System Identification, Lecture 10

Kristiaan Pelckmans (IT/UU, 2338)

Course code: 1RT880, Report code: 61800 - Spring 2011
F, FRI Uppsala University, Information Technology

2 Mai 2011

Overview Part II

1. State Space Systems.
2. Subspace Identification.
3. Further Topics.
4. Identification of Nonlinear Models.
5. Wider View.

Overview Further Topics

1. Design of Experiments.
2. Closed Loop Identification.
3. Preprocessing.
4. User Choices.

1. Design of Experiments

1. General.

2. Informative Experiment.

3. Optimal Experiments.

4. Sampling.

General Considerations

- Purpose and Norms.
- Physical versus Black-box.
- Placement of sensors.
- Manipulate Signals?
- Which signals are to be considered in/out?
- Sampling period?
- Operation Point?
- Length n?

Problem: for given $\mathcal{S} \in \mathcal{M}$:

$$\max_{M \in \{\mathcal{M}\}, \mathbf{u} \in \sqcap} \mathcal{I}(M) \text{ s.t. } M(\mathbf{u}) = \mathcal{S}(\mathbf{u})$$

MINIMAX:

$$\max_{M \in \{\mathcal{M}\}, \mathbf{u} \in \sqcap} \min_{\mathcal{S} \in \mathcal{M}} \mathcal{I}(M) \text{ s.t. } M(\mathbf{u}) = \mathcal{S}(\mathbf{u})$$

where

- The system $\mathcal{S}$ to be identified.
- $\mathbf{u}$ represents the input signal to be injected into $\mathcal{S}$.
- A class of allowed input signals $\sqcap$.
- The model class $\mathcal{M}$ of candidate models.
- The model $M \in \mathcal{M}$ identified ('$=$') based on $\mathbf{u}$ and $\mathcal{S}(\mathbf{u})$
- The use (information content) of a model M is $\mathcal{I}(M)$.

A common choice of $\mathcal{I}(M)$ is based on the covariance of the parameters of M.

- $\hat{\theta} \rightarrow M$.
- $\theta_0 \rightarrow \mathcal{S}$.
- $\mathbf{P}_{\theta_0}(u) \propto \left[\mathbb{E}\left(\frac{d\hat{y}_t(\theta_0)}{d\theta_0}\right)\left(\frac{d\hat{y}_t(\theta_0)}{d\theta_0}\right)^T\right]^{-1}$.
- $\alpha : \mathbb{R}^{d\times d} \rightarrow \mathbb{R}$ measures 'size' of $\mathbf{P}_{\theta_0}$.

Then

$$\min_{\mathbf{u}\in\sqcap} \max_{\theta_0\in\Theta} \alpha(\mathbf{P}_{\theta_0}(\mathbf{u}))$$

For FIR systems of order d, covariance of $\hat{\theta}$ given as

$$\mathbf{P}_{\theta_0}(\mathbf{u}) = \frac{1}{n} \begin{bmatrix} r_0(\mathbf{u}) & r_1(\mathbf{u}) & \dots & r_{d-1}(\mathbf{u}) \\ r_1(\mathbf{u}) & r_0(\mathbf{u}) & & \\ \vdots & & \ddots & \\ r_{d-1}(\mathbf{u}) & r_{d-2}(\mathbf{u}) & & r_0(\mathbf{u}) \end{bmatrix}^{-1}$$

- Invertible $\Leftrightarrow$ Informative.
- Independent of $\theta_0, \hat{\theta}$!
- Stochastic Issues.

Crest factor

$$C_r^1(\mathbf{u}) = \frac{\max_t \mathbf{u}_t^2}{\frac{1}{n}\sum_t \mathbf{u}_t^2}$$

minimum 1.

Relation to frequency content.

$$\mathbf{P}_\theta^{-1}(\mathbf{u}) \propto \int_{\pi}^{\pi} \mathbf{H}(\omega)\phi_u(\omega)d\omega + \mathbf{H}_e$$

