \frac{d}{d\tau} V(x(\tau),\tau) d\tau \leq -k_3|x(t)|^2. \end{aligned}$$

Then, $x(t)$ converges exponentially to 0.

Estimating exponential convergence rate

$$\alpha = \frac{1}{2T} \ln \left[\frac{1}{1 - \frac{2\gamma\alpha_1}{(1+\sqrt{2n}\gamma\alpha_2)^2}} \right]$$

- Increasing the gain $\gamma > 0$ cannot make it arbitrarily fast.
- When γ is sufficiently small, the rate $\alpha \propto \gamma$ approximately

Outline

Least Squares Algorithms

Regressor $y(t) = \phi^\top(t)\theta$
Estimation error $\tilde{\theta}(t) := \hat{\theta}(t) - \theta$
Output error $e := \phi^\top(t)\hat{\theta} - y(t)$.
Gradient estimator optimizes $J(\hat{\theta}, t) = |e(\theta, t)|^2$

Intuition: Least Squares Algorithms

Find the parameter θ to minimize the integral-squared-error (ISE)

$$\text{ISE} = \int_0^t e^2(s)ds = \int_0^t (\phi^\top(s)\hat{\theta} - y(s))^2 ds$$

The estimate of θ may be obtained from

$$\frac{\partial}{\partial\hat{\theta}}\text{ISE}(\hat{\theta}, t) = 2\int_0^t \phi(s)\left(\phi^\top(s)\hat{\theta} - y(s)\right) ds = 0$$

The least squares estimate is given by

$$\hat{\theta}_{\text{LS}}(t) = \left(\int_0^t \phi(s)\phi^\top(s) \right)^{-1} \left(\int_0^t \phi(s)y(s)ds \right)$$

if the inverse exists, and $\hat{\theta} = \theta$.

To get recursive formulations, let us define

$$P(t) = \left(\int_0^t \phi(s)\phi^\top(s) \right)^{-1},$$

so that[5]

$$\frac{d}{dt}[P^{-1}(t)] = \phi(t)\phi^\top(t).$$

[5] Calculation $\frac{d}{dt}[P^{-1}(t)]$: Since

$$\begin{aligned} 0 = \frac{d}{dt}I &= \frac{d}{dt}\left[P(t)P^{-1}(t)\right] \\ &= \frac{d}{dt}[P(t)]P^{-1}(t) + P(t)\frac{d}{dt}[P^{-1}(t)]. \end{aligned}$$

Then,

$$\dot{P} = -P\frac{d}{dt}[P^{-1}]P = -P\phi(t)\phi^\top(t)P, \quad P(t_0) = \left(\int_0^{t_0} \phi(s)\phi^\top(s)\right)^{-1}.$$

The least-square estimate can be represented as

$$\theta_{\text{LS}} = P(t)\int_0^t \phi(s)y(s)ds,$$

whose dynamics is[6]

$$\begin{aligned}\dot{\theta}_{\text{LS}} &= -P\phi(t)\phi^\top(t)\theta_{\text{LS}} + P\phi(t)y(t) \\ &= -P(t)\phi(t)\left(\phi^\top(t)\theta_{\text{LS}} - y\right) \\ &= -P\phi(t)e(t)\end{aligned}$$

[6] In practice, the recursive least-squares algorithms starts with arbitrary initial condition at $t_0 = 0$.

Recursive Least Squares Estimator

$$\dot{\hat{\theta}} = -\gamma P\phi(t)\left(\phi^\top(t)\hat{\theta} - y\right), \quad \hat{\theta}(0) = \hat{\theta}_0$$

$$\dot{P} = Q - \gamma P\phi(t)\phi^\top(t)P, \quad P(0) = P_0 \succ 0.$$

Adaptation gain $\gamma > 0$

Design parameter $Q \succeq 0$ (usually $Q = 0$)

Viewing the regression model as the LTV system

$$\dot{\theta} = 0$$

$$y = \phi^\top\theta.$$

RLS estimator is nothing but just the well-known **Kalman-Bucy filter**.

In the Standard RSL estimator with $Q = 0$,

$$\dot{P} = -\gamma P \phi \phi^\top P \implies \frac{d}{dt}(P^{-1}) = \gamma \phi \phi^\top.$$

Covariance Wind-up: This means that P^{-1} may become unbounded as $t \to \infty$ and thus P^{-1} may become arbitrarily small in some directions – adaptation becoming very slowly.

Least-Squares with Forgetting Factor

$$\dot{\hat{\theta}} = -\gamma P \phi \left(\phi^\top \hat{\theta} - y\right), \quad \hat{\theta}(0) = \hat{\theta}_0$$

$$\dot{P} = \gamma P(\lambda P - \phi \phi^\top P), \quad P(0) = P_0 \succ 0$$

or $$\frac{d}{dt}(P^{-1}) = \gamma \left(-\lambda P^{-1} + \phi \phi^\top\right)$$

- BIBO stability from $\phi \phi^\top$ to P^{-1}
- Another possible remedy: covariance resetting. $P(t_r^+) = k_0 I$

Normalized Least-Squares Estimator

$$\dot{\theta} = -\gamma \frac{P\phi(\phi^\top \hat{\theta} - y)}{1 + \alpha\phi^\top P\phi}$$

$$\dot{P} = -\gamma \frac{P\phi\phi^\top P}{1 + \alpha\phi^\top P\phi}, \quad P(0) \succ 0$$

with fixed parameters $\gamma, \alpha > 0$.

- The modification with forgetting factor can be combined to avoid covariance wind-up.
- RLS is complicated but has faster convergence rates compared to the gradient estimator.

Theorem (Normalized LS Estimator with Covariance Resetting)

Consider the regressor $y(t) = \phi^\top(t)\theta$ with the normalized LS estimator with covariance resetting

$$\dot{\theta} = -\gamma \frac{P\phi(\phi^\top\hat{\theta} - y)}{1 + \alpha\phi^\top P\phi}$$

$$\dot{P} = -\gamma \frac{P\phi\phi^\top P}{1 + \alpha\phi^\top P\phi}, \quad P(0) = P(t_r^+) = k_0 I$$

and $\gamma, \alpha > 0$, $t_r := \{t | \lambda_{\mathtt{min}}(P(t)) \leq k_1 < k_0\}$, and $\phi \in PC[0, \infty)$.

- $\frac{e}{\sqrt{1+\alpha\phi^\top P\phi}} \in L_2 \cap L_\infty$
- $\phi \in L_\infty,\ \dot{\phi} \in L_2 \cap L_\infty$
- $\beta = \frac{\phi^\top\tilde{\theta}}{1+\|\phi(\cdot)\|_\infty} \in L_2 \cap L_\infty$.
- If ϕ is PE, the estimate $\hat{\theta}$ satisfies $\lim_{t\to\infty} |\hat{\theta}(t) - \theta| = 0 \quad$ (exp.).

The proof is given in (Sastry & Bodson, pages 67 and 75).

System Identification

To identify an LTI system

$$\dot{x} = Ax + Bu$$
$$y = Cx + Du,$$

we need to represent the unknown "parameter" $\theta = (A, B, C, D)$ in a regression model.

We will study it next week.

What have we learned today?

- Motivation and formulation of real-time parameter estimation
- Regressor models and least squares problem
- Recursive estimator
  1) Gradient algorithm (standard, normalized, projection), $J(\theta) = |y - \phi^\top(t)\theta|^2$
  2) Recursive LS (standard, normalized, with forgetting factor, covariance resetting), , $J(\theta) = \int_0^t |y - \phi^\top(s)\theta|^2 ds$
- Pesistency of excitation (PE)
- Stability properties of online estimators

Homework

- Review the slides; Read (Sastry & Bodson, Ch. 2.3-2.5 and pp. 48-50)
- Find a regression model and numerically test two estimators in Matlab, Julia or Python (preparing for your first report).

ELE6214 - Commande de Systèmes Incertains

Lecture 3: Online System Identification

Bowen Yi

bowen.yi@polymtl.ca

Polytechnique Montréal

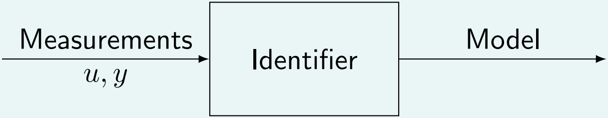
Hiver 2025

Outline

Online System Identification

We are concerned with the following questions with SISO systems:

- How to parameterize a system to get a linear regressor?
- How to generate the data to satisfy the PE condition?
- How to online estimate these parameters? [Solved in the last lecture]



Problem 1: Plant Parameterization

Given an *unknown* SISO plant

$$\dot{x} = Ax + Bu$$
$$y = Cx + Du,$$

or represented in the frequency domain

$$\frac{\hat{y}(s)}{\hat{u}(s)} = H(s) = \frac{n(s)}{m(s)},$$

we have the measurements (u, y) and unknown parameters θ included in (A, B, C, D) (or equivalently the coefficients in $n(s)$ and $m(s)$).

Can we get a linear regressor

$$Y(t) = \phi^{\top}(t)\theta$$

to estimate these parameters θ? (y has been used for plant output.)

Technical Challenges

1. No access to the internal state x
2. Overparameterization: n-dimensional SISO systems
   - State space model has $(n+1)^2$ parameters (n^2 in A, n in B and C, and 1 in D)
   - Transfer function has $2n$ parameters

$$H(s) = \frac{\theta_1 s^n + \ldots + \theta_n}{\theta_{n+1} s^n + \ldots + \theta_{2n}}$$

3. No access to derivative $\dot{x}$. For the special case that we have x,

$$\dot{x} = \underbrace{\begin{bmatrix} A & B \end{bmatrix}}_{:=\theta^\top} \begin{bmatrix} x \\ u \end{bmatrix}$$

4. ...

Problem 2: Data Generation (Identification Input Design)

In order to be able to identify θ online, we require ϕ persistently excited.

Can we operate the plant (A, B, C, D), or equivalently $H(s)$, to generate $\{u, y\}$ such that ϕ is PE?

Technical Challenges

1. Reformulate the PE condition of ϕ into some requirements on $u(t)$;
2. Operate the plant safely; Input should not be too large. [Not the focus of our course]
3. $\cdots$

Some Definitions of Transfer Function

- Monic: a polynomial in s is monic if the coefficient of the highest power in s is 1
- Hurwitz: if its roots lie in $\mathbb{C}_<$
- Stable: transfer function has its denominator polynomial Hurwitz
- Minimum phase: if the numerator polynomial is Hurwitz
- Relative degree: difference between the degrees of the denominator and numerator polynomials
- Proper: relative degree ≥ 0
- Strictly proper: relative degree > 0

Assumptions

A1 **Plant Assumptions**

SISO LTI system, whose transfer function $\hat{P}(s) = \frac{\hat{y}_p(s)}{\hat{r}(s)}$ is

$$\hat{P}(s) = k_p \frac{\hat{n}_p(s)}{\hat{d}_p(s)}$$

$\hat{r}(s), \hat{y}_p(s)$ - Laplace transforms of input/output
$\hat{n}_p(s), \hat{d}_p(s)$ - monic, coprime polynomials of degrees m and n
n is known, but m is unknown
Plant is strictly proper $m \leq n - 1$

We do *not* assume the stability of the plant.

A2 **Reference Input Assumptions**
Input $r(\cdot)$ is piecewise continuous and bounded on $\mathbb{R}_+$.

A3 **Bounded Output Assumption**
The plant is located in a control loop such that $r, y_p \in L_\infty$.[a]

[a] This can be further relaxed as the regular signal assumption. (Sastry & Bodson, page 70).

Objective: Estimate k_p and the coefficients of the polynomials $\hat{n}_p(s), \hat{d}_p(s)$ from measurements of input $r(t)$ and output $y_p(t)$ only.

Outline

Generating Linear Regression Models

Parameterization of Unknown Plants

Transfer function $\hat{P}(s)$ can be explicitly written as

$$\frac{\hat{y}_p(s)}{\hat{r}(s)} = \hat{P}(s) = \frac{\alpha_n s^{n-1} + \cdots + \alpha_1}{s^n + \beta_n s^{n-1} + \cdots + \beta_1} \tag{1}$$

with $2n$ unknown coefficients.

Express as the linear regressor

$$s^n \hat{y}_p(s) = \left(\alpha_n s^{n-1} + \cdots + \alpha_1\right) \hat{r}(s) - \left(\beta_n s^{n-1} + \cdots + \beta_1\right) \hat{y}_p(s)$$

Not practical: Require explicit differentiations to be implemented!

Introduce a monic n-th order Hurwitz (but arbitrary) polynomial

$$\hat{\lambda}(s) = s^n + \lambda_n s^{n-1} + \cdots + \lambda_1.$$

Then, from

$$\hat{P}(s) = \frac{\hat{y}_p(s)}{\hat{r}(s)} := k_p \frac{\hat{n}_p(s)}{\hat{d}_p(s)} \implies k_p \hat{n}_p(s)\hat{r}(s) - \hat{d}_p(s)\hat{y}_p(s) = 0.$$

this leads to a linear regressor, but $\hat{n}_p(s)$ and $\hat{d}_p(s)$ are not proper! To address this implementation issue, we rewrite it as

$$\hat{\lambda}(s)\hat{y}_p(s) = k_p \hat{n}_p(s)\hat{r}(s) + (\hat{\lambda}(s) - \hat{d}_p(s))\hat{y}_p(s),$$

or equivalently

$$\hat{y}_p(s) = \frac{\alpha_n s^{n-1} + \cdots + \alpha_1}{\hat{\lambda}(s)}\hat{r}(s) + \frac{(\lambda_n - \beta_n)s^{n-1} + \cdots + (\lambda_1 - \beta_1)}{\hat{\lambda}(s)}\hat{y}_p(s).$$

$$\hat{y}_p(s) = \frac{\alpha_n s^{n-1} + \cdots + \alpha_1}{\hat{\lambda}(s)}\hat{r}(s) + \frac{(\lambda_n - \beta_n)s^{n-1} + \cdots + (\lambda_1 - \beta_1)}{\hat{\lambda}(s)}\hat{y}_p(s)$$

Define the polynomials

$$\hat{a}^*(s) = \alpha_n s^{n-1} + \cdots + \alpha_1$$
$$\hat{b}^*(s) = (\lambda_n - \beta_n)s^{n-1} + \cdots + (\lambda_1 - \beta_1).$$

and (unknown) parameters

$$\theta_a := \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}, \quad \theta_b := \begin{bmatrix} \lambda_1 - \beta_1 \\ \vdots \\ \lambda_n - \beta_n \end{bmatrix}$$

New representation

The plant can be equivalently represented as

$$\hat{y}_p(s) = \frac{\hat{a}^*(s)}{\hat{\lambda}(s)}\hat{r}(s) + \frac{\hat{b}^*(s)}{\hat{\lambda}(s)}\hat{y}_p(s) \tag{2}$$

Linear Regressor Form

$$\hat{y}_p(s) = \frac{\hat{a}^*(s)}{\hat{\lambda}(s)}\hat{r}(s) + \frac{\hat{b}^*(s)}{\hat{\lambda}(s)}\hat{y}_p(s)$$

$$= \theta_a^\top \underbrace{\begin{bmatrix} 1 \\ s \\ \vdots \\ s^{n-1} \end{bmatrix} \frac{1}{\hat{\lambda}(s)}\hat{r}(s)}_{\hat{w}_p^{(1)}(s)} + \theta_b^\top \underbrace{\begin{bmatrix} 1 \\ s \\ \vdots \\ s^{n-1} \end{bmatrix} \frac{1}{\hat{\lambda}(s)}\hat{y}_p(s)}_{\hat{w}_p^{(2)}(s)}$$

Rewrite in time domain leading to a linear regressor

$$y_p(t) = w_p(t)^\top \theta, \quad \theta := \begin{bmatrix} \theta_a \\ \theta_b \end{bmatrix}, \; w_p := \begin{bmatrix} w_p^{(1)} \\ w_p^{(2)} \end{bmatrix}$$

How can we obtain $w_p(t)$?

