ECE 2695: Adaptive Control (3 Credits, Fall 2008)

## Lecture 4: System Identification (II)

September 22, 2008

Instructor: Zhi-Hong Mao
Assistant Professor of ECE and Bioengineering
University of Pittsburgh, Pittsburgh, PA

1

## Outline

- Homework problems
- Review of last lecture: more about Lyapunov stability theory
- Basic identification methods
- General identification problem

2

## Homework problems

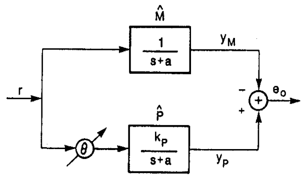
- Homework 3
  - Problems 0.3 and 2.3(a)
  - Due on 9/29 (Monday)
- About last homework

3

## More about Lyapunov stability theory

- An example

MRAC using the MIT rule and the Lyapunov redesign




4

## More about Lyapunov stability theory

- An example
- Control Lyapunov function
  - One way to design a nonlinear controller is to begin with a candidate Lyapunov function $v(x)$ and a control system $dx/dt = f(x, u)$. We say that $v(x)$ is a control Lyapunov function if for every $x$ there exists a $u$ such that $dv(x)/dt < 0$. In this case, it may be possible to find a function $g(x)$ such that $u = g(x)$ stabilizes the system

5

## More about Lyapunov stability theory

- An example
- Control Lyapunov function

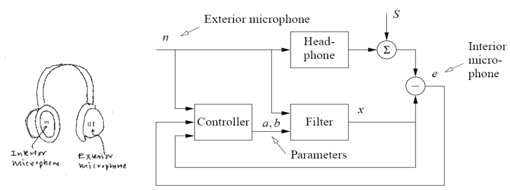
Adaptive noise cancellation


6

## More about Lyapunov stability theory

- An example
- Control Lyapunov function

Adaptive noise cancellation



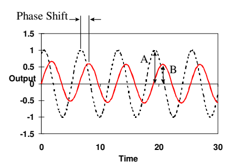

7

## Basic identification methods

- Frequency domain approach
  - Frequency response: steady-state response of systems to sinusoidal inputs

The figure compares the output response of a system (red solid line) with a sinusoidal input (black dashed line)

Both the magnitude and the phase shift of a system will change with the frequency of the input into the system




Amplitude Ratio = B / A

8

## Basic identification methods

- Frequency domain approach
  - Frequency response: steady-state response of systems to sinusoidal inputs
  - Frequency response function
    - Given a system with transfer function $\hat{P}(s)$, its frequency response function is $\hat{P}(j\omega)$
    - The steady-state gain of a system for a sinusoidal input $\sin(\omega_0 t)$ is the magnitude of the transfer function evaluation at $s = j\omega_0$, and the phase shift of the output sinusoid relative to the input sinusoid is the angle of $\hat{P}(j\omega_0)$

9

## Basic identification methods

- Frequency domain approach
  - Frequency response: steady-state response of systems to sinusoidal inputs
  - Frequency response function
  - Example

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

10

## Basic identification methods

- Frequency domain approach
  - Frequency response: steady-state response of systems to sinusoidal inputs
  - Frequency response function
  - Example

Question: How about identification of a system of higher order?

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

11

## Basic identification methods

- Frequency domain approach
- Time domain approach
  - Example

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

Question: What is the differential equation for the above frequency-domain description?

12

## Basic identification methods

- Frequency domain approach
- Time domain approach
  - Example

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

$$\dot{y}_p(t) = -a_p y_p(t) + k_p r(t)$$

Question: If we have measurement of $r(t)$ and $dy_p(t)/dt$ at $t_1$ and $t_2$, how can we estimate $a_p$ and $k_p$?