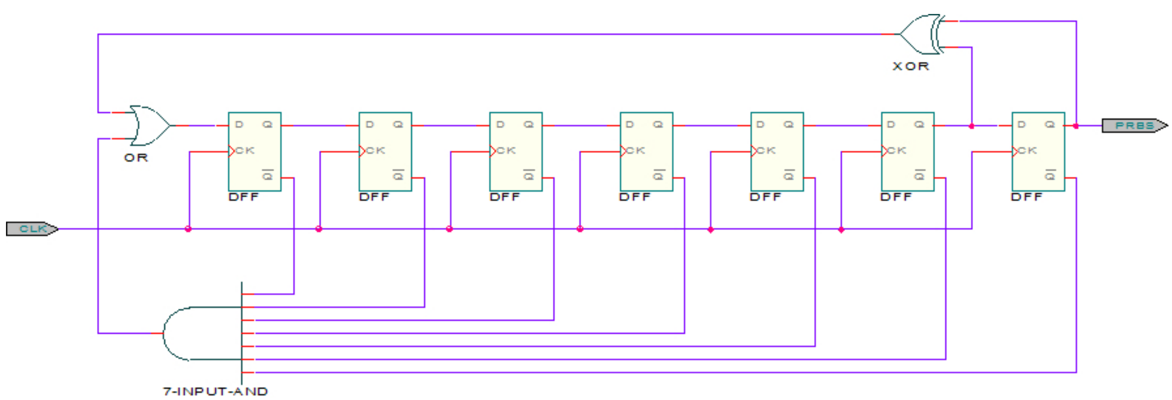
where

- $\mathbf{H}(\omega)$ denotes how sensitive the system $\mathcal{S}$ is to frequency ω
- $\mathbf{H}_e \propto \int_{-\pi}^{\pi} \frac{1}{\phi_e(\omega)} H'(e^{i\omega})H'(e^{i\omega})^T d\omega$
- So put frequencies of ϕ_u on places where $\mathbf{H}(\omega)$ is large.
- If a parameter is of interest, vary it, look at where changes bode plot, and put input power there.

- Choice of function α
  - $\alpha(\mathbf{P}) = \mathrm{tr}(\mathbf{PW})$ (A-Optimal Design)
  - $\alpha(\mathbf{P}) = \det(\mathbf{P})$ (D-Optimal Design)
  - $\alpha(\mathbf{P}) = -\underline{\lambda}(\mathbf{P})$ (E-Optimal Design)
- Optimize over $\sqcap$.

Common choices:

1. White Noise.
2. Filtered White Noise.
3. PRBS.
4. Swept Sinusoids.
5. Periodic vs. aperiodic (PE, Averaging, Transient).



Intersample D/A signal u_t from $\mathbf{u}$.

- ZOH or FOH.
- Trigonometric interpolation (band-limited).

Sampling Period:

- Information content.
- Nyquist.
- Computational.
- Higher model orders and delays.
- $10T$

Length of Experiment:

- Critical mass.
- Averaging.
- Computational.

Prefiltering:

- Low-pass and differencing.
- Antialiasing.

Outliers and Missing variables:

- Unknown 'parameters'
- Averages.
- Norms.

2. Closed Loop Identification

Consider the system:

$$\begin{cases} y_t = G(q^{-1})u_t + H(q^{-1})e_t \\ u_t = -F(q^{-1})y_t + L(q^{-1})v_t \end{cases}$$

where

- The input u_t is determined through feedback.
- F and L are called regulators.
- The signal v_t can be the reference signal or noise entering the regulator.

Why?

- Many realworld systems have feedback.
- The open-loop system is unstable.
- Feedback is required due to safety reasons.

What

- The input u_t depends on past y_t (and hence on past e_t).
- The aim of control is to apply a u_t which minimizes the deviation between y_t and a reference signal v_t. Good control often requires a u_t of bounded energy.
- SI requires PE, hence substantial energy of u_t.
- The frequency content of u_t is limited by the true system.

An example

System:

$$\begin{cases} y_t + a y_{t-1} = b u_{t-1} + e_t, \; E[e_t^2] = \lambda^2 \\ u_t = -f y_t \end{cases}$$

Model structure:

$$y_t + \hat{a} y_{t-1} = \hat{b} u_{t-1} + \epsilon(t)$$

Estimate by PEM

$$\begin{cases} \hat{a} = a + f\gamma \\ \hat{b} = b - \gamma \end{cases}$$

where γ is any scalar. There is no unique solution, hence the parameters are not estimated consistently.

Closed-loop behavior

Open-loop system:

$$\begin{cases} y_t = G(q^{-1})u_t + H(q^{-1})e_t \\ u_t = -F(q^{-1})y_t + L(q^{-1})v_t. \end{cases}$$

Closed loop system:

$$\begin{cases} y_t = (I + GF)^{-1}GLv_t + (I + GF)^{-1}He_t \\ u_t = \left(L - F(I + GF)^{-1}GL\right) v_t - F(I + GF)^{-1}He_t. \end{cases}$$

Some assumptions

- The open loop system is strictly proper: y_t depends only on past values of the input u_s or $s < t$.
- The closed loop system is asymptotically stable.
- The external signal v_t is stationary and PE of sufficiently high order.
- The external signal v_t and the disturbance e_s are independent $\forall s, t$.

Prediction Error Methods

- In most cases it is not necessary to assume that the external signal v_t is measurable.
- Gives statistically efficient estimates under mild conditions.
- Computationally demanding.

Notation $\hat{G}$ denotes $G(q^{-1}, \hat{\theta})$.

Different Approaches

- *Direct Identification*. Feedback is neglected during identification - the system is treated as an open loop system.

- *Indirect Identification*. It is assumed that v_t is measured and the feedback law is known. First the closed loop behavior is modeled, then the open-loop system is identified by 'subtracting' the effect of the regulators from this model.

- *Joint Identification*. The signals u_t and y_t are both considered as the outputs of a multivariate system driven by white noise.