State space realization can be found in controllable canonical form

$$\dot{w}_p^{(1)} = \Lambda w_p^{(1)} + b_\lambda r$$
$$\dot{w}_p^{(2)} = \Lambda w_p^{(2)} + b_\lambda y_p$$

with[a]

$$\Lambda = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -\lambda_1 & -\lambda_2 & \cdots & \cdots & -\lambda_n \end{bmatrix} \quad b_\lambda = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

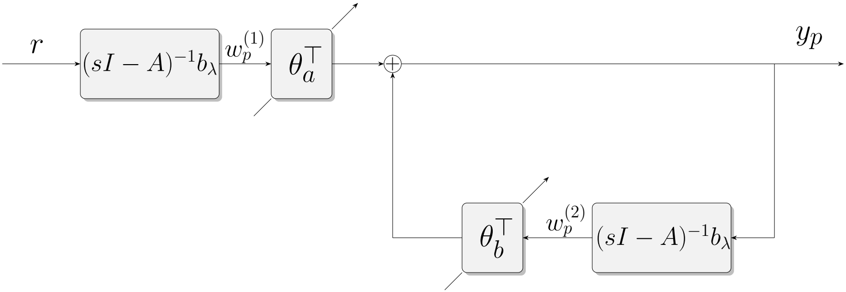
with initial conditions $w_p^{(1)}(0)$, $w_p^{(2)}(0)$.

---

[a] Recall

$$(sI - \Lambda)^{-1} b_\lambda = \frac{1}{\hat{\lambda}(s)} \begin{bmatrix} 1 & s & \ldots & s^{n-1} \end{bmatrix}^\top .$$

$$\hat{y}_p(s) = \frac{\hat{a}^*(s)}{\hat{\lambda}(s)}\hat{r}(s) + \frac{\hat{b}^*(s)}{\hat{\lambda}(s)}\hat{y}_p(s)$$



Implementation

We have the state-space realization of w_p:

$$\dot{w}_p^{(1)} = \Lambda w_p^{(1)} + b_\lambda r$$
$$\dot{w}_p^{(2)} = \Lambda w_p^{(2)} + b_\lambda y_p.$$

However, we do *not* have the initial condition $w_p(0) \in \mathbb{R}^{2n}$.

Filter Design

$$\dot{w}^{(1)} = \Lambda w^{(1)} + b_\lambda r$$
$$\dot{w}^{(2)} = \Lambda w^{(2)} + b_\lambda y_p.$$

with $w(0) \in \mathbb{R}^{2n}$, which only uses the available signals r, y_p, without knowledge of the plant parameters.

$$\text{Hurwitz } \Lambda \implies \lim_{t\to\infty} |w(t) - w_p(t)| = 0 \text{ (exp.)}$$

Implementation (cont'd)

With the available signals r, y_p, apply the filter

$$\dot{w}^{(1)} = \Lambda w^{(1)} + b_\lambda r$$
$$\dot{w}^{(2)} = \Lambda w^{(2)} + b_\lambda y_p.$$

with $w(0) \in \mathbb{R}^{2n}$ leading to the (perturbed) linear regressor

$$y(t) = w(t)^\top \theta + \epsilon(t)$$

with the error $\epsilon(t)$ exponentially decaying to zero.

Use the gradient estimator

$$\dot{\hat{\theta}} = -\gamma w(w^\top \hat{\theta} - y), \quad \gamma > 0$$

or the RLS estimator to estimate θ!

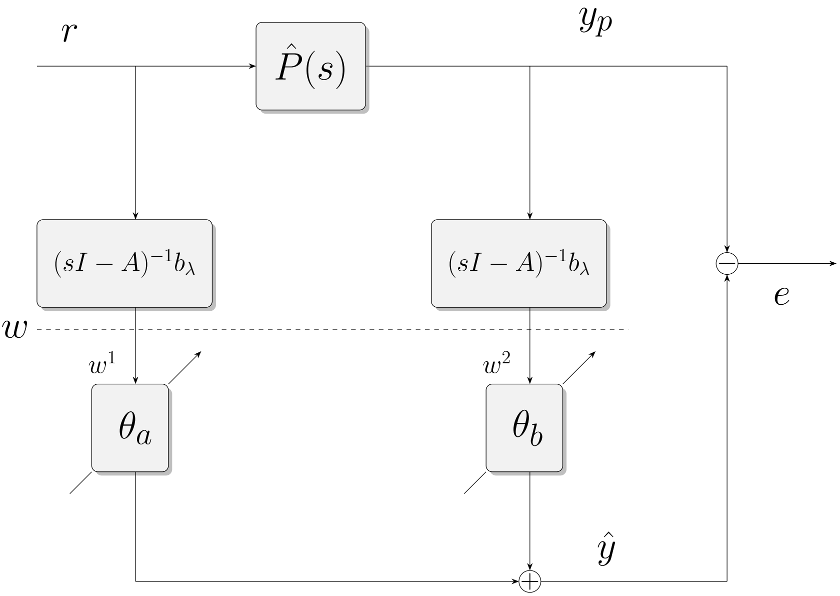


Figure: Identifier Structure 1: Error Equation Identifier

Outline

Different ways to get regression models:

- Equation error identifier (discussed)
- Output error approach (Landau, 1979)
- Model reference approach (Luders & Narendra, 1973)

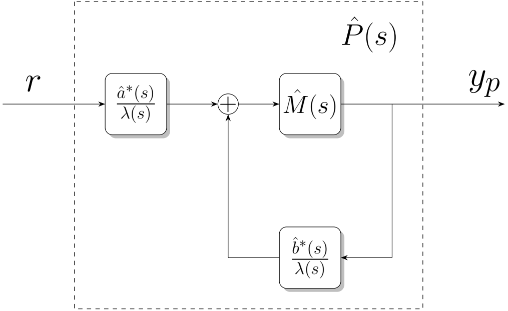


Figure: Model Reference Reparameterization

Modifying the reference model through feedback $\frac{\hat{b}^*(s)}{\lambda(s)}$ and feedforward $\frac{\hat{a}^*(s)}{\lambda(s)}$ action, so as to match the plant transfer function.

Homework 1

Numerically test the first identification structure and the gradient estimator for your model.

Further Assumptions

A4 **Reference Model Assumptions**

The reference model is an SISO LTI system (selected by us)

$$\hat{M}(s) = k_m \frac{\hat{n}_m(s)}{\hat{d}_m(s)}$$

- $\hat{n}_m(s), \hat{d}_m(s)$ are monic, coprime polynomials of degrees $l, k \leq n$.
- $M(s)$ is strictly proper
- Its relative degree is no greater than the relative degree of the plant $\hat{P}(s)$, i.e. $1 \leq k - l \leq n - m$
- $\hat{d}_m(s)$ is Hurwitz

A5 **Postive Real Model**

$\hat{M}(s)$ is strictly positive real. [Show you def later.]

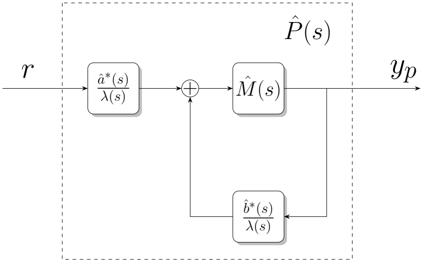
Objective: Estimate polynomials $\hat{a}^*(s), \hat{b}^*(s) \iff$ estimate $\hat{P}(s)$

Polynomial $\hat{\lambda}(s)$

- The polynomial $\hat{\lambda}(s)$ is a monic, Hurwitz of degree $n-1$.
- Similar role as that in the equation error identifier.
- The zeros of $\hat{\lambda}(s)$ should contain those of $\hat{n}_m(s)$ in the reference model:

$$\boxed{\hat{\lambda}(s) = \hat{n}_m(s)\hat{\lambda}_0(s)}$$

with another monic, Hurwitz polynomial $\hat{\lambda}_0(s)$ of degree $n-l-1$.



Theorem (Model Matching)

There exist *unique* $\hat{a}^*(s)$ and $\hat{b}^*(s)$, in the above figure, such that the transfer function $r \to y_p$ is the plant transfer function $\hat{P}(s)$.

Proof.

Existence. The transfer function $r \to y_p$ is given by

$$\frac{\hat{y}_p(s)}{\hat{r}(s)} = \frac{\hat{a}^*}{\hat{\lambda}} \frac{\frac{k_m \hat{n}_m}{\hat{d}_m}}{1 - k_m \frac{\hat{n}_m}{\hat{d}_m} \frac{\hat{b}^*}{\hat{\lambda}}} = \frac{\hat{a}^* k_m \hat{n}_m}{\hat{\lambda} \hat{d}_m - k_m \hat{n}_m \hat{b}^*} = \frac{k_m \hat{a}^*}{\hat{\lambda}_0 \hat{d}_m - k_m \hat{b}^*},$$

which equals to $\hat{P}(s)$ iff

$$\boxed{\hat{\lambda}_0 \hat{d}_m - k_m \hat{b}^* = \frac{k_m}{k_p} \hat{d}_p \frac{\hat{a}^*}{\hat{n}_p}} \tag{C1}$$

The problem is therefore to find polynomials $\hat{a}^*, \hat{b}^*$ of degrees $\leq n-1$.

A solution can be found by inspection. Divide $\hat{\lambda}_0 \hat{d}_m$ by $\hat{d}_p$: denote by $\hat{q}$ the quotient of degree $k-l-1$ and let $k_m \hat{b}^*$ be the remainder of degree $n-1$. In other words, let

$$\hat{\lambda}_0 \hat{d}_m = \hat{q} \hat{d}_p + k_m \hat{b}^*.$$

This defines $\hat{b}^*$ appropriately. Eq. (C1) is satisfied if $\hat{a}^*$ is

$$\hat{a}^* = \frac{k_p}{k_m}\hat{q}\hat{n}_p$$

The degree of the polynomial in RHS is $m + k - l - 1$, which is at most $n - 1$ by assumption, s.t. the degree requirements are satisfied.

Uniqueness. Assume that $\exists \hat{a}^* + \delta\hat{a}, \hat{b}^* + \delta\hat{b}$ satisfying

$$\hat{\lambda}_0\hat{d}_m - k_m\left(\hat{b}^* + \delta\hat{b}\right) = \frac{k_m}{k_p}\hat{d}_p\frac{(\hat{a} + \delta\hat{a})}{\hat{n}_p}$$

We find that

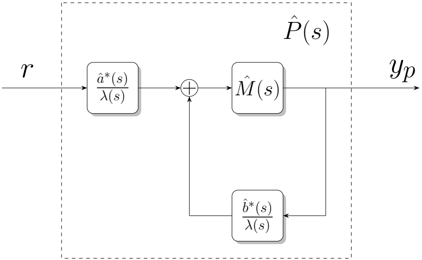
$$\frac{\delta\hat{a}}{\delta\hat{b}} = -k_p\frac{\hat{n}_p}{\hat{d}_p} = -\hat{P}$$

Recall that $\hat{n}_p, \hat{d}_p$ are assumed to be coprime, while the degree of $\hat{d}_p$ and $\delta\hat{b}$ are n and at most $n-1$, respectively. Therefore, Eq. (C1) *cannot* have any solution. ∎

$$\frac{\hat{a}^*(s)}{\hat{\lambda}(s)} = a_0^* + a^{*\top}(sI - \Lambda)^{-1}b_\lambda$$

$$\frac{\hat{b}^*(s)}{\hat{\lambda}(s)} = b_0^* + b^{*\top}(sI - \Lambda)^{-1}b_\lambda$$

parameters $a_0, b_0 \in \mathbb{R}$ and $a^*, b^* \in \mathbb{R}^{n-1}$ (Same Λ, b_λ as previous but with different dimensionality)



Filter:

$$\dot{w}^{(1)} = \Lambda w^{(1)} + b_\lambda r$$
$$\dot{w}^{(2)} = \Lambda w^{(2)} + b_\lambda y_p$$

The filter dimension is lower than that in the error equation identifier – saving in computations.

The regressor vector ϕ is

$$\phi(t)^\top := \begin{bmatrix} r(t) & w^{(1)\top}(t) & y_p(t) & w^{(2)^\top}(t) \end{bmatrix} \in \mathbb{R}^{2n}$$

and the parameter vector $\theta^{*\top} := \begin{bmatrix} a_0^* & a^{*\top} & b_0^* & b^{*\top} \end{bmatrix} \in \mathbb{R}^{2n}$

The signal coming into the reference model $\hat{M}(s)$ is

$$\mathcal{L}^{-1}\left\{ \frac{\hat{a}^*(s)}{\hat{\lambda}(s)}\hat{r}(s) + \frac{\hat{b}^*(s)}{\hat{\lambda}(s)}\hat{y}_p(s) \right\} := \phi^\top(t)\theta$$

The output of $\hat{M}(s)$ is y_p. Therefore,[1]

$$\boxed{y_p(t) = \hat{M}\left[\phi^\top(t)\theta\right]}$$

It is similar to the linear regression model, but with the transfer function $\hat{M}$.

[1]We omit the exponentially decaying term stemming from the filter.

SPR-based Linear Regressor

The equation

$$\boxed{y_p(t) = \hat{M}\left[\phi^\top(t)\theta\right]}$$

is a linear regressor with a strictly positive real (SPR) transfer function.

Questions:

- What is SPR?
- How to verify SPR?
- Any properties for SPR transfer functions?
- How can we design an online estimator for SPR-based linear regressors?