13

## Basic identification methods

- Frequency domain approach
- Time domain approach
  - Example

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

$$\dot{y}_p(t) = -a_p y_p(t) + k_p r(t)$$

$$\begin{bmatrix} -a_p \\ k_p \end{bmatrix} = \begin{bmatrix} y_p(t_1) & r(t_1) \\ y_p(t_2) & r(t_2) \end{bmatrix}^{-1} \begin{bmatrix} \dot{y}_p(t_1) \\ \dot{y}_p(t_2) \end{bmatrix}$$

14

## Basic identification methods

- Frequency domain approach
- Time domain approach
  - Example

$$\frac{\hat{y}_p(s)}{\hat{r}(s)} = \hat{P}(s) = \frac{k_p}{s+a_p} \rightarrow \dot{y}_p(t) = -a_p y_p(t) + k_p r(t)$$

$$\begin{bmatrix} -a_p \\ k_p \end{bmatrix} = \begin{bmatrix} y_p(t_1) & r(t_1) \\ y_p(t_2) & r(t_2) \end{bmatrix}^{-1} \begin{bmatrix} \dot{y}_p(t_1) \\ \dot{y}_p(t_2) \end{bmatrix}$$

Question: Why we avoid the measurement of $dy_p(t)/dt$ and how to do that?

15

## Basic identification methods

- Frequency domain approach
- Time domain approach
  - Example

$$\frac{\hat{y}_p(s)}{\hat{r}(s)} = \hat{P}(s) = \frac{k_p}{s+a_p} \rightarrow \frac{s+a_p}{s+\lambda}\hat{y}_p = \frac{k_p}{s+\lambda}\hat{r}$$

Define $\hat{w}^{(1)} = \frac{1}{s+\lambda}\hat{r}$ $\quad \hat{w}^{(2)} = \frac{1}{s+\lambda}\hat{y}_p$

$$\begin{aligned} y_p(t) &= k_p w^{(1)}(t) + (\lambda - a_p)w^{(2)}(t) \\ \dot{w}^{(1)} &= -\lambda w^{(1)} + r \\ \dot{w}^{(2)} &= -\lambda w^{(2)} + y_p \end{aligned}$$

16

## Basic identification methods

- Frequency domain approach
- Time domain approach
  - Example

$$\frac{\hat{y}_p(s)}{\hat{r}(s)} = \hat{P}(s) = \frac{k_p}{s+a_p} \rightarrow \begin{aligned} y_p(t) &= k_p w^{(1)}(t) + (\lambda - a_p)w^{(2)}(t) \\ \dot{w}^{(1)} &= -\lambda w^{(1)} + r \\ \dot{w}^{(2)} &= -\lambda w^{(2)} + y_p \end{aligned}$$

Define $\theta^* := \begin{bmatrix} k_p \\ \lambda - a_p \end{bmatrix}$

$$w(t) := \begin{bmatrix} w^{(1)}(t) \\ w^{(2)}(t) \end{bmatrix}$$

$$y_p(t) = \theta^{*T} w(t) = w^T(t)\theta^*$$

17

## Basic identification methods

- Frequency domain approach
- Time domain approach
  - Example

$$y_p(t) = \theta^{*T} w(t) = w^T(t)\theta^*$$

Nominal identifier parameter

$$e_1(t) = \theta^T w(t) - y_p(t) = \left(\theta^T - \theta^{*T}\right) w(t)$$

Identification error

Adaptive identifier parameter

Linear error equation

18

## Basic identification methods

- Frequency domain approach
- Time domain approach
  - Example
  - Gradient algorithm

$$e_1(t) = \theta^T w(t) - y_p(t) = \left(\theta^T - \theta^{*T}\right) w(t)$$

Objective: minimize $e_1^2(t)$

Gradient: $\frac{\partial}{\partial \theta}\left(e_1^2\right) = 2e_1 \frac{\partial}{\partial \theta}\left(e_1\right) = 2e_1 w$

It is a vector!

