Modeling and Learning for Dynamical Systems

Lecture 10

Martin Enqvist

LINKÖPING UNIVERSITY

Closed-Loop System Identification

Why Closed-Loop Experiments?

- The system is unstable and open-loop experiments are therefore not feasible.
- Data comes from a system in normal closed-loop operation. It is too expensive to perform an open-loop experiment just for the purpose of system identification.
- The feedback is inherent in the system.

Setup

Consider a closed-loop setup where the **true system** can be written

$$y(t) = G_0(q)u(t) + \underbrace{H_0(q)e(t)}_{=v(t)},$$

where $e(t)$ is white noise with variance λ_0. The **controller** can be written

$$u(t) = \tilde{r}(t) - F_y(q)y(t),$$

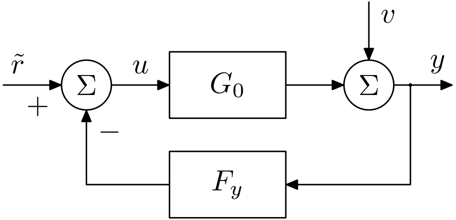
where $\tilde{r}$ is a (filtered) reference signal.

Consider also a **generic model**

$$y(t) = G(q,\theta)u(t) + H(q,\theta)e(t),$$

where θ is a **parameter vector**.

Setup...



Assumptions:

- Either $F_y(q)$ or both $G_0(q)$ and $G(q, \theta)$ contain a delay.
- The closed-loop system is stable.

Setup...

Closed-loop equations:

$$y(t) = S_0(q)G_0(q)\tilde{r}(t) + S_0(q)v(t),$$
$$u(t) = S_0(q)\tilde{r}(t) - S_0(q)F_y(q)v(t),$$

where

$$S_0(q) = \frac{1}{1 + G_0(q)F_y(q)}$$

is the sensitivity function.

Setup. . .

The input consists of two components originating from the reference signal and the noise, respectively:

$$u(t) = u^{\tilde{r}}(t) + u^{v}(t)$$

Spectrum ($\tilde{r}$ and v independent):

$$\Phi_u(\omega) = \underbrace{|S_0(e^{i\omega})|^2 \Phi_{\tilde{r}}(\omega)}_{\Phi_u^{\tilde{r}}(\omega)} + \underbrace{|F_y(e^{i\omega})|^2 |S_0(e^{i\omega})|^2 \Phi_v(\omega)}_{\Phi_u^{e}(\omega)}$$

Closed-loop Challenges

Key question: Why is closed-loop identification more challenging than open-loop identification?

- There will be correlation between the input signal at time t and past values of the noise signal $v(t)$. Because of this, several methods that work well in open loop will give biased estimates in closed loop.
- A closed-loop experiment may contain less information about the system, making it impossible to estimate a model uniquely

Simplified Noise Models

The system dynamics from input to output can be estimated consistently with a simplified noise model in open-loop identification using PEM. For example, an output-error model

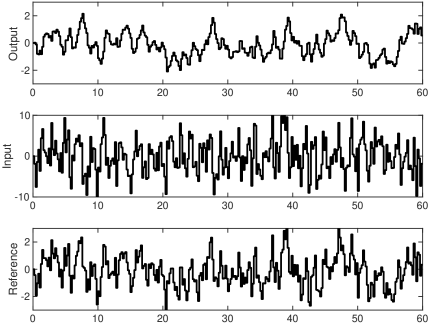
$$y(t) = G(q,\theta)u(t) + e(t) \quad \text{where} \quad G(q,\theta) = \frac{B(q)}{F(q)}$$

can be used to get consistent (estimates that converge to the true values) estimates of $G_0(q)$ even if the additive noise is not white.

However, with data from a closed-loop experiment, the use of a simplified noise model in PEM will usually result in a (asymptotically) biased estimate of $G_0(q)$.