Positive Real Function

We require $\hat{M}(s)$ to be strictly positive real (SPR), then get the SPR regression model:[2]

$$y_p(t) = \hat{M}\left[\phi^\top(t)\theta\right].$$

Definition (Positive Real)

A rational function $\hat{M}(s)$ of the complex variable $s = \sigma + j\omega$ is *positive real* (SR), if

- $\hat{M}(\sigma) \in \mathbb{R}$ for all $\sigma \in \mathbb{R}$
- $\mathrm{Re}[\hat{M}(\sigma + j\omega)] \geq 0$ for all $\sigma > 0,\ \omega \geq 0$

It is *strictly positive real* (SPR) if, for some $\epsilon > 0$, $\hat{M}(s - \epsilon)$ is PR.

[2] The SPR/PR originates from Network theory. A rational transfer function is the driving point impedance of a passive network iff it is PR.

Criterion for SPR

(Ioannou & Tao, 1987) A strictly proper function $\hat{M}(s)$ is SPR if and only if

- $\hat{M}(s)$ is stable
- $\mathrm{Re}(\hat{M}(j\omega)) > 0$, for all $\omega \geq 0$
- $\lim_{\omega\to\infty} \omega^2 \mathrm{Re}(\hat{M}(j\omega)) > 0$.

For example, the transfer function

$$\hat{M}(s) = \frac{s+c}{(s+a)(s+b)}$$

is SPR is and only if $a > 0, b > 0$ and $a + b > c > 0$.

Positive Real Lemma

Let

$$G(s) = C(sI - A)^{-1}B + D$$

where (A, B) is controllable and (A, C) is observable (i.e. (A, B, C, D) is a minimal realization of $G(s)$). Then, $G(s)$ is positive real $\exists P = P^\top \succ 0$, L and W s.t.

$$\begin{aligned} PA + A^\top P &= -L^\top L \\ PB &= C^\top - L^\top W \\ W^\top W &= D + D^\top. \end{aligned}$$

Kalman-Yakubovich-Popov (KYP) Lemma

In the above lemma, $G(s)$ is *strictly* positive real if and only if we replace the first equation with

$$PA + A^\top P = -L^\top L - \epsilon P, \quad \text{for some} \quad \epsilon > 0$$

Gradient Estimator for SPR-based Linear Regressors

We have the linear regressor

$$y_p(t) = \hat{M}\left[\phi^\top(t)\theta\right].$$

with $\hat{M}(s)$ SPR.

Gradient Algorithm with SPR

The gradient

$$\dot{\hat{\theta}} = -\gamma\phi(t)\left(\hat{M}\left[\phi^\top(t)\hat{\theta}\right] - y(t)\right)$$

with the adaptation gain $\gamma > 0$.

The SPR error equation $e = \hat{y} - y := \hat{M}[\phi^\top\hat{\theta}] - y$ is

$$e = \hat{M}\left[\phi^\top(t)\tilde{\theta}\right]$$

with $\tilde{\theta} := \hat{\theta} - \theta$.

Since $\hat{M}(s)$ is SPR, a state space realization is

$$\begin{aligned} \dot{e}_m &= Ae_m + b\phi^\top(t)\tilde{\theta} \\ e &= c^\top e_m \\ \dot{\tilde{\theta}} &= -\gamma c^\top e_m \phi(t). \end{aligned}$$

with $(A, b, c^\top)$ a minimal realization of $\hat{M}(s)$ and e_m its internal state.

Theorem (Stability of SPR Gradient Estimator)

Assume $\phi : \mathbb{R}_+ \to \mathbb{R}^{2n}$ piecewise continuous. The above realization guarantees

- $e_m, e \in L_2$
- $e_m, e, \tilde{\theta} \in L_\infty$
- If ϕ is PE and $\phi, \dot{\phi} \in L_\infty$, then the above system is globally exponentially stable.

Proof.

$\hat{M}(s)$ is SPR, thus $\exists P, Q \succ 0$ such that

$$PA + A^\top P = -Q$$
$$Pb = c$$

Consider the Lyapunov function

$$V = \gamma e_m^\top P e_m + \tilde{\theta}^\top \tilde{\theta},$$

whose derivative is

$$\begin{aligned}\dot{V} &= \gamma e_m^\top P A e_m + \gamma e_m^\top P b \phi^\top \tilde{\theta} + \gamma e_m^\top A^\top P e_m \\ &\quad + \gamma \phi^\top \tilde{\theta} b^\top P e_m - 2\gamma c^\top e_m \phi \top \tilde{\theta} \\ &= -\gamma e_m^\top Q e_m \leq 0\end{aligned}$$

The first two claims follow immediately. The last point can be found in (Sastry & Bodson, page 86). ■

Outline

Lemma. PE through LTI Systems

Let $\phi : \mathbb{R}_+ \to \mathbb{R}^n$. If

- ϕ is PE, and $\phi, \dot{\phi} \in L_\infty$
- $\hat{H}(s)$ is a stable, minimum phase, rational transfer function,

then $\hat{H}[\phi(t)]$ is also PE.

Lemma. PE and L_2 Signals

Let $\phi_1, \phi_2 : \mathbb{R}_+ \to \mathbb{R}^n$ be piecewise continuous. If

- ϕ_1 is PE
- $\phi_2 \in L_2$

then $\phi_1 + \phi_2$ is PE.

Reference Input Design

- Exponential convergence of estimation error $\tilde{\theta}$ builds upon the PE of ϕ, but the reference input to the plant is $r(t)$.
- $\phi \in$ PE $\iff$ r is sufficiently rich of order $2n$.

  A single sinusoid in the input r contributes 2 points to the spectrum: at $\pm\omega_0$.
- If $r(t)$ is stationary and sufficiently rich of order $2n$, the identified parameter $\hat{\theta}$ (gradient or normalized LS estimator with covariance resetting) will converge to θ exponentially faster.
- Any ideas on the design of identification references r in experiments?

Frequency domain conditions (Sastry & Bodson, page 90)

What have we learned?

- Parameterization of a given plant
  1) Error equation approach
  2) Reference model (or SPR error equation) approach
- Given a plant $G(s)$, how to generate a linear regressor?
- Positive real function
- Positive real and KYP lemmata
- SPR gradient estimator
- Sufficient Rich Input

Homework

- Read Ch2 of (Sastry & Bodson); Understand the sufficient richness on page 92.
- Verify if your selected plant satisfies the assumptions.
- Start your first report – applying the learned two approaches to identify your plant and doing simulations

Additional Reading Materials

- PE condition is necessary and sufficient to the GES of

$$\dot{\tilde{\theta}} = -\phi(t)\phi^\top(t)\tilde{\theta}$$

See for the GAS case:

N. Barabanov, R. Ortega, On global asymptotic stability of $\dot{x} = -\phi(t)\phi^\top(t)x$ with ϕ not persistently exciting. *Syst. Control Lett.*, 109 (2017): 24-29.

ELE6214 - Commande de Systèmes Incertains

Lecture 4: Model Reference Adaptive Control

Bowen Yi

bowen.yi@polymtl.ca

Polytechnique Montréal

Hiver 2025

Outline

Adaptive Control

- Adaptive control: a technique of applying system identification techniques to obtain a model of a plant and its environment and using this model to design a controller.
- Large basket of tools: for LTI systems, the most popular may refer to
  - Self-tuning regulator (STR)

    Identifier-based: separate the estimation of unknown parameters from the design of controllers (certainty equivalence principle)

  - Model reference adaptive control (MRAC) – our focus

    Behavior of the controlled plant remains close to the one of a *desired model*, despite uncertainties or variations in the plant.

MRAC: Direct Approach

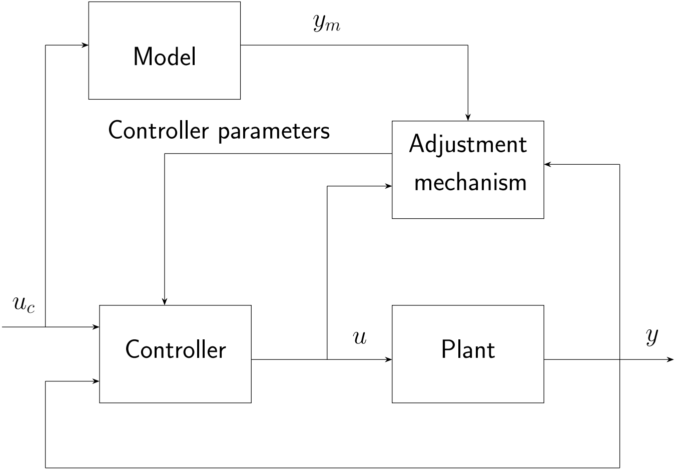


Figure: Controller parameters are adjusted directly

MRAC: Indirect Approach

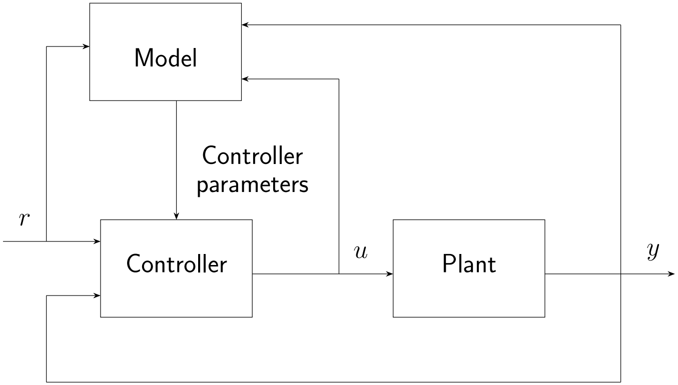


Figure: Controller parameters are adjusted indirectly by first estimating parameters of a plant/process model and then designing a controller

Motivating Example

Consider the plant

$$\hat{P}(s) = \frac{k_p}{s + a_p}$$

with k_p, a_p unknown. We design a *reference model*

$$\hat{M}(s) = \frac{k_m}{s + a_m}, \quad a_m > 0$$

and hope that the given model can behave like the reference model.

- We may apply the control

  $$u(t) = c_0 r(t) + d_0 y(t)$$

  with $r(t)$ the reference input of $\hat{M}(s)$.
- Two models are matched $\iff$ $c_0 = \frac{k_m}{k_p},\ d_0 = \frac{a_p - a_m}{k_p}$ (unknown)

cont'd

- Direct parameterization: Viewing the controller parameters c_0, d_0 as the unknown parameters, i.e.

$$\theta = \begin{bmatrix} c_0 & d_0 \end{bmatrix}^\top ,$$

  it belongs the direct approach.

- Indirect parameterization: Viewing the plant parameters k_p, a_p as unknown parameters, i.e.

$$\theta = \begin{bmatrix} k_p & a_p \end{bmatrix}^\top$$

  and using θ to solve the controller parameters c_0, d_0, it belongs the indirect approach.

Caveat: It is convenient to divide the algorithms into direct and indirect, but the distinction should not be overemphasized.

MRAC: Input-Error and Output-Error

In model reference adaptive control (MRAC), we need to generate a linear regression equation on the unknown parameter vector θ, or equivalently an error equation.

Two basic approaches to get error equations:

- Input Error
- Output Error

Outline

Assumptions

A1 Plant: The plant is a SISO LTI system

$$\frac{\hat{y}_p(s)}{\hat{u}(s)} = \hat{P}(s) = k_p \frac{\hat{n}_p(s)}{\hat{d}_p(s)}$$

- $\hat{n}_p(s), \hat{d}_p(s)$ are monic, coprime polynomials of degree m and n
- Strictly proper and minimum phase[a]
- The sign of the high-frequency gain k_p is known (w.l.g. $k_p > 0$).

A2 Reference Model:

$$\frac{\hat{y}_m(s)}{\hat{r}(s)} = \hat{M}(s) = k_m \frac{\hat{n}_m(s)}{\hat{d}_m(s)}$$

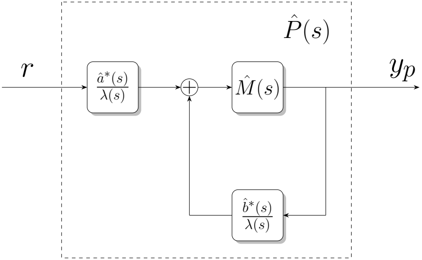
- $\hat{n}_m(s), \hat{d}_m(s)$ with the same orders as plant (i.e. n and m)
- Stable and minimum phase and $k_m > 0$

A3 Reference signal: $r(\cdot) \in PC[0, \infty) \cap L_\infty$.

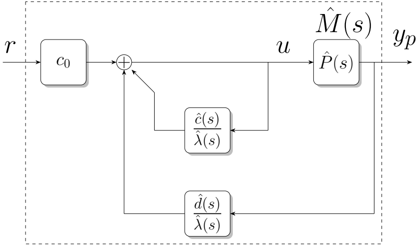
[a] Plant may be unstable.

Model Matching

Model Reference Identification



Model Reference Adaptive Control



Controller structure: linear combination of r, u, y_p

$$u = \begin{bmatrix} c_0 & \frac{\hat{c}(s)}{\hat{\lambda}(s)} & \frac{\hat{d}(s)}{\hat{\lambda}(s)} \end{bmatrix} \begin{bmatrix} r \\ u \\ y_p \end{bmatrix}$$

Control Objectives

The reference model $\hat{M}(s)$ – under the reference input $r(t)$ – will generate the reference output

$$y_m(t) = M[r(t)].$$

We want to design a control $u(t)$ for the plant $\hat{P}(s)$ such that

- all the states are bounded
- the plant output y_p asymptotically converges to the reference output y_m, i.e.
$$\lim_{t\to\infty} |y_p(t) - y_m(t)| = 0.$$
- if possible, we can estimate some unknown parameters.

Transfer function from r to y_p

From

$$u = c_0 r + \frac{\hat{c}(s)}{\hat{\lambda}(s)}[u] + \frac{\hat{d}(s)}{\hat{\lambda}(s)}[y_p]$$

we have

$$u = \frac{\hat{\lambda}}{\hat{\lambda} - \hat{c}} \left(c_0 r + \frac{\hat{d}}{\hat{\lambda}}(y_p) \right)$$

Combining $y_p = k_p \frac{\hat{n}_p}{\hat{d}_p}(u)$, we have

$$\frac{\hat{y}_p}{\hat{r}} = \frac{c_0 k_p \hat{\lambda} \hat{n}_p}{(\hat{\lambda} - \hat{c})\hat{d}_p - k_p \hat{n}_p \hat{d}} \quad \overset{?}{=} \quad M(s) = k_m \frac{\hat{n}_m}{\hat{d}_m}$$

A necessary condition: $\lambda(s)$ is Hurwitz (for implementation) and can be decomposed into

$$\hat{\lambda}(s) = \hat{\lambda}_0(s)\hat{n}_m(s)$$

with $\hat{\lambda}_0(s)$ an abribtrary Hurwitz polynomial of degree $(n - m - 1)$.