Parameter update law: $\frac{d\theta}{dt} = -g e_1 w$

19

## Basic identification methods

- Frequency domain approach
- Time domain approach
  - Example
  - Gradient algorithm
    - Advantages of using the gradient algorithm over using the following calculation

$$\begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} w_1(t_1) & w_2(t_1) \\ w_1(t_2) & w_2(t_2) \end{bmatrix}^{-1} \begin{bmatrix} y_p(t_1) \\ y_p(t_2) \end{bmatrix}$$

20

## Basic identification methods

- Frequency domain approach
- Time domain approach
  - Example
  - Gradient algorithm
  - Least-squares algorithm

Objective: minimize the integral-squared-error (ISE)

$$\text{ISE} = \int_0^t \left[\theta^T(\tau) w(\tau) - y_p(\tau)\right]^2 d\tau$$

$$\theta_{LS}(t) = \left[\int_0^t w(\tau)\, w^T(\tau)\, d\tau\right]^{-1} \left[\int_0^t w(\tau)\, y_p(\tau)\, d\tau\right]$$

Least-squares estimate

21

## Basic identification methods

- Frequency domain approach
- Time domain approach
  - Example
  - Gradient algorithm
  - Least-squares algorithm

Recursive formulation

$$\dot{\theta}(t) = -P(t)w(t)\left[\theta^T(t)w(t) - y_p(t)\right] \quad \theta(0) = \theta_0$$
$$\dot{P}(t) = -P(t)w(t)w^T(t)P(t) \quad P(0) = P^T(0) = P_0$$

Remark: It can be shown that $\theta(t)$ converges asymptotically to $\theta^*$ if $\int_0^t w(\tau)w^T(\tau)d\tau$ is unbounded as $t \to \infty$.

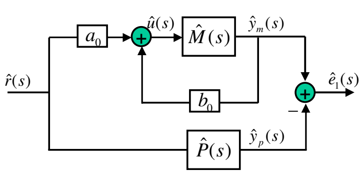
22

## Basic identification methods

- Frequency domain approach
- Time domain approach
  - Example
  - Gradient algorithm
  - Least-squares algorithm
  - Model reference identification

Reference model

$$\frac{\hat{y}_p(s)}{\hat{r}(s)} = \hat{P}(s) = \frac{k_p}{s + a_p} \qquad \frac{\hat{y}_m(s)}{\hat{u}(s)} = \hat{M}(s) = \frac{k_m}{s + a_m}$$




Question: What is the difference between this architecture and the one for MRAC?

23

## General identification problem

- A few concepts
  - Monic
    - A polynomial in $s$ is called monic if the coefficient of the highest power of $s$ is 1
  - Hurwitz
    - A polynomial in $s$ is called Hurwitz if its roots lie in the open left-half plane
  - Minimum phase
    - Rational transfer functions are called stable if their denominator polynomial is Hurwitz and minimum phase if their numerator polynomial is also Hurwitz

24

## General identification problem

- A few concepts
  - Monic
  - Hurwitz
  - Minimum phase
  - Relative degree
    - The relative degree of a transfer function is the difference between the degrees of the denominator and numerator polynomials
  - Proper
    - A rational transfer function is called proper if its relative degree is at least $0$ and strictly proper if its relative degree is at least $1$

25

## General identification problem

- A few concepts
- A few assumptions
  - Plant assumptions
    - The plant is SISO LTI system, described by a transfer function

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

where $\hat{n}_p(s)$ and $\hat{d}_p(s)$ are monic, coprime polynomials of degrees $m$ and $n$ respectively. The degree $m$ is unknown, but the plant is strictly proper $(m < n)$

26

## General identification problem

- A few concepts
- A few assumptions
  - Plant assumptions $\boxed{\frac{\hat{y}_p(s)}{\hat{r}(s)} = \hat{P}(s) = k_p \frac{\hat{n}_p(s)}{\hat{d}_p(s)}}$
  - Reference input assumptions
    - The input $r(t)$ is piecewise continuous and bounded on $\mathbf{R}_+$

27

## General identification problem

- A few concepts
- A few assumptions
- Identifier structure

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

  - Objective
    - Find an expression which depends linearly on the unknown parameters (why?)



## General identification problem

- A few concepts
- A few assumptions
- Identifier structure

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

  - Objective
    - Find an expression which depends linearly on the unknown parameters

One candidate expression

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

Question: What is the advantage or disadvantage of using the above expression?



## General identification problem

- A few concepts
- A few assumptions
- Identifier structure

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

  - Objective
    - Find an expression which depends linearly on the unknown parameters

Another expression

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

Question: What is the requirement on $\hat{\lambda}(s)$?