Direct Identification

Model structure:

$$\begin{cases} y_t = Gu_t + He_t \\ E[e^2(t)] = \lambda^2 \end{cases}$$

Use the signals $(u_t)_t$ and $(y_t)_t$

Goal: estimate (SISO)

$$\begin{cases} \hat{\theta} = \operatorname{argmin}_\theta V_N(\theta) \\ V_N(\theta) = \frac{1}{N}\sum_{t=1}^N \epsilon_t^2(\theta) \\ \epsilon_t(\hat{\theta}) = \hat{H}^{-1}\Big(y_t - \hat{G}u_t\Big) \end{cases}$$

Question: Identifiability? Desired solution $\hat{G} = G$ and $\hat{H} = H$.

Consistency: Analyze the asymptotic cost function:

$$V(\theta) = \lim_{N\to\infty} V_N(\theta) = E[\epsilon(t,\theta)]$$

- Will $\hat{G} = G$ and $\hat{H} = H$ be a global minimum to $V(\theta)$ (system identifiability)?
- Is the solution $\hat{G} = G$ and $\hat{H} = H$ unique (parameter identifiability)?

An Example

System:

$$y_t + a y_{t-1} = b u_{t-1} + e_t,\ E[e^2(t)] = \lambda^2$$

Model structure:

$$y_t + \hat{a} y_{t-1} = \hat{b} u_{t-1} + \epsilon(t)$$

Input

$$u_t = \begin{cases} -f_1 y_t & \text{for a fraction } \gamma_1 \text{ of the total time.} \\ -f_2 y_t & \text{for a fraction } \gamma_2 \text{ of the total time.} \end{cases}$$

Then (for $i = 1, 2$) we get

$$\begin{cases} y_t^i + (a + b f_i) y_{t-1}^i = e_t \\ y_t^i + (\hat{a} + \hat{b} f_i) y_{t-1}^i = \epsilon_t^i \end{cases}$$

which gives

$$\begin{aligned}V(\hat{a},\hat{b}) &= \gamma_1 E[\epsilon_1^2(t)] + \gamma_2 E[\epsilon_2^2(t)] \\ &= \lambda^2 + \gamma_1 \lambda^2 \frac{(\hat{a}+\hat{b}f_1 - a - bf_1)^2}{1-(a+bf_1)^2} \\ &\quad + \gamma_2 \lambda^2 \frac{(\hat{a}+\hat{b}f_2 - a - bf_2)^2}{1-(a+bf_2)^2}\end{aligned}$$

Consequently

$$V(\hat{a},\hat{b}) \geq \lambda^2 = V(a,b)$$

we get

- A global minimum is obtained if $\hat{a} = a$ and $\hat{b} = b$
- Unique minimum?
- Solve $V(\hat{a},\hat{b}) = \lambda^2$

$$\begin{bmatrix} 1 & f_1 \\ 1 & f_2 \end{bmatrix} \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = \begin{bmatrix} a + bf_1 \\ a + bf_2 \end{bmatrix}$$

- Unique solution if and only if $f_1 \neq f_2$ (Compare to our previous example).

The General Case

- The desired solution $\hat{G} = G$ and $\hat{H} = H$ will be a global minimum to $V(\theta)$

- Unique global minimum is necessary for parameter identifiability (consistency). Consistency is assured by
  - Using an external input signal v_t
  - Using a regulator $F(q^{-1})$ that shifts between different settings during the experiment.

Indirect Identification

- Two step approach
  1. **Step 1** Identify the closed loop system using v_t as input and y_t as output.
  2. **Step 2** Determine the open loop system parameters from the closed loop parameters, *using knowledge of the feedback F and L*.

- Closed-loop system:

$$y_t = \bar{G} v_t + \bar{H} e_t$$

where

$$\begin{cases} \bar{G} = (I + GF)^{-1} GL \\ \bar{H} = (I + GF)^{-1} H \end{cases}$$

- Estimate $\bar{G}$ and $\bar{H}$ from v_t and y_t with a PEM.

- From the estimated $\bar{G}$ and $\bar{H}$, form the $\hat{G}$ and $\hat{H}$

- Identifiability conditions are the same as for the direct approach.
- Same identifiability conditions do not imply that both direct as indirect approach give the same result.
- Drawback of indirect approach: one needs to know v_t and the regulators.

Joint input-output identification.

- Regard u_t and y_t as outputs from a multivariable system, driven by white noise and the reference input v_t.

$$\begin{cases} y_t = H_{11}(q^{-1,\theta})e_t + H_{12}(q^{-1,\theta})v_t \\ u_t = H_{21}(q^{-1,\theta})e_t + H_{22}(q^{-1,\theta})v_t \end{cases}$$

- Innovations model: let $z_t = \Big(y_t, u_t\Big)^T$ and $\bar{e}_t = \Big(e_t, v_t\Big)^T$, then

$$z_t = \mathbf{H}(q^{-1}, \theta)\bar{e}_t$$

with $E[\bar{e}_s\bar{e}_t^T] = \Lambda_{\bar{e}}(\theta)\delta_{t,s}$.

- Use PEM to identify θ in $\mathbf{H}$ and $\Lambda_{\bar{e}}$.

Properties

- Same identifiability conditions as for the direct method.
- Both system and the regulator can be identified.
- The spectral characterization of v_t can be identified;
- the drawback is the computational demand.

Conclusions

To remember

- Design of Experiments.
- Closed Loop Identification.
- Preprocessing.