Matching Equality

Theorem

There exist unique $c_0^*, \hat{c}^*(s), \hat{d}^*(s)$ s.t. the transfer function from $r \to y_p$ is $\hat{M}(s)$.

Proof. $\star$

(*Existence*) TF r to y_p is $\hat{M} \iff$ *matching equality* is satisfied

$$\boxed{(\lambda - \hat{c}^*)\hat{d}_p - k_p \hat{n}_p \hat{d}^* = c_0^* \frac{k_p}{k_m} \hat{\lambda}_0 \hat{n}_p \hat{d}_m} \tag{S1}$$

The solution can be found by inspection. Divide $\hat{\lambda}_0 \hat{d}_m$ by $\hat{d}_p$, let $\hat{q}$ be the quotient (of degree $n - m - 1$) and $-k_p \hat{d}^*$ the remainder (of degree $n - 1$). Thus, a feasible solution is

$$\hat{d}^* = \frac{1}{k_p}\left(\hat{q}\hat{d}_p - \hat{\lambda}_0 \hat{d}_m\right), \quad \hat{c}^* = \lambda - \hat{q}\hat{n}_p, \quad c_0^* = \frac{k_m}{k_p}.$$

cont'd

(*Uniqueness*) Assume $\exists c_0 = c_0^* + \delta c_0$, $\hat{c} = \hat{c}^* + \delta\hat{c}$, $\hat{d} = \hat{d}^* + \delta\hat{d}$ satisfying (S1). The following equality must then be satisfied

$$\delta\hat{c}\hat{d}_p + k_p\hat{n}_p\delta\hat{d} = -\delta c_0 \frac{k_p}{k_m}\hat{\lambda}_0\hat{n}_p\hat{d}_m \qquad (3.2.10)$$

Recall that $\hat{d}_p$, $\hat{n}_p$, $\hat{\lambda}_0$ and $\hat{d}_m$ have degrees n, m, $n-m-1$ and n, respectively, with $m \leq n-1$, and $\delta\hat{c}$ and $\delta\hat{d}$ have degrees at most $n-2$ and $n-1$. Consequently, the RHS has degree $2n-1$ and LHS has degree at most $2n-2$.

No solution exists unless $\delta c_0 = 0$, so that c_0^* is unique. Let, then, $\delta c_0 = 0$, so

$$\frac{\delta\hat{c}}{\delta\hat{d}} = -k_p\frac{\hat{n}_p}{\hat{n}_p} = -\hat{P},$$

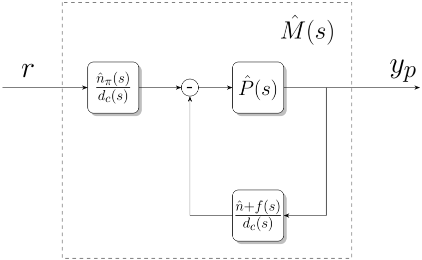
which has no solution since $\hat{n}_p, \hat{d}_p$ are coprime. ■

Remarks

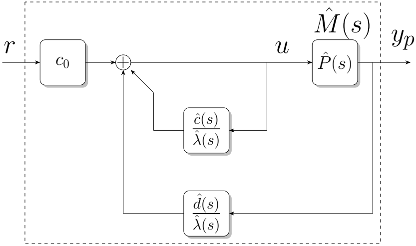
- The coprimeness of $\hat{n}_p, \hat{d}_p$ guarantees the uniqueness. Otherwise, there still exist feasible solutions.
- When model matching occurs, the forward block actually cancels the zeros of $\hat{P}(s)$ and replaces them by the zeros of $\hat{M}(s)$, i.e.

$$\frac{\hat{\lambda}}{\hat{\lambda} - \hat{c}^*} = \frac{\hat{\lambda}_0 \hat{n}_m}{\hat{q}\hat{n}_p}$$

- Alternative structures: (Callier & Desoer, 1982)



Linear Parameterized Controller



$$u = \begin{bmatrix} c_0 & \frac{\hat{c}(s)}{\hat{\lambda}(s)} & \frac{\hat{d}(s)}{\hat{\lambda}(s)} \end{bmatrix} \begin{bmatrix} r \\ u \\ y_p \end{bmatrix} = \underbrace{\begin{bmatrix} c_0 & c^\top & d_0 & d^\top \end{bmatrix}}_{\theta^\top} \underbrace{\begin{bmatrix} r \\ \frac{\hat{\pi}(s)}{\hat{\lambda}(s)} u \\ y_p \\ \frac{\hat{\pi}(s)}{\hat{\lambda}(s)} y_p \end{bmatrix}}_{\phi(t)}$$

State-Space Realization

Recall

$$(sI - \Lambda)^{-1} b_\lambda = \frac{1}{\hat{\lambda}(s)} \begin{bmatrix} 1 & s & \dots & s^{n-1} \end{bmatrix}^\top := \pi(s).$$

with

$$\Lambda = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -\lambda_1 & -\lambda_2 & \cdots & \cdots & -\lambda_n \end{bmatrix} \quad b_\lambda = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

We have

$$\frac{\hat{c}(s)}{\hat{\lambda}(s)} = c^\top (sI - \Lambda)^{-1} b_\lambda$$

$$\frac{\hat{d}(s)}{\hat{\lambda}(s)} = d^\top (sI - \Lambda)^{-1} b_\lambda + d_0$$

The state-space realization is

$$\dot{w}_1 = \Lambda w_1 + b_\lambda u$$
$$\dot{w}_2 = \Lambda w_2 + b_\lambda y_p$$

and the controller is

$$\boxed{u = \phi^\top(t)\theta}$$

where

$$\theta := \begin{bmatrix} c_0^* \\ c^* \\ d_0^* \\ d^* \end{bmatrix} \in \mathbb{R}^{2n}, \quad \phi(t) := \begin{bmatrix} r \\ w_1 \\ y_p \\ w_2 \end{bmatrix}.$$

- No plant model or information of θ. Instead, we use

$$u = \phi^\top(t)\hat{\theta}$$

- We need to design online identifiers and get the error equations.

Outline

Input Error vs Output Error

- In direct adaptive control, identification is designed directly identifying the *the controller parameters*

$$\theta = \mathrm{col}(c_0^*, c^*, d_0, d^*).$$

- To get linear error equation, we may define

  Output Error: The difference between the plant output y_p and the reference output

$$e_o := y_p - y_m = y_p - \hat{M}[r]$$

  Input Error:

$$\begin{aligned} e_i &:= \hat{M}^{-1}(e_o) = M^{-1}(y_p) - r \\ &:= r_p - r \end{aligned}$$

  with $r_p = \hat{M}^{-1}(\hat{P}(u))$.

Matching plant and reference model requires the matching equation:

$$\boxed{(\hat{\lambda} - \hat{c}^*)\hat{d}_p - k_p \hat{n}_p \hat{d}^* = c_0^* \frac{k_p}{k_m} \hat{\lambda}_0 \hat{n}_p \hat{d}_m}$$

$$\implies \quad 1 - \frac{\hat{c}^*}{\hat{\lambda}} - k_p \frac{\hat{n}_p}{\hat{d}_p} \frac{\hat{d}^*}{\hat{\lambda}} = c_0^* \frac{k_p \hat{n}_p}{\hat{d}_p} \frac{\hat{d}_m}{k_m \hat{n}_m}$$

$$\implies \quad \boxed{1 - \frac{\hat{c}^*(s)}{\hat{\lambda}(s)} - \frac{\hat{d}^*(s)}{\hat{\lambda}(s)} \hat{P}(s) = c_0^* \hat{M}^{-1}(s) \hat{P}(s)}$$

$$\implies \quad \hat{M}(s) = c_0^* \hat{P}(s) + \hat{M}(s) \left(\frac{\hat{c}^*(s)}{\hat{\lambda}(s)} + \frac{\hat{d}^*(s)}{\hat{\lambda}(s)} \hat{P}(s) \right)$$

where we used $\hat{\lambda}(s) = \hat{\lambda}_0(s)\hat{n}_m(s)$.

Applying to the plant input u:

$$\hat{M}[u] = c_0^* y_p + \hat{M}\left[\underbrace{\frac{\hat{c}^*(s)}{\hat{\lambda}(s)}[u] + \frac{\hat{d}^*(s)}{\hat{\lambda}(s)}[y_p]}_{\bar{\phi}^\top(t)\bar{\theta}}\right], \quad \bar{\theta} := \begin{bmatrix} c^* \\ d_0^* \\ d^* \end{bmatrix}, \ \bar{\phi}(t) := \begin{bmatrix} w_1 \\ y_p \\ w_2 \end{bmatrix}.$$

Note $\theta = \text{col}(c_0^*, \bar{\theta})$. Thus, we have an SPR regressor-like equation:

$$\boxed{y_p := \frac{1}{c_0^*}\hat{M}[u - \bar{\phi}^\top \bar{\theta}]}$$

Caveat:

- $\hat{M}(s)$ is strictly positive real (SPR) – relative degree 1
- Unknown c_0^*
- True for any u

Output Error Equation (Relative Degree 1)

Select the input u as the predefined structure

$$u(t) = \phi^\top \hat{\theta}$$

with $\hat{\theta}$ the estimate of θ.

Identifier error equation:

$$\begin{aligned} e_o = y_p - y_m &= \frac{1}{c_0^*}\hat{M}[u - \bar{\phi}^\top \bar{\theta}] - \hat{M}(r) \\ &= \frac{1}{c_0^*}\hat{M}\left[(\hat{c}_0 - c_0^*)r + \bar{\phi}^\top(\hat{\bar{\theta}} - \bar{\theta}^*)\right] \\ &= \frac{1}{c_0^*}\hat{M}[\phi^\top(t)\tilde{\theta}(t)] \end{aligned}$$

with the estimation error $\tilde{\theta} := \hat{\theta} - \theta$

Gradient identifier $\dot{\hat{\theta}} = -\gamma e_o \phi$.

Output-Error Direct MRAC (Relative Degree 1)

- Plant: $\hat{P}(s)$, relative degree 1, Assumption A1 (known direction $k_p > 0$, strictly proper, minimum phase)
- Reference signal: $r(t)$, A3: $PC[0,\infty) \cap L_\infty$
- Control Law:

$$\dot{w}_1 = \Lambda w_1 + b_\lambda u$$
$$\dot{w}_1 = \Lambda w_2 + b_\lambda y_p$$
$$u = \phi^\top(t)\hat{\theta}, \quad \phi = \mathsf{col}(r, w_1, y_p, w_2)$$

- Adaptive Law:

$$\dot{\hat{\theta}} = -\gamma e_o \phi, \quad \gamma > 0$$
$$e_o = y_p - y_m, \quad y_m = \hat{M}[r]$$

- Design parameters:
  1. Adaptation gain $\gamma > 0$
  2. Reference model: $\hat{M}(s)$ satisfying A2 (stable, minimum phase)
  3. Filter parameters: Λ, b_λ s.t. $\det(sI - \Lambda) = \hat{n}_m(s)$

Stability for OE Direct MRAC (Relative Degree 1)

Theorem

The output error direct model reference adaptive control, summarized above, guarantees that:

i) (*Lyapunov stable*) All signals in the closed-loop plant are bounded;

ii) (*Asymptotic convergence of OE*) Output tracking error e_o converges to zero asymptotically for any $r \in L_\infty$, i.e.[a]

$$\lim_{t\to\infty} |y_m(t) - y_p(t)| = 0.$$

iii) If ϕ is persistently exciting (PE), then the adaptive system is exponentially stable and

$$\lim_{t\to\infty} |\hat{\theta}(t) - \theta| = 0 \quad \text{(exp.)}.$$

[a] Attention: Not a claim on asymptotic stability!

State-Space Realization

To prove the above theorem, we need to write down the state space realization of the closed-loop dynamics.

The plant $\hat{P}(s)$ has a minimal realization $(A_p, b_p, c_p^\top)$, thus the plant and the filters are given by

$$\begin{aligned}\dot{x}_p &= A_p x_p + b_p u \\ y &= c_p^\top x_p \\ \dot{w}_1 &= \Lambda w_1 + b_\lambda u \\ \dot{w}_2 &= \Lambda w_2 + b_\lambda c_p^\top x_p\end{aligned}$$

with the control

$$u = \phi^\top(t)\hat{\theta}.$$

Define $\chi := \text{col}(x_p, w_1, w_2)$ and write compactly as

$$\dot{\chi} = A_o\chi + B_c u$$
$$y_p = C_c\chi$$

with

$$A_o := \begin{bmatrix} A_p & 0 & 0 \\ 0 & \Lambda & 0 \\ b_\lambda c_p^\top & 0 & \lambda \end{bmatrix}, \quad B_c := \begin{bmatrix} b_p \\ b_\lambda \\ 0 \end{bmatrix}$$
$$C_c := \begin{bmatrix} c_p^\top & 0 & 0 \end{bmatrix}$$

The input $= \phi^\top\hat{\theta}$, so we add and subtract the desired input $\phi^\top\theta$:

$$\dot{\chi} = A_o\chi + B_c\phi^\top\theta + B_c(\underbrace{u - \phi^\top\theta}_{\text{adaptation error}})$$
$$= A_m\chi + B_m c_o^* r + B_c(u - \phi^\top\theta)$$

Without the last error term, it would become the *reference model*!

Hurwitz Stability of A_c

That is[a]

$$\hat{M}(s) = C_m(sI - A_m)^{-1}B_m c_o^*.$$

Note that this is not a minimal realization.

- However, using the matching equation to calculate the transfer function, by avoiding cancellation, its denominator is

$$c_0^* \cdot \hat{d}_m(s)\hat{\lambda}(s)\hat{\lambda}_0(s)\hat{n}_p(s)$$

- The additional term is $\hat{\lambda}(s)\hat{\lambda}_0(s)\hat{n}_p(s)$ is stable. Hence, A_c is stable.
- The reference model (not minimal realization) is

$$\dot{\chi}_m = A_m\chi_m + B_m c_o^* r.$$

[a] The matrices (A_m, B_m, C_m) can be found in (Sastry & Bodson, page 135).

Overall Closed-Loop Dynamics

Defining the state error $e := \chi_m - \chi$ and the parameter error $\tilde{\theta} := \hat{\theta} - \theta$, and the output error $e_o := y_m - y_p$, the overall dynamics is

$$\begin{bmatrix} \dot{e} \\ \dot{\tilde{\theta}} \end{bmatrix} = \begin{bmatrix} A_m & B'_m \tilde{\phi}^\top(t) \\ -\gamma C_m & 0 \end{bmatrix} \begin{bmatrix} e \\ \tilde{\theta} \end{bmatrix}$$

$$e_o = C_m e$$

with $B'_m := B_m c_0^*$

It looks like LTV, but it is not! This is because ϕ is indeed a function of states.