Closed-Loop Example

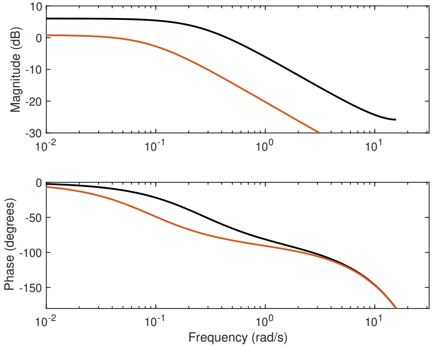
Consider a particular closed-loop system where $G_0(q)$ is a first-order system, $H_0(q)$ is a second-order system and the controller is a PI controller. A dataset with $N = 10000$ samples of r, u and y has been collected.



(part of the dataset)

Closed-Loop Example: OE

An OE model structure with correct orders ($n_b = n_f = n_k = 1$) results in a biased estimate



(estimate in red, true frequency response in black)

Standard Subspace Methods

The standard subspace methods have turned out to give accurate model estimates in many open-loop settings. For example, the subspace approach is particularly appealing for large MIMO systems, and for initialization of PEM algorithms.

> However, with data from a closed-loop experiment, the standard subspace methods will typically *not* result in consistent estimators.

Spectral Analysis

For open-loop identification problems, spectral analysis can be used to to obtain an accurate nonparametric estimate

$$\hat{G}_N(e^{i\omega}) = \frac{\hat{\Phi}_{yu}^N(\omega)}{\hat{\Phi}_u^N(\omega)}$$

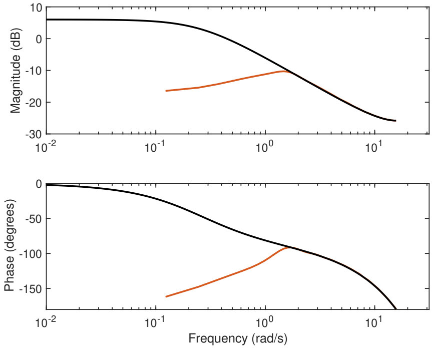
of the frequency response of the system, for example for validation purposes.

However, with data from a closed-loop experiment, the spectral analysis estimator will converge to

$$\frac{G_0(e^{i\omega})\Phi_{\tilde{r}}(\omega) - F_y(e^{-i\omega})\Phi_v(\omega)}{\Phi_{\tilde{r}}(\omega) - |F_y(e^{i\omega})|^2\Phi_v(\omega)}.$$

Closed-Loop Example: SPA

The standard spectral analysis estimate is biased



(estimate in red, true frequency response in black)

Correlation Analysis

For open-loop data, the impulse response of a system can be estimated using correlation analysis. The impulse response estimator is

$$\hat{g}_\tau = \frac{\hat{R}^N_{y_F u_F}(\tau)}{\hat{R}^N_{u_F}(0)}$$

where u_F is the pre-whitened input signal and y_F the corresponding output signal.

However, with data from a closed-loop experiment, the correlation analysis impulse response estimator will be biased.

PEM in Closed Loop

Finally, some good news:

The prediction-error method (PEM) results in a consistent estimator if

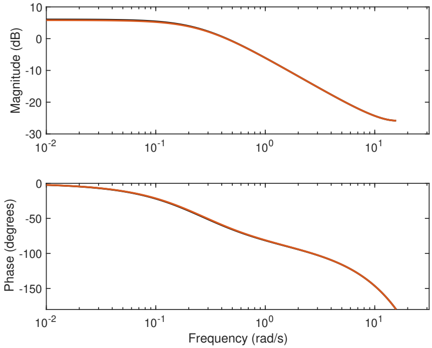
- The experimental data are informative (for example, thanks to a reference signal with enough variations)
- The model structure is flexible enough such that it can be used to describe the true system

regardless if the input-output data have been collected under feedback

N.B. The second assumption means that *both* $G_0(q)$ and $H_0(q)$ must be possible to describe using the chosen model structure.