## General identification problem

- A few concepts
- A few assumptions
- Identifier structure
  - Objective
  - Linear error equation

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

$$\boxed{\begin{aligned} y_i(t) &= \theta^T(t)\, w(t) \\ y_p(t) &= \theta^{*'} w(t) \end{aligned}}$$

$$\phi(t) := \theta(t) - \theta^*$$

$$\boxed{e_1(t) := y_i(t) - y_p(t) = \phi^T(t)\, w(t)}$$

31

## General identification problem

- A few concepts
- A few assumptions
- Identifier structure
- Identification algorithms
  - Gradient algorithms

Standard gradient algorithm

$$\dot{\theta} = -g\, e_1\, w \qquad g > 0$$

Normalized gradient algorithm

$$\dot{\theta} = -g \frac{e_1 w}{1 + \gamma w^T w} \qquad g, \gamma > 0$$

32

## General identification problem

- A few concepts
- A few assumptions
- Identifier structure
- Identification algorithms
  - Gradient algorithms
  - Least-squares algorithms

Iterative least-squares algorithm

$$\dot{\theta} = -g\, P\, w\, e_1$$

$$\frac{dP}{dt} = -gPww^T P$$

$$\text{or} \quad \frac{d(P^{-1})}{dt} = gww^T \qquad g > 0$$

33

## General identification problem

- A few concepts
- A few assumptions
- Identifier structure
- Identification algorithms
  - Gradient algorithms
  - Least-squares algorithms

Normalized least-squares algorithm

$$\dot{\theta} = -g \frac{P\, w\, e_1}{1 + \gamma\, w^T P\, w} \qquad g, \gamma > 0$$

$$\frac{dP}{dt} = -g \frac{P\, w\, w^T P}{1 + \gamma\, w^T P\, w}$$

$$\text{or} \quad \frac{d(P^{-1})}{dt} = g \frac{w\, w^T}{1 + \gamma\, w^T (P^{-1})^{-1} w}$$

34

## General identification problem

- A few concepts
- A few assumptions
- Identifier structure
- Identification algorithms
- Persistency of excitation (PE)
  - Definition
    - A vector $w$: $\mathbf{R}_+ \to \mathbf{R}^{2n}$ is persistently exciting (PE) if there exists $\alpha_1$, $\alpha_2$, $\delta > 0$ such that

$$\alpha_2 I \ \geq \int_{t_0}^{t_0+\delta} w(\tau)\, w^T(\tau)\, d\tau \ \geq \ \alpha_1 I \qquad \text{for all } t_0 \geq 0$$

35

## General identification problem

- A few concepts
- A few assumptions
- Identifier structure
- Identification algorithms
- Persistency of excitation (PE)
  - Definition
    - A vector $w$: $\mathbf{R}_+ \to \mathbf{R}^{2n}$ is persistently exciting (PE) if there exists $\alpha_1$, $\alpha_2$, $\delta > 0$ such that

$$\alpha_2 \ \geq \int_{t_0}^{t_0+\delta} (w^T(\tau)\, x)^2\, d\tau \ \geq \ \alpha_1 \qquad \text{for all } t_0 \geq 0,\ |x| = 1$$

36

## General identification problem

- A few concepts
- A few assumptions
- Identifier structure
- Identification algorithms
- Persistency of excitation (PE)
  - Definition
  - PE and exponential stability
    - Let $w$: $\mathbf{R}_+ \to \mathbf{R}^{2n}$ be piecewise continuous and PE, then the differential equation

$$\dot{\phi}(t) = -g\, w(t)\, w^T(t)\, \phi(t) \qquad g > 0$$

is globally exponential stable

37

## References


- K. J. Astrom and R. M. Murray, Feedback Systems: An Introduction for Scientists and Engineers, Manuscript, 2007.
- S. Sastry and M. Bodson, Adaptive Control: Stability, Convergence, and Robustness, Prentice-Hall, 1989.
- G. Strang, Introduction to Applied Mathematics, Wellesley-Cambridge Press, 1986.
- http://static.howstuffworks.com/gif/noise-canceling-headphone-6.jpg


38