Proof. (Stability of OE Direct MRAC, Relative Degree 1)

Since the reference model $\hat{M}(s)$ is SPR, using the MKY Lemma (the version of KYP lemma for non-minimal realizations): $\exists P_m = P_m^\top$, a vector q, a small constant $\rho > 0$ and $L_c = L_c^\top \succ 0$ s.t.

$$\begin{aligned} P_m A_m + A_m^\top P_m &= -qq^\top - \rho L_c \\ P_m B_m' &= C_m. \end{aligned}$$

Select the candidate Lyapunov function

$$V(\tilde{\theta}, e) = e^\top P_m e + \frac{1}{\gamma}\tilde{\theta}^\top \tilde{\theta},$$

then

$$\begin{aligned} \dot{V} &= -e^\top q q^\top e - \rho e^\top L_c e + 2e^\top P_m B_m' \phi^\top \tilde{\theta} + \frac{2}{\gamma}\tilde{\theta}\dot{\tilde{\theta}} \\ &= -e^\top q q^\top e - \rho e^\top L_c e \leq 0. \end{aligned}$$

It leads to $\tilde{\theta}, e \in L_\infty$.

Integrating $\dot{V}$ and taking $t \to \infty$,

$$V(\infty) - V(0) = \int_0^\infty \left(-e^\top q q^\top e - \rho e^\top L_c e\right) dt.$$

Since $V(\infty) := \lim_{t\to\infty} V(t)$ is bounded, we have $e \in L_2$. It is also straightforward to show $\dot{e} = A_m e + B'_m \tilde{\phi}(t)\tilde{\theta} \in L_\infty$. According to the Barbalat's lemma, the signal $e(t) \to 0$ as $t \to 0$. Note that $e_o = C_m e$, thus we have

$$\lim_{t\to\infty} |y_m(t) - y_p(t)| = 0.$$

In the identification part, we have shown that if ϕ is PE, then $\tilde{\theta} \to 0$ exponentially as $t \to 0$.[a] ∎

[a]Here, it is not practically useful to impose the PE condition on the intermediate signal ϕ. Instead, we are more interested in the conditions on $r(\cdot)$.

Strict Lyapunov Function $\star$

We are indeed interested in the stability of the system

$$\begin{bmatrix} \dot{e} \\ \dot{\tilde{\theta}} \end{bmatrix} = \begin{bmatrix} A & B_o(t) \\ -\gamma B_o(t) P & 0 \end{bmatrix} \begin{bmatrix} e \\ \tilde{\theta} \end{bmatrix}$$

If $B_o(t)$ is PE, selecting a sufficient large $\rho > 0$, we may select the following strict Lyapunov function, via Mazenc's method[a]

$$\begin{aligned} V(t, e, \tilde{\theta}) =& \rho \left(e^\top P e + \frac{1}{\gamma} |\tilde{\theta}|^2 \right) - e^\top B_o(t) \tilde{\theta} \\ & - \frac{1}{4} \tilde{\theta}^\top \left(\int_0^\infty e^{t-\tau} B_o(\tau)^\top B_o(\tau) d\tau \right) \tilde{\theta} \end{aligned}$$

satisfying

$$a_1 |\mathrm{col}(e, \tilde{\theta})|^2 \leq V(t, e, \tilde{\theta}) \leq a_2 |\mathrm{col}(e, \tilde{\theta})|^2$$

$$\dot{V}(t, e, \tilde{\theta}) \leq -a_3 |\mathrm{col}(e, \tilde{\theta})|^2$$

[a](Loria, Pantely & Maghenem, 2020)

Output Error Direct MRAC (Relative Degree >1) ⋆

- For the case of relative degree one, the output error is

$$e_o := y_p - y_m = \frac{1}{c_0^*}\hat{M}(\phi^\top \tilde{\theta}).$$

- We impose the condition on relative degree, since we need the SPR of $\hat{M}(s)$ for stability proof.
- It can be saved for higher relative degree!

Modifier SPR Output Error

$$e_o' = \frac{1}{c_0^*}\hat{M}\hat{L}\left[\bar{v}^\top \tilde{\theta} - \rho \bar{v}^\top \bar{v} e_o'\right]$$

$$\bar{v} = \hat{L}^{-1}(\bar{\phi})$$

$$\dot{\hat{\theta}} = -\gamma e_o' \bar{v} \quad \text{Modified SPR gradient}$$

$$e'_o = \frac{1}{c_0^*}\hat{M}\hat{L}\left[\bar{v}^\top\tilde{\theta} - \rho\bar{v}^\top\bar{v}e'_o\right]$$

$$\bar{v} = \hat{L}^{-1}(\bar{\phi})$$

$$\dot{\hat{\bar{\theta}}} = -\gamma e'_o\bar{v} \quad \text{Modified SPR gradient}$$

- $\hat{L}(s)$ is designed by us to make $\hat{M}(s)\hat{L}(s)$ SPR
- Its inverse $\hat{L}(s)$ is stable, minimum phase of relative degree $n - m - 1$
- We may verify

$$e'_o = \underbrace{y_p - y_m}_{e_o} + y_a$$

with the augmented error:

$$y_a := \frac{1}{c_o^*}\hat{M}\hat{L}\left[(\hat{L}^{-1}\hat{\bar{\theta}} - \hat{\bar{\theta}}^\top\hat{L}^{-1})[\bar{\phi}] + \rho\bar{v}^\top\bar{v}e'_o\right]$$

In the Input-Error method, we need not consider relative degree!

Outline

Input Error Equation: Motivation

Output Error: $e_o := y_p - y_m = y_p - \hat{M}[r]$
Input Error: $e_i := \hat{M}^{-1}(e_o) = M^{-1}(y_p) - r := r_p - r$
Different ways to obtain error equations (or linear regressors).

In OE, we use the matching equation to get

$$\boxed{(\hat{\lambda} - \hat{c}^*)\hat{d}_p - k_p \hat{n}_p \hat{d}^* = c_0^* \frac{k_p}{k_m} \hat{\lambda}_0 \hat{n}_p \hat{d}_m}$$

$$\implies \quad 1 - \frac{\hat{c}^*}{\hat{\lambda}} - k_p \frac{\hat{n}_p}{\hat{d}_p} \frac{\hat{d}^*}{\hat{\lambda}} = c_0^* \frac{k_p \hat{n}_p}{\hat{d}_p} \frac{\hat{d}_m}{k_m \hat{n}_m}$$

$$\implies \quad \boxed{1 - \frac{\hat{c}^*(s)}{\hat{\lambda}(s)} - \frac{\hat{d}^*(s)}{\hat{\lambda}(s)} \hat{P}(s) = c_0^* \hat{M}^{-1}(s) \hat{P}(s)}$$

$$\implies \quad \hat{M}(s) = c_0^* \hat{P}(s) + \hat{M}(s) \left(\frac{\hat{c}^*(s)}{\hat{\lambda}(s)} + \frac{\hat{d}^*(s)}{\hat{\lambda}(s)} \hat{P}(s) \right)$$

In IE, we apply an arbitrary input u to the second last equation:

$$u - \underbrace{\left(\frac{\hat{c}^*}{\hat{\lambda}}[u] + \frac{\hat{d}^*}{\hat{\lambda}}[y_p] \right)}_{\bar{\phi}^\top(t)\bar{\theta}} \;=\; c_0^* \hat{M}^{-1}[y_p]$$

By fixing $u = \phi^\top \hat{\theta} = \hat{c}_o r + \bar{\phi}^\top(t)\hat{\bar{\theta}}$, we have

$$\begin{aligned} \bar{\phi}^\top \bar{\theta} &= \hat{c}_0 r + \bar{\phi}\hat{\bar{\theta}} - c_0^* \hat{M}^{-1}[y_p] \\ \implies \qquad \hat{c}_0 \underbrace{\left(\hat{M}^{-1}[y_p] - r\right)}_{\text{Input error: } e_i} &= \underbrace{\left[\hat{M}^{-1}[y_p] \quad \bar{\phi}^\top\right]}_{\psi_1^\top(t)} \tilde{\theta} \end{aligned}$$

(Preliminary) Input Error Equation

$$e_i = \frac{1}{\hat{c}_o} \psi_1^\top(t) \tilde{\theta}$$

Input Error Equation: Motivation

This is the motivation of IE equations, but it is not implementable:

- Relative degree of $\hat{M}(s)$ is at least 1, so its inverse is not proper. (Make sense in analysis but not implementable)
- In order to get the IE error equation, we fixed $u = \phi^\top \hat{\theta}$. This is not crucial, but can be avoided to decouple identification and control.
- We are careless about the initial condition, since the plant $\hat{P}(s)$ may be unstable.

Implementable Input-Error Equation

- $\hat{M}(s)$ is minimum phase with relative degree $n-m$, so we may select any stable minimum phase TF $\hat{L}^{-1}$ of relative degree $n-m$, then

  $$\boxed{(\hat{M}\hat{L})^{-1} \text{ is proper and stable.}}$$

  and

  $$\boxed{\hat{L}^{-1}[r_p] \;=\; (\hat{M}\hat{L})^{-1}[y_p]}$$

  which is implementable.
- In the motivating case, we have

  $$1 - \frac{\hat{c}^*(s)}{\hat{\lambda}(s)} - \frac{\hat{d}^*(s)}{\hat{\lambda}(s)}\hat{P}(s) = c_0^*\hat{M}^{-1}(s)\hat{P}(s)$$

  Applying $\hat{L}^{-1}(s)[\cdot]$,

  $$\hat{L}^{-1}(s) - \hat{L}^{-1}(s)\frac{\hat{c}^*(s)}{\hat{\lambda}(s)} = \left[\hat{L}^{-1}(s)\frac{\hat{d}^*(s)}{\hat{\lambda}(s)} + c_0^*(\hat{M}(s)\hat{L}^{-1}(s))^{-1}\right]\hat{P}(s)$$

- We are able to use $\hat{L}^{-1}$ to deal with the unstable poles of $\hat{P}(s)$.
- Applying to an arbitrary input $u(t)$, we have

$$\hat{L}^{-1}\frac{\hat{d}^*}{\hat{\lambda}}[y_p] + c_0^*(\hat{M}\hat{L})^{-1}[y_p] = L^{-1}[u] - \hat{L}^{-1}\frac{\hat{c}^*}{\hat{\lambda}}[u] + \epsilon(t)$$

- Since $\theta := \text{col}(c_0^*, \bar{\theta})$ is constant, we may take it outside the TF:

$$\begin{aligned}\hat{L}^{-1}[\bar{\phi}^\top]\bar{\theta} &= \hat{L}^{-1}[\bar{\phi}^\top\bar{\theta}] \\ &= \hat{L}^{-1}\left[\frac{\hat{c}^*}{\hat{\lambda}}[u] - \frac{\hat{d}^*}{\hat{\lambda}}[y_p]\right] \\ &= \hat{L}^{-1}[u] - c_0^*(\hat{M}\hat{L})^{-1}[y_p] + \epsilon(t)\end{aligned}$$

Implementable Input-Error Regressor

$$\hat{L}^{-1}[u] = v^\top(t)\theta + \epsilon(t),$$

with

$$v^\top(t) := \begin{bmatrix}\hat{L}^{-1}(r_p) & \hat{L}^{-1}(\bar{\phi}^\top)\end{bmatrix}$$

Input Error Identifier Structure

We define the modified input error

$$e_o := v^\top(t)\hat{\theta} - \hat{L}^{-1}(u)$$

which satisfies the linear error equation

$$e_o = v^\top(t)\tilde{\theta} + \epsilon(t).$$

- All signals here are available.
- When obtaining the linear error equation, we do not fix $u = \phi^\top\hat{\theta}$ – decoupling identification and control!
- Useful in the presence of actuator saturation.

Input-Error Direct MRAC

- Plant: $\hat{P}(s)$, ~~relative degree 1~~, Assumption A1 (known direction $k_p > 0$, strictly proper, minimum phase) and A4

> A4 **Bound on the High-Frequency Gain k_p**
> Assume that an upper bound on k_p is known, i.e.
>
> $$k_p \leq k_{\mathtt{max}}.$$

- Reference signal: $r(t)$, A3: $PC[0, \infty) \cap L_\infty$
- Control Law:

$$\dot{w}_1 = \Lambda w_1 + b_\lambda u$$
$$\dot{w}_1 = \Lambda w_2 + b_\lambda y_p$$
$$u = \phi^\top(t)\hat{\theta}, \quad \phi = \text{col}(r, w_1, y_p, w_2)$$

- Identifier Structure:

$$v^\top(t) := \left[(\hat{M}\hat{L})^{-1}[y_p] \quad \hat{L}^{-1}(w_1^\top) \quad \hat{L}^{-1}[y_p] \quad \hat{L}^{-1}[w_2^\top]\right]$$

$$e_i := v^\top(t)\hat{\theta} - \hat{L}^{-1}[u] \qquad \text{(Input Error)}$$

- Adaptive Law (Normalized Gradient with Projection):

$$\dot{\hat{\theta}} = \begin{cases} -\gamma \dfrac{e_i v}{1+\alpha|v|^2} & \text{otherwise} \\ 0 & \text{if } \dot{\hat{c}}_0 < 0 \text{ or } \hat{c}_0 = \frac{k_m}{k_{\max}} \end{cases}$$

$$e_i = y_p - y_m, \quad y_m = \hat{M}[r]$$

- Design parameters:
  1. Adaptation gain $\gamma > 0$
  2. Normalization parameter $\alpha > 0$
  3. Reference model: $\hat{M}(s)$ satisfying A2 (stable, minimum phase)
  4. Filter parameters: $\Lambda,\ b_\lambda$ s.t. $\det(sI - \Lambda) = \hat{n}_m(s)$
  5. $\hat{L}^{-1}$ stable, minimum phase TF of relative degree $n - m$

Normalization & projection are important for stability!

Stability of IE Direct MRAC

Theorem

Consider the input error direct MRAC described above, with initial condition in an arbitrary B_h. Then,

i) (*Lyapunov stable*) All internal states are bounded;

ii) (*Asymptotic convergence of OE*) Output tracking error e_o converges to zero asymptotically for any $r \in L_\infty$, i.e.[a]

$$\lim_{t \to \infty} |y_m(t) - y_p(t)| = 0.$$

iii) If v is persistently exciting, then the adaptive system is exponentially stable and

$$\lim_{t \to \infty} |\hat{\theta}(t) - \theta| = 0 \quad \text{(exp.)}.$$

[a] Attention: Not a claim on asymptotic stability!

Sketch of Proof

Full proof can be found in (Sastry & Bodson, page 143) and (Ioannou & Sun, page 390).