Closed-Loop Example: BJ

A Box-Jenkins (BJ) model structure with correct model orders ($n_b = n_c = n_f = n_k = 1$, $n_d = 2$) results in an accurate estimate



(estimate in red, true frequency response in black)

Different Approaches to Closed-loop Identification

Three Categories

The available methods for closed-loop identification belong to the following categories:

- The direct approach
- The indirect approach
- The joint input-output approach (including the two-stage method)

The Direct Approach

The **direct PEM approach**:

- Apply the basic prediction-error method using only the input $u(t)$ and the output $y(t)$ in the same way as for an open-loop system.
- Ignore any possible feedback.
- Do not use the reference signal $r(t)$.

Other direct approaches: Some special subspace methods for closed-loop identification can also be viewed as direct approaches.

Direct PEM Approach...

It can be shown that the bias $B(q, \theta)$ of the G model when direct PEM is used is:

$$|B(e^{i\omega}, \theta)|^2 = \frac{\lambda_0}{\Phi_u(e^{i\omega})} \frac{\Phi_u^e(e^{i\omega})}{\Phi_u(e^{i\omega})} |H_0(e^{i\omega}) - H(e^{i\omega}, \theta)|^2$$

This means that $G(e^{i\omega}, \theta)$ will approximate $G_0(e^{i\omega}) + B(e^{i\omega}, \theta)$ instead of $G_0(e^{i\omega})$ (with independently parameterized G and H).

Observations: The bias in the estimate of $G_0(q)$ will be small in frequency regions where either (or all) of the following holds:

- The noise model is accurate
- The feedback contribution (Φ_u^e/Φ_u) is small
- The signal to noise ratio is high (λ_0/Φ_u is small)

In particular, the bias will be small if a flexible enough noise model is used.

Direct PEM Approach...

As mentioned earlier, a simplified noise model cannot be used in the direct PEM approach. Hence, a simplified model cannot be fitted to the true system in a certain frequency region by prefiltering the input-output data as in the open-loop case.

However, this can be handled by first estimating a high-order model and then reducing the model order

High-order ARX Models

High-order ARX models can be used to approximate any linear system arbitrarily well since it can be shown that the least-squares estimates satisfy

$$\frac{\hat{B}_N^M(e^{i\omega})}{\hat{A}_N^M(e^{i\omega})} \to G_0(e^{i\omega})$$
$$\frac{1}{\hat{A}_N^M(e^{i\omega})} \to H_0(e^{i\omega})$$

uniformly in ω as $N >> M \to \infty$. (M is the model order.)

Model Reduction

Model reduction of a high-order model can be done in several ways. For example:

- Using balanced model reduction
- By simulating the high-order model with an input with suitable spectrum and estimating a low-order OE model (unstable models may cause problems).

Of course, the first alternative can be used as a way to initialize the optimization in the second.

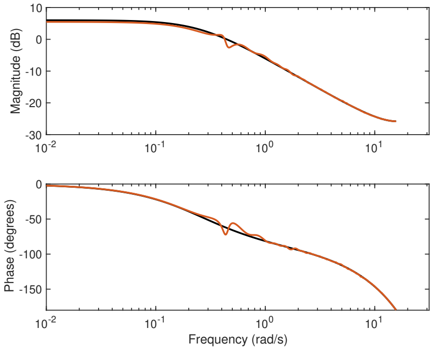
Balanced Model Reduction

A **balanced model reduction** is obtained by changing basis in a state-space model such that the states are ordered according to how much input energy is required to control them and how much energy they provide to the output.

- The first $n_1 < n$ states can then be kept and the remaining ones eliminated in order to obtain a lower-order model
- Matlab: `balred`

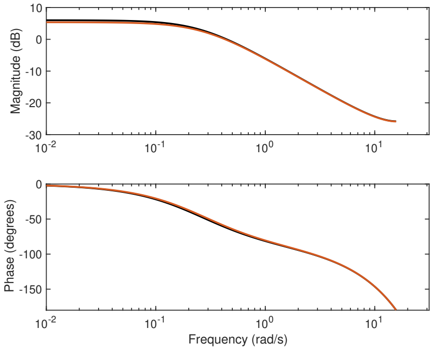
Closed-Loop Example: High-Order ARX

An ARX model structure with high model orders ($n_a = 50$, $n_b = 100$, $n_k = 1$) results in a rather accurate model estimate



Closed-Loop Example: ARX + Model Reduction

Computing a first order model from the high-order ARX model using balanced model reduction improves the accuracy



Direct PEM Approach...