Step 1. Regressor bound & existence of solutions

From the properties of the projected, normalized identifier, we have

$$|v^\top \tilde{\theta}| = |\beta(t)| \|v(\cdot)\|_\infty + |\beta(t)|$$

$$\beta := \frac{v^\top \tilde{\theta}}{1 + \|v(\cdot)\|_\infty} \in \ L_2 \cap L_\infty$$

$$\tilde{\theta} \in L_\infty, \quad \dot{\tilde{\theta}} \in \ L_2 \cap L_\infty$$

$$|\tilde{\theta}(t)| \leq |\tilde{\theta}(0)|$$

$$\hat{c}_0 \geq c_{\mathtt{min}} > 0 \qquad \text{(due to projection)}$$

Similar to the OE case, the closed-loop dynamics can be "viewed" as an LTI system with an LTV controller and bounded feedback gains – solutions are well-defined.

Sketch of Proof

Step 2. Express the plant input & output in terms of control error $\phi^\top \tilde{\theta}$

$$y_p = y_m + \frac{1}{c_0^*}\hat{M}(\phi^\top \tilde{\theta})$$

$$u = \hat{P}^{-1}\hat{M}\left(r + \frac{1}{c_0^*}\phi^\top \tilde{\theta}\right)$$

Define an auxiliary signal $m_f^2 := 1 + \|u\|^2 + \|y_p\|^2$ with L_{2e}-norm in some interval, we have

$$m_f \leq c + c\|\phi^\top \tilde{\theta}\|$$

Using the Swapping Lemma for TF, we have

$$\|\phi^\top \tilde{\theta}\| \leq \frac{c}{\alpha_0} m_f + c\alpha_0^{*n}\|\tilde{g} m_f\|$$

with $\tilde{g}^2 := \epsilon^2 n_s^2 + |\dot{\tilde{\theta}}|^2 + \epsilon^2 \in L_2$. (Tricky for this step)

(*Swapping Lemma*) Let $\phi, \tilde{\theta} : \mathbb{R}_+ \to \mathbb{R}^n$ and $\dot{\tilde{\theta}} \in C^1$. Let $\hat{H}(s)$ be a proper rational TF. If $\hat{H}(s)$ is stable, with a minimal realization

$$\hat{H}(s) = c^\top (sI - A)^{-1} b + d,$$

then

$$\hat{H}(\phi^\top \tilde{\theta}) = \hat{H}(\phi^\top)\tilde{\theta} + \hat{H}_c(\hat{H}_b(\phi^\top)\dot{\tilde{\theta}})$$

with

$$\hat{H}_b = (sI - A)^{-1} b, \quad \hat{H}_c = -c^\top (sI - A)^{-1}.$$

Sketch of Proof

Step 3. Apply the Bellman-Gronwall Lemma to establish boundedness.

From the last step, we may get

$$m_f^2 \leq c + c\int_0^t \alpha_0^{*2n}\tilde{g}(\tau)^2 m_f^2(\tau)d\tau.$$

Using the B-G lemma and $g \in L_2$, we have $m_f \in L_\infty$. Then, considering all the transfer functions, we can show that all signals are bounded.

Step 5. Show tracking error $e_o := y_p - y_m \to 0$.

Check $e_o \in L_2$ and $\dot{e}_o \in L_\infty$ and apply the Barbalat's lemma.

Step 4. Verifying $\tilde{\theta} \to 0$ if v is PE.

Comparison: IE vs OE

- Traditional starting point in MRAC is to study $e_o := y_p - y_m$
- Stability proof in OE-MRAC requires the SPR condition of $\hat{P}(s)$
  - Limited to the systems with relative degree 1
  - Otherwise using the augmented error (complicated and non-robust)

  IE-MRAC does not have this issue.
- Error term:
  - OE error equation relies on the input equal to $u = \phi^\top(t)\hat{\theta}$
  - IE error equation can use arbitrary inputs (still work under saturation)

  Decoupling identifier and controller (possible to replace the gradient estimator by other estimators)

Outline

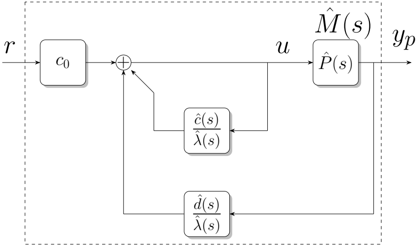
Indirect Adaptive Control

- Use the identifier in the last lecture to estimate plant parameters

$$k_p, \hat{n}_p(s), \hat{d}_p(s)$$

- Use the matching equation to compute the controller parameters

$$c_0, \hat{c}(s), \hat{d}(s)$$



Recall the plant

$$\frac{\hat{y}_p(s)}{\hat{u}(s)} = \hat{P}(s) = \frac{\alpha_n s^{n-1} + \cdots + \alpha_1}{s^n + \beta_n s^{n-1} + \cdots + \beta_1}$$

Introduce a monic n-th order Hurwitz (but arbitrary) polynomial

$$\hat{\lambda}(s) = s^n + \lambda_n s^{n-1} + \cdots + \lambda_1.$$

Then, from

$$\frac{\hat{y}_p}{\hat{u}} := k_p \frac{\hat{n}_p}{\hat{d}_p} \implies k_p \hat{n}_p(s)\hat{u}(s) - \hat{d}_p(s)\hat{y}_p(s) = 0$$

$$\implies \hat{\lambda}(s)\hat{y}_p(s) = k_p \hat{n}_p(s)\hat{u}(s) + (\hat{\lambda}(s) - \hat{d}_p(s))\hat{y}_p(s).$$

Or equivalently

$$\hat{y}_p(s) = \frac{\alpha_n s^{n-1} + \cdots + \alpha_1}{\hat{\lambda}(s)} \hat{r}(s) + \frac{(\lambda_n - \beta_n) s^{n-1} + \cdots + (\lambda_1 - \beta_1)}{\hat{\lambda}(s)} \hat{y}_p(s),$$

with the associated the plant parameters

$$a = \begin{bmatrix} \alpha_1 & \dots & \alpha_n \end{bmatrix}, \quad b = \begin{bmatrix} \lambda_1 - \beta_1 & \dots & \lambda_n - \beta_n \end{bmatrix}.$$

Linear Regressor

Using the filters

$$\begin{aligned} \dot{w}_1 &= \Lambda w_1 + b_\lambda u \\ \dot{w}_2 &= \Lambda w_2 + b_\lambda y_p \quad \text{(Shared by Identifier \& Controller)} \end{aligned}$$

We have the linear regressor on unknown parameters $[a^\top \quad b^\top]^\top$.

For simplicity, we define

- Plant parameter $\eta := [a_1, \ldots, a_{m+1}, 0, \ldots, b_1, \ldots, b_n]^\top$
- "Perfect" control parameter
  $\theta := \text{col}(c_0^*, \bar{\theta}) := [c_0^*, c_1, \ldots, c_n, d_1, \ldots, d_n]^\top$

Attention:

1. No d_0 term here!
2. Since the relative degree of $\hat{P}(s)$ is known, we needn't estimate $a_{m+2}, \ldots$

Linear Regressor (cont'd)

$$y_p = \phi_w^\top(t)\eta$$

with

$$\phi_w = \text{col}(w_{1,1}, \ldots, w_{1,m}, 0, \ldots, 0, w_2)$$

Indirect MRAC

- Plant: $\hat{P}(s)$, ~~relative degree 1~~, Assumption A1 (known direction $k_p > 0$, strictly proper, minimum phase) and A4

> A7 Bound on the High-Frequency Gain k_p
> Assume that an upper bound on k_p is known, i.e.
>
> $$k_p \geq k_{\mathtt{min}} > 0.$$

- Reference signal: $r(t)$, A3: $PC[0, \infty) \cap L_\infty$
- Control Law:

$$\begin{aligned} \dot{w}_1 &= \Lambda w_1 + b_\lambda u \\ \dot{w}_1 &= \Lambda w_2 + b_\lambda y_p \\ u &= \phi^\top(t)\hat{\theta}, \quad \phi = \text{col}(r, w_1, w_2) \text{ (No } y_p\text{!)} \end{aligned}$$

Be careful about the dimensionality.

- Identifier Structure:

$$\phi_w^\top = \text{col}(w_{1,1}, \ldots, w_{1,m}, 0, \ldots, w_2)$$
$$e = \phi^\top \hat{\eta} - y_p$$

- Adaptive Law (Normalized Gradient with Projection):

$$\dot{\hat{\eta}} = -\gamma \frac{e\phi_w}{1 + \rho|\phi_w|^2}$$

If $\hat{a}_{m+1} = k_{\mathtt{min}}$ and $\dot{\hat{a}}_{m+1} < 0$, then let $\dot{\hat{a}}_{m+1} = 0$.

- Design parameters:
  1. Adaptation gain $\gamma > 0$
  2. Normalization parameter $\rho > 0$
  3. Reference model: $\hat{M}(s)$ satisfying A2 (stable, minimum phase)
  4. Filter parameters: $\Lambda,\ b_\lambda$ s.t. $\det(sI - \Lambda) = \hat{\lambda}(s)$ and $\hat{\lambda} = \hat{\lambda}_0 \hat{n}_m$

- Translating Identifier Parameter $\rightarrow$ Control Parameter

  Define $\hat{q} := \frac{\hat{\lambda}_0 \hat{d}_m}{(\hat{\lambda} - \hat{b})}$:

$$\hat{c} = \hat{\lambda} - \frac{1}{\hat{a}_{m+1}} \hat{q}\hat{a}$$

$$\hat{d} = \frac{1}{a_{m+1}} (\hat{q}\hat{\lambda} - \hat{q}\hat{b} - \hat{\lambda}_0 \hat{d}_m)$$

$$\hat{c}_0 = \frac{k_m}{a_{m+1}}$$

Stability of Indirect MRAC

Theorem

Consider the indirect MRAC described above, with initial condition in an arbitrary B_h. Then,

i) (*Lyapunov stable*) All internal states are bounded;

ii) (*Asymptotic convergence of OE*) Output tracking error e_o converges to zero asymptotically, i.e.

$$\lim_{t \to \infty} |y_m(t) - y_p(t)| = 0,$$

and the regressor error converges to zero.

See (Sastry & Bodson, page 341) for proof.

Why need the sign and upper bound of k_p?

In MRAC, we need to identify c_0^* or $\frac{1}{c_0^*}$:

- If we identify $\frac{1}{c_0^*}$ (as done in *indirect* MRAC), the input $u = \hat{c}_0 r + \bar{\phi}^\top(t)\hat{\bar{\theta}}$ will be unbounded if $\frac{1}{\hat{c}_0} \to 0$

  The sign and lower bound on $\frac{1}{c_0^*}$ (i.e. those of k_p) help us avoid the zero-cross issue with projection.

- If we identify c_0 (in direct MRAC), then

$$\hat{c}_0 = 0,\ \hat{\bar{\theta}} = 0 \implies u = 0,\ e_i = 0$$

  No adaptation will occur $\dot{\hat{\theta}} = 0$, although $y_p - y_m$ does not tend necessarily to zero and may even be unbounded.

To avoid this above issues, we require the sign and upper bound of k_p.

Discussions

- Alternate Model Reference Schemes: Flexible by combining various identification and control structures [1]

- Adaptive Pole Placement Control: Choose a particular reference model

$$\hat{M}(s) = k_m \frac{\hat{n}_p(s)}{\hat{d}_m(s)}, \quad \text{i.e. } \hat{n}_m(s) = \hat{n}_p(s).$$

Only assigning closed-loop poles – $\hat{n}_p(s)$ is replaced by its estimate in the reference model $\hat{M}(s)$. The matching equation becomes a *Diophantine* equation

$$(\hat{\lambda} - \hat{c}^*)\hat{d}_p - k_p \hat{n}_p \hat{d}^* = \left(c_0^* \frac{k_p}{k_m} \right) \hat{\lambda} \hat{d}_m,$$

which is not always solvable, unlike MRAC.

[1] G.C. Goodwin & D.Q. Mayne. A parameter estimation perspective of continuous time model reference adaptive control, *Automatica* 23.1 (1987): 57-70.

What have we learned?

- Model reference adaptive control
  - Reference model
  - Similar control structure
  - Different identifiers
- Direct MRAC
  - Input error direct MRAC
  - Output error direct MRAC (relative degree 1)
- Indirect MRAC
- Stability:
  - Asymptotic convergence for general reference $r(t)$;
  - Exponential stability with PE.
  - Same stability and convergence properties for these three schemes
- IE Direct MRAC and Indirect MRAC are attractive:
  - Linear error equations
  - SPR conditions
  - Decoupling between identification and control

ELE6214 - Commande de Systèmes Incertains

Lecture 5: Robustness of Adaptive Systems

Bowen Yi

bowen.yi@polymtl.ca

Polytechnique Montréal

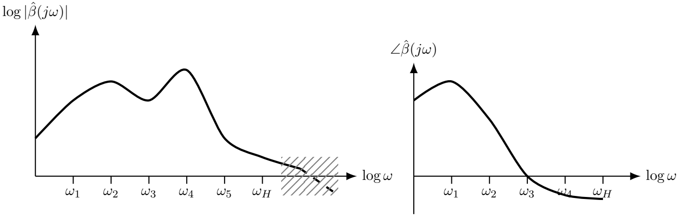
Hiver 2025

Outline

Parametric and Unstructured Uncertainty

- Designers do not have a detailed model:
  - Too complex
  - Not completely understood of its dynamics
  - Model reduction

- For *stable* systems, we may obtain Bode diagrams.
  - Data beyond a certain frequency ω_H is unreliable – measurements are poor (noise)
  - Referred as *high-order dynamics* – wish to neglect



Adaptive Control vs (Non-adaptive) Robust Control

Assume the *goal* is the following: select a reference model $\hat{M}(s)$ such that the plant output $y_p(t)$ tracks the reference output $y_r(t)$.

Non-Adaptive Robust Control: Use poor data at high-frequencies to get a nominal model $\hat{P}^*(s)$. The actual plant $\hat{P}(s)$ satisfies

$$\hat{P}(s) = \hat{P}^*(s) + \hat{H}_a(s) \qquad \text{(Additive uncertainty)}$$

or

$$\hat{P}(s) = \hat{P}^*(s)[1 + \hat{H}_m(s)] \qquad \text{(Multiplicative uncertainty)}$$

- $|\hat{H}_a(j\omega)$ and $|\hat{H}_m(j\omega)|$ are unknown but bounded.
- Design an LTI controller (feedforward + feedback) to match the reference model $\hat{M}(s)$ over the frequency range of interest
- At least preserve stability and reduce sensitivity

Adaptive Control vs (Non-adaptive) Robust Control

Adaptive Control:

- Designer distinguishes the parametric uncertainty in the pole/zero locations and unstructured uncertainty.

$$\hat{P}(s) = \hat{P}_\theta(s) + \hat{H}_{au}(s)$$

or

$$\hat{P}(s) = \hat{P}_\theta(s)[1 + \hat{H}_{mu}(s)]$$

Plant model $\hat{P}_\theta(s)$ – still unstructured uncertainty $\hat{H}_{au}$ & $\hat{H}_{mu}$

- Identify the pole-zero locations on-line – during operation
- Better match to $\hat{M}(s)$ but yielding nonlinear time-varying control
- Added complexity is made worthwhile when non-adaptive control has unsatisfactory performance

Adaptive Control vs (Non-adaptive) Robust Control

Example

Unstable plant:

$$\hat{P}(s) = \frac{m}{(s-1+\epsilon)(s+m)}$$

with $\epsilon > 0$ small and $m > 0$ large.

- Robust Control: Select nominal model $\hat{P} = \frac{1}{s-1}$, with uncertainty

$$\hat{H}_m(s) = \frac{-s^2 + s - \epsilon(s+m)}{(s-1+\epsilon)(s+m)} \qquad \text{(unstable)}$$

- Adaptive Control: Parameterized nominal model $\hat{P}_\theta(s) = \frac{1}{s+\theta}$ ($\theta = -1+\epsilon$ unknown), with uncertainty

$$\hat{H}_{mu}(s) = -\frac{s}{s+m} \qquad \text{(stable)}$$

Robustness of Adaptive Systems

How will the adaptive algorithms behave with the true plant $\hat{P}(s)$?

How can we maintain stability in the presence of uncertainties?

Outline

A Young PhD's Story

Not to mention lively sessions at conferences

> *"I went to a fight, and an adaptive control session broke out!"*
>
> — Bob Bitmead

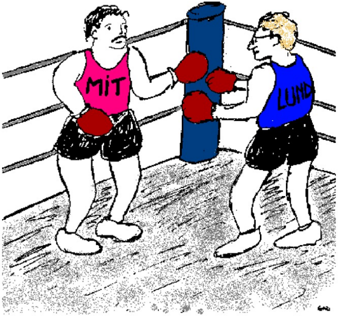
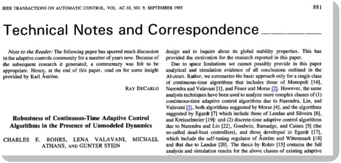


Figure: 1981 IEEE CDC, San Diego

"Earlier versions of the above paper have been presented at several conferences since 1980. These presentations have certainly contributed towards making the sessions on adaptive control at the CDC conferences lively and fun. The discussions have also inspired a lot of work on robustness of adaptive systems which have significantly contributed to our understanding of such systems.

...

I would like to thank you, Charles, for sticking your neck out as a young Ph.D. and challenging 'the adaptive establishment.' ..." — Karl J. Astrom



IEEE TRANSACTIONS ON AUTOMATIC CONTROL, VOL. AC-30, NO. 9, SEPTEMBER 1985 881

Technical Notes and Correspondence

Note to the Reader: The following paper has spurred much discussion in the adaptive controls community for a number of years now. Because of the subsequent research it generated, a commentary was felt to be appropriate. Hence, at the end of this paper, read on for some insight provided by Karl Åström.

RAY DECARLO

Robustness of Continuous-Time Adaptive Control Algorithms in the Presence of Unmodeled Dynamics

CHARLES E. ROHRS, LENA VALAVANI, MICHAEL ATHANS, AND GUNTER STEIN

design and to inquire about its global stability properties. This has provided the motivation for the research reported in this paper.

Due to space limitations we cannot possibly provide in this paper analytical and simulation evidence of all conclusions outlined in the Abstract. Rather, we summarize the basic approach only for a single class of continuous-time algorithms that includes those of Monopoli [14], Narendra and Valavani [1], and Feuer and Morse [2]. However, the same analysis techniques have been used to analyze more complex classes of (1) continuous-time adaptive control algorithms due to Narendra, Lin, and Valavani [3], both algorithms suggested by Morse [4], and the algorithms suggested by Egardt [7] which include those of Landau and Silveira [6], and Kreisselmeier [19]: and (2) discrete-time adaptive control algorithms due to Narendra and Lin [22], Goodwin, Ramadge, and Caines [5] (the so-called dead-beat controllers), and those developed in Egardt [17], which include the self-tuning regulator of Åström and Wittenmark [18] and that due to Landau [20]. The thesis by Rohrs [15] contains the full analysis and simulation results for the above classes of existing adaptive

The Rohrs Example

Consider a first-order plant

$$\hat{P}_\theta(s) = \frac{k_p}{s + a_p}, \quad k_p = 2,\ a_p = 1$$

and select the SPR reference model

$$\hat{M}(s) = \frac{3}{s+3}$$

Matching control parameters $c_0^* = \frac{k_m}{k_p} = 1.5,\ d_o^* = \frac{a_p - a_m}{k_p} = -1$

Output Error MRAC

$$u = \hat{c}_0 r + \hat{d}_0 y_p$$

$$e_o = y_p - y_m$$

$$\dot{\hat{c}}_0 = -\gamma e_o r$$

$$\dot{\hat{d}}_0 = -\gamma e_o y_p$$

The Rohrs Example (cont'd)

Real Plant

Actual plant and nominal model not exactly matched:

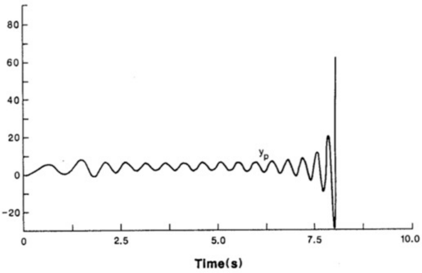
$$\hat{P}(s) = \frac{2}{s+1} \cdot \underbrace{\frac{229}{s^2 + 30s + 229}}_{\text{unmodeled dynamics}}$$

- Poles of uncertainty: $-15 \pm 2j$
- Approximately equal to 1 at low frequency
- Unmodeled dynamics is well-damped, stable (Traditional view: it should be innocuous)

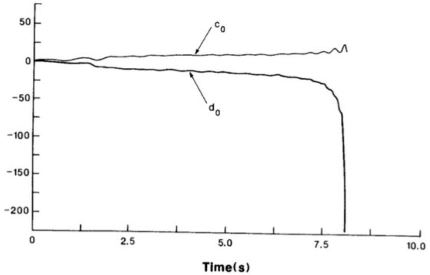
The Rohrs Example: Simulation 1

A large constant reference input $r(t) = 4.3$ and noise free

Plant output y_p



Controller paremater estimates

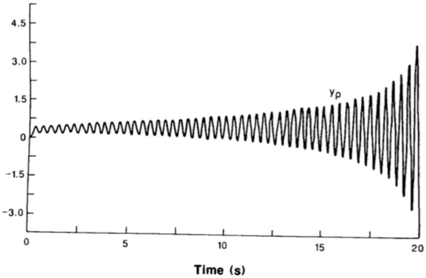


The output error initially converges to zero, but eventually diverges to infinity.

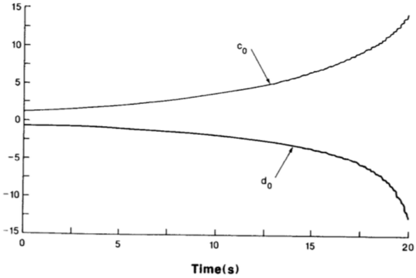
The Rohrs Example: Simulation 2

Reference $r(t) = 0.3 + 1.85 \sin 16.1t$ has a small constant and large high-frequency component, and noise free

Plant output y_p



Controller paremater estimates

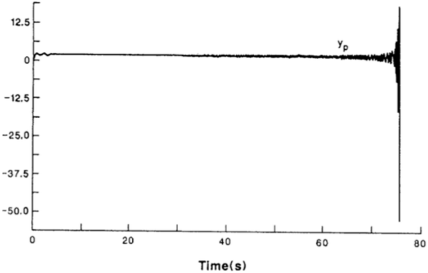


The output error diverges at first slowly, and then more rapidly to infinity.

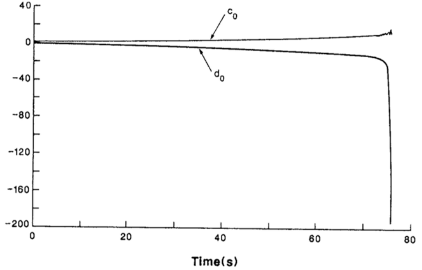
The Rohrs Example: Simulation 3

Constant reference input $r = 2$ and a small output disturbance $n = 0.5 \sin 16.1t$

Plant output y_p



Controller paremater estimates



The output error initially converges to zero. After staying in the neighborhood of zero for an extend period of time, it diverges to infinity.

A Lesson from Rohrs Examples

- The Rohrs examples stimulated much research about the robustness of adaptive systems.
- It shows the bounded-input-bounded-state stability properties for the MRAC is not robust to uncertainties. Even an arbitrary small disturbance can destabilize an adaptive system.
- Instability mechanisms are related to the identifier.

Outline

Nominal and Perturbed Systems

We consider properties of the nominal (or unperturbed) system

$$\dot{x} = f(t, x, 0), \quad x(0) = x_0,$$

and relate to properties of the perturbed system

$$\dot{x} = f(t, x, u), \quad x(0) = x_0.$$

Roughly, exponential stability of the nominal system implies the robustness of the perturbed system w.r.t. external perturbation u.

Theorem (Small Signal I/O Stability)

Consider the perturbed system $\dot{x} = f(t, x, u)$ and the unperturbed system $\dot{x} = f(t, x, 0)$ that has a zero equilibrium. Assume $f \in C^1$ and Lipschitz w.r.t. x for $x \in B_h,\ u \in B_c$, and $u \in L_\infty$.

If $x = 0$ is exponentially stable for the unperturbed system. Then, the perturbed system

- is small-signal L_∞ stable, i.e. $\exists \gamma_1, c_1 > 0$ s.t. for $\|u\|_\infty < c_1$

$$\|x\|_\infty \leq \gamma_1 \|u\|_\infty < h.$$

- $\exists m \geq$ s.t. $\forall |x_0| < \frac{h}{m}$, $0 < \|u\|_\infty < c_1$ implies

$$\lim_{t \to \infty} \text{dist}(x(t), B_\delta) = 0, \quad \delta := \gamma_1 \|u\|_\infty < h.$$

Also consider the tool of input-to-state (ISS) stability.

Sketch of Proof

- Invoking the converse Lyapunov theorem: for exponential stability $\exists$ Lyapunov function $V(t,x)$ s.t.

$$\alpha_1|x|^2 \le V(t,x) \le \alpha_2|x|^2$$
$$\frac{\partial V}{\partial t} + \frac{\partial V}{\partial x} f(t,x,0) \le -\alpha_3|x|^2$$
$$\left|\frac{\partial V(t,x)}{\partial x}\right| \le \alpha_4|x|.$$

- Lie derivative along the perturbed system:

$$\begin{aligned}\dot{V} &\le -\alpha_3|x|^2 + \frac{\partial V}{\partial x}(t,x)[f(t,x,u) - f(t,x,0)] \\ &\le -\alpha_3|x|^2 + \alpha_4\ell_u|x|\|u\|_\infty \\ &\le -\frac{1}{2}\alpha_3|x|^2 + \frac{(\alpha_4\ell_u)^2}{2\alpha_3}\|u\|_\infty^2\end{aligned}$$

- Intuitively, a large $|x|$ will make $\dot{V} < 0$.

Exponential Stability and Robustness of Adaptive Systems

- Let's consider the output error MRAC with relative degree 1. The closed-loop dynamics is the nonlinear (bilinear) time-varying system

$$\begin{bmatrix} \dot{e} \\ \dot{\tilde{\theta}} \end{bmatrix} = \begin{bmatrix} A_m & B_m \tilde{\phi}^\top(t) \\ -\gamma C_m & 0 \end{bmatrix} \begin{bmatrix} e \\ \tilde{\theta} \end{bmatrix}$$

- We have learned that if $\tilde{\phi}(t)$ is PE, then the above system is exponentially stable.
- This means that if the reference signal is sufficiently exciting, we may achieve robust performance.
- We will consider the robustness of OE-MRAC w.r.t.
  - Output measurement noise
  - Unmodeled dynamics

Robustness of OE-MRAC to Noise

- y_p^* - output of the plant $\hat{P}_\theta(s)$; y_p - measured output affected by noise

$$y_p(t) = y_p^*(t) + n(t) = \hat{P}_\theta[u] + n(t)$$

- In OE-MRAC, the following terms are affected by $n(t)$:
  - 2nd part in the filter

$$\dot{w}_2 = \Lambda w_2 + b_\lambda y_p^* + b_\lambda n$$

  - Update law

$$\dot{\tilde{\theta}} = -\gamma(y_p^* + n - y_m)\phi = -\gamma c_m^\top e\phi - \gamma n\phi$$

  - Regression vector

$$\phi = \begin{bmatrix} r \\ w_1 \\ y_p^* \\ w_2 \end{bmatrix} = w_\star + q_n n, \quad q_n := \begin{bmatrix} 0 \\ 0 \\ n \\ 0 \end{bmatrix}$$

Robustness of OE-MRAC to Noise

Now, the error dynamics becomes

$$\dot{\chi} = \underbrace{f(t,\chi)}_{\text{Nominal part}} + p_1(t,n) + P_2(t,n)\chi(t)$$

See (Sastry & Bodson, page 228) for the formulas of $p_1(\cdot), P_2(\cdot)$, s.t.

$$n \in L_\infty \quad \Longrightarrow \quad \|p_1\|_\infty + \|P_2\|_\infty \le k_n \|n\|_\infty.$$

If the noise $n \in L_\infty$, and $\tilde{\phi}$ is PE, then $\exists \gamma_n, c_n > 0$ and $m \ge 1$, s.t. $\|n\|_\infty < c_n$ and $|x(0)| < \frac{h}{m}$ implies

$$\lim_{t\to\infty} \text{dist}(\chi(t), B_\delta) = 0, \quad \delta = \gamma_n \|n\|_\infty,$$

and $|x(t)| \le m|x_0| < h$ for all $t \ge 0$.

Robustness of OE-MRAC to Unmodeled Dynamics

Assume the existence of some additive uncertainty

$$y_p(t) = y_p^*(t) + \underbrace{H_a[u(t)]}_{:=\Delta(t)}$$

satisfying

$$\|H_a[u]_t\|_\infty \leq \gamma_a \|u_t\|_\infty + \beta_a, \quad \forall t \geq 0.$$

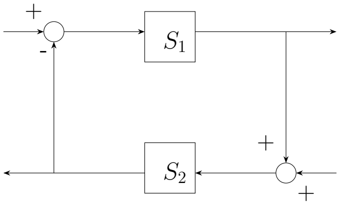
- The perturbation Δ affects the plant input u (using u^* to present the nominal case):

$$u = u^* + \theta^\top q_n \Delta + \tilde{\theta}^\top q_n n \implies \|u_t\|_\infty \leq \gamma_u \|\Delta_t\|_\infty + \beta_u$$

Small gain theorem:

$$\gamma_a \gamma_u < 1$$

$$\frac{\beta_a + \gamma \beta_u}{1 - \gamma_a \gamma_u} < c_n$$



Theorem (Robustness to Unmodeled Dynamics)

Consider the OE direct MRAC with relative degree 1. If the additive perturbation H_a satisfies the above assumptions.

If $\tilde{\phi}$ is PE, then, for x_0, γ_a, β sufficiently small, the overall states

$$\chi \in L_\infty.$$

We will learn the Small Gain theorem in the robust control part.

Outline

Rohrs Example 1: High-Gain Identifier Instability

Example 1 uses a large reference input $r(t) = 4.3$, which is related to the high-gain identifier instability.

- Adaptation law:

$$\dot{\hat{\theta}} = -\gamma e_o \begin{bmatrix} r \\ y_p \end{bmatrix}.$$

- Although not directly using a large $\gamma > 0$, multiplying r by 2 means twice of y_m, y_p and r, or equivalently multiplying the gain by 4.
- Applying high-gain to LTI systems with relative degree > 2 yields instability
- Can be simply fixed using a normalized algorithm

Rohrs Example 2: Exciting at High-Frequency

Example 2 uses a sufficiently exciting reference $r = 0.3 + 1.85 \sin 16t$. Without unmodeled dynamics, $\exists!$ values of c_0^*, d_0^* to match $r \to y_p$. With unmodeled dynamics, there still exists *unique* values of c_0', d_0' but at the high-frequency ω_0, i.e.

$$\left. \frac{458c_0'}{(s+1)(s^2+30s+229) - 458d_0'} \right|_{j\omega_0} = \left. \frac{3}{s+3} \right|_{j\omega_0}$$

For this case, c_0', d_0' depends on $\hat{P}(s), \hat{M}(s)$ and also the reference r.

By attempting to match the reference model at a high frequency, the adaptive system leads to an unstable closed-loop system.

Rohrs Example 3: Slow Drift Instability / Insufficiently Rich

- Example 3 uses the reference not sufficiently rich (thus $\tilde{\phi}$ not PE). We have shown that for the PE case, the adaptive system is robust to noise.
- Rohrs example with no unmodeled dynamics and $\hat{\theta}$ fixed

$$\frac{\hat{y}_p(s)}{\hat{r}(s)} = \frac{2\hat{c}_0}{s + 1 - 2\hat{d}_0}$$

If a constant reference r is used, then the transfer function should be matched with the DC gain of $\hat{M}(s)$, i.e.

$$\frac{2\hat{c}_0}{1 - 2\hat{d}_0} = 1.$$

$\exists$ infinite numbers of feasible $\hat{c}_0$ and $\hat{d}_0$ satisfying the above, thus

$$\lim_{t\to\infty} |y_p(t) - y_m(t)| = 0.$$

- If the noise $n(t)$ appears in y_p, then $\hat{c}_0, \hat{d}_0$ would move along the line

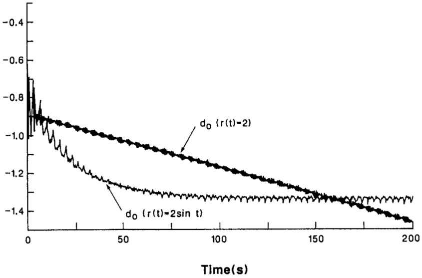
$$\frac{2\hat{c}_0}{1 - 2\hat{d}_0} = 1$$

leaving $e_o = y_p - y_m$ at zero.

- Part of adaptive law becomes

$$\dot{\hat{d}}_0 = -\gamma y_p^*(y_p^* - y_m) - \gamma y_m n - \gamma n^2$$

slowly drifting $\hat{d}_0$ toward the negative direction



Outline

Robustness Enhancement

PE relies on the reference $r(t)$ – not realistic for many scenarios.

Intuitive Ideas

- The Usage of Prior Information

  If the plant is fairly well modeled, except for a few unknown/uncertain components, we may integrate them in adaptive control – reducing complexity and excitation requirements.

- Choice of Reference Model and Reference Plant

  The reference model must be chosen to reflect a desirable response of the closed-loop plant – *should have a bandwidth no greater than that of the identifier, and should not have large gains in those frequency regions* (reducing effects from unmodeled dynamics).

  Dual control: reference $r(t)$ affects both the control target and excitation conditions.

Intuitive Ideas (cont'd)

- Time Variation of the Parameters

  Plant parameters may slowly vary over time – estimator needs to discount old input-output data.

- Robust Identification Schemes

  Parameter convergence is not guaranteed in general, but is unnecessary to output convergence. The identifier robustness is fundamental to the adaptive system robustness.

  Careful selection of plant order: Large number leads to numerical issue for identification, but should be sufficient to model the plant dynamics.

  Further filter the regression vectors: to reduce the effect of noise

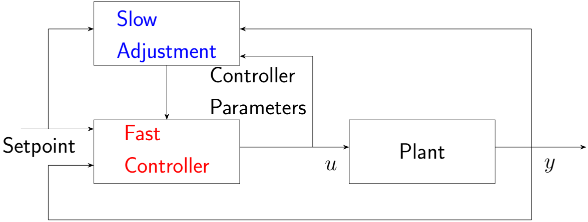
  Monitor the excitation in the identification loop.

Robustness: Slow Adaptation

Rabbit-vs-Tortoise in adaptive systems:

- Slow adaptation
- Fast control loop

to generate two-times scale separation for stability analysis.



Large adaptation gain γ may freeze estimates (Ortega, ACC 2013)

$$\lim_{\gamma\to\infty}\lim_{t\to\infty}[|\tilde{\theta}(0)| - |\tilde{\theta}(t)|] = 0.$$

Robustness: Deadzone

Stop updating the parameters when excitation is insufficient to distinguish between regressor signals and noise – turn off setting a threshold.

For example, we replace the gradient estimator by

$$\dot{\hat{\theta}} = \begin{cases} -\gamma \dfrac{e\phi}{1+\gamma\phi|\phi|^2} & \text{if } \ |e| > \rho \\ 0 & \text{if } \ |e| \leq \rho \end{cases}$$

This can be combined with different types of estimators, e.g. gradient, RLS, normalized, with projection ...

Difficulty in selecting $\rho > 0$

Robustness: Leakage Term (σ-Modification)

Replace the parameter identifier by

$$\dot{\hat{\theta}} = -\gamma\phi e - \sigma\hat{\theta}$$

with a small $\sigma > 0$

- This is stable estimator to keep $\hat{\theta}$ from growing unbounded. However, with PE it cannot guarantee $\hat{\theta} \to \theta$ as $t \to \infty$.
- With a prior estimate of θ as $\hat{\theta}_0$, it can be modified as

$$\dot{\hat{\theta}} = -\gamma\phi e - \sigma(\hat{\theta} - \hat{\theta}_0).$$

  Try to bias the direction of the drift towards $\hat{\theta}_0$ rather than 0.
- Another interesting modification is

$$\dot{\hat{\theta}} = -\gamma\phi e - \sigma|e|\hat{\theta}.$$

  Retain the feature without leakage.

Robustness: Dynamic Regression Extension

Introducing dynamic extension, it may

- act as low-pass filtering of process input/output, thus removing the effects of noise and high-frequency unmodeled dynamics (Witenmark & Astrom, 1984)
- Improve transient performance of parameter estimation – fast convergence rate (Kreisselmeier, TAC 1977) [1]
- Recently new method: Dynamic Regressor Extension and Mixing (DREM) (Ortega et al., Ann. Rev. Control 2020) [2]

[1] G. Kreisselmeier, Adaptive observers with exponential rate of convergence, *IEEE Trans Autom. Control*, vol. 22, pp. 2–8, 1977.

[2] Ortega, Nikiforov & Gerasimov, On modified parameter estimators for identification and adaptive control: A unified framework and some new schemes, *Ann. Rev. Control*, vol. 50, pp. 278–293, 2020.

Beyond

- Multi-input-multi-output systems

$$G(s) \in \mathbb{C}^{n\times m}$$

- Nonlinear adaptive control

$$\dot{x} = f(x) + \phi(x,t)^\top \theta + g(x)u.$$

- Self-tuning adaptive control
- Transient performance and robustness under weak excitation
- Machine learning and adaptation

What have we learned?

- Two methods to deal with uncertainty: Adaptive & (Non-adaptive) Robust
- The Rohrs examples: adaptive systems may not be robust to different types of uncertainties
- Heuristic analysis to the Rohrs examples
- Challenge "well-established" theory for young researchers
- Exponential stability implies robustness
- Some methods to improve robustness of adaptive systems
  - Slow adaptation
  - Dynamic regressor extension
  - Leakage term (σ-modification)
  - Deadzone

Nomenclature for Lecture 3

Bowen.Yi@polymtl.ca

Nomenclature

Monic	A polynomial in s with the coefficient of the highest power is 1, e.g. s^2+2s+5
Hurwitz	If its roots lie in $\mathbb{C}_{<0}$
Minimum phase	A transfer function has *numerator* polynomial Hurwitz, e.g. $\frac{(s+1)(s+2)}{s^2+5s+2}$
Relative degree	Difference between the degrees of the denominator and numerator, e.g. Relative deg$\frac{s+2}{s^2+2}=1$
Proper	relative degree ≥ 0, e.g. $\frac{s+1}{s+2}$
Strictly proper	relative degree >0, e.g. $\frac{1}{s+2}$
$\hat{P}(s)=k_p\frac{\hat{n}_p(s)}{\hat{d}_p(s)}$	Plant transfer function (to estimate or control)
α_i,β_j	Coefficient parameters in the transfer function $\hat{P}(s)=\frac{\alpha_n s^{n-1}+\ldots+\alpha_1}{s^n+\beta^n s^{n-1}+\ldots+\beta_1}$
$y_p,\ r$	Plant output and input
$\hat{y}_p(s),\hat{r}(s)$	Laplace transform of the plant's input and output, i.e. i.e. $\hat{P}(s)=\frac{\hat{y}_p(s)}{\hat{r}(s)}$
$\hat{\lambda}(s)$	$\frac{1}{\hat{\lambda}(s)}$ is the introduced stable filter to deal with not proper terms $\hat{\lambda}(s)=s^n+\lambda_n s^{n-1}+\cdots+\lambda_1$
$\hat{a}^*(s)$	$\hat{a}^*(s)=\alpha_n s^{n-1}+\cdots+\alpha_1$
$\hat{b}^*(s)$	$\hat{b}^*(s)=(\lambda_n-\beta_n)s^{n-1}+\cdots+(\lambda_1-\beta_1)$
$\theta_a,\ \theta_b$	Parameter vectors $\theta_a:=\begin{bmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{bmatrix},\quad \theta_b:=\begin{bmatrix}\lambda_1-\beta_1\\ \vdots\\ \lambda_n-\beta_n\end{bmatrix}$
$\Lambda,\ b_\lambda$	Matrices in state-space realization $\Lambda=\begin{bmatrix}0&1&0&\cdots&0\\0&0&1&\cdots&0\\ \vdots&\vdots&\vdots&\ddots&\vdots\\0&0&0&\cdots&1\\-\lambda_1&-\lambda_2&\cdots&\cdots&-\lambda_n\end{bmatrix}\quad b_\lambda=\begin{bmatrix}0\\ \vdots\\0\\1\end{bmatrix}$
$w_p^{(1)},w_p^{(2)}$	State space variable for the filtered signals
$w^{(1)},w^{(2)}$	States in the designed filter to estimate $w_p^{(1)},w_p^{(2)}$
e	Output error

$\hat{M}(s)=k_m\frac{\hat{n}_m(s)}{\hat{d}_m(s)}$	Reference model (transfer function)
$\hat{\lambda}(s),\ \hat{\lambda}_0(s)$	$\frac{1}{\hat{\lambda}(s)}$ is the introduced filter to deal with not proper terms, but deg $\hat{\lambda}(s)=n-1$ $\hat{\lambda}(s)=\hat{n}_m(s)\lambda_0(s)$
$\hat{a}^*(s),\ \hat{b}^*(s)$	Polynomials to match the reference model to the given plant (to estimate) $\frac{\hat{a}^*(s)}{\hat{\lambda}(s)}=a_0^*+a^{*\top}\frac{1}{\hat{\lambda}(s)}\begin{bmatrix}1&s&s^{n-2}\end{bmatrix}$ $\frac{\hat{b}^*(s)}{\hat{\lambda}(s)}=b_0^*+b^{*\top}\frac{1}{\hat{\lambda}(s)}\begin{bmatrix}1&s&s^{n-2}\end{bmatrix}$
ϕ	Regression $\phi(t)^\top:=\begin{bmatrix}r(t)&w^{(1)\top}(t)&y_p(t)&w^{(2)^\top}(t)\end{bmatrix}\in\mathbb{R}^{2n}$
θ	Unknown parameters $\theta^{*\top}:=\begin{bmatrix}a_0^*&a^{*\top}&b_0^*&b^{*\top}\end{bmatrix}\in\mathbb{R}^{2n}$
e_m	States for the realization of $\hat{M}(s)$

Assumptions on the Plant

A1 Plant Assumption: SISO LTI system, whose transfer function $\hat{P}(s) = \frac{\hat{y}_p(s)}{\hat{r}(s)}$ is

$$\hat{P}(s) = k_p \frac{\hat{n}_p(s)}{\hat{d}_p(s)}$$

- $\hat{r}(s), \hat{y}_p(s)$ - Laplace transforms of input/output
- $\hat{n}_p(s), \hat{d}_p(s)$ - monic, coprime polynomials of degrees m and n
- n is known, but m is unknown
- Plant is strictly proper $m \leq n - 1$

A2 Reference Input Assumption: Input $r(\cdot)$ is piecewise continuous and bounded on $\mathbb{R}_+$.

A3 Output Boundedness Assumption: The plant is located in a control loop such that $r, y_p \in L_\infty$.

Assumptions on the Reference Model

A4 The reference model is an SISO LTI system (selected by us)

$$\hat{M}(s) = k_m \frac{\hat{n}_m(s)}{\hat{d}_m(s)}$$

- $\hat{n}_m(s), \hat{d}_m(s)$ are monic, coprime polynomials of degrees $l, k \leq n$.
- $M(s)$ is strictly proper
- Its relative degree is no greater than the relative degree of the plant $\hat{P}(s)$, i.e. $1 \leq k - l \leq n - m$
- $\hat{d}_m(s)$ is Hurwitz

A5 Positive Real Model: $\hat{M}(s)$ is strictly positive real