Advantages:

- The direct PEM approach gives consistency and optimal accuracy provided that the model set contains the true system (including the noise description).
- The method works regardless of the complexity of the controller, and requires no knowledge about the character of the feedback.
- Unstable systems can be handled without problems as long as the closed-loop system is stable and the predictor is stable.
- No special algorithms and software are required.

Disadvantages:

- The order of the noise model must be high enough such that an accurate noise model is obtained.

The Indirect Approach

The **indirect approach**:

- Assume that the reference signal $r(t)$ is measured and that the controller is known.
- Identify the closed-loop system from reference signal $r(t)$ to output signal $y(t)$ (an open-loop problem).
- Retrieve the open-loop system by making use of the known controller.

The Indirect Approach...

The closed-loop system can be written

$$y(t) = \frac{G_0(q)}{1 + G_0(q)F_y(q)}\tilde{r}(t) + \frac{1}{1 + G_0(q)F_y(q)}v(t).$$

In the **indirect approach**, a model $\hat{G}_c(q)$ of the closed-loop system is estimated first from measurements of $\tilde{r}$ and y (an open-loop problem).

In a second step, an estimate $\hat{G}(q)$ of the open-loop transfer function is computed from the relation

$$\hat{G}_c(q) = \frac{\hat{G}(q)}{1 + \hat{G}(q)F_y(q)},$$

where the LTI controller $F_y(q)$ is assumed to be known.

The Indirect Approach...

In many cases (for example PEM), the model of the closed-loop system can be parameterized using a parameterization of the model of the open-loop system:

$$G_c(q, \theta) = \frac{G(q, \theta)}{1 + G(q, \theta) F_y(q)}$$

The retrieval of the open-loop model from the closed-loop one is then trivial.

The Indirect Approach...

Advantages:

- Any open-loop method can be used, including spectral analysis, instrumental variables and standard subspace methods.
- Consistent estimates of the system dynamics can be obtained also with a fixed (simplified) noise model.
- In the case of undermodeling, the resulting model $\hat{G}(q)$ will be a compromise between approximation of the true system dynamics and minimization of the model sensitivity function. This might be advantageous if the model is used for control design.

The Indirect Approach...

Disadvantages:

- The controller has to be known
- The reference signal has to be measured
- Any error in $F_y(q)$ (including saturation and anti-windup) will result in reduced accuracy of the model estimate $\hat{G}(q)$.
- The accuracy is typically worse than for the direct PEM approach (higher variance).

The Joint Input-output Approach

The **joint input-output approach**:

- Consider $y(t)$ *and* $u(t)$ as outputs of a system driven by $r(t)$ and noise.
- Recover information about the system from this joint model.
- Some methods assume that the reference signal is measured, but not all.

The Joint Input-output Approach...

Consider the complete closed-loop system:

$$y(t) = S_0(q)G_0(q)\tilde{r}(t) + S_0(q)v(t) = G_c(q)\tilde{r}(t) + \nu_1(t),$$
$$u(t) = S_0(q)\tilde{r}(t) - S_0(q)F_y(q)v(t) = G_{\tilde{r}u}(q)\tilde{r}(t) + \nu_2(t),$$

In **joint input-output approaches**, models that describe how both $u(t)$ and $y(t)$ depend on $\tilde{r}(t)$ are estimated (an open-loop problem).
Two options:

- Work with the complete model from $\tilde{r}(t)$ to $u(t)$ and $y(t)$ and consider the fact that the noise on the u channel is correlated with the noise on the y channel.
- Disregard the correlation between the two noise terms above and use the equations to define separate estimation problems.

The Joint Input-output Approach...

Since

$$G_c(q) = G_{\tilde{r}u}(q)G_0(q),$$

an estimate of the open-loop system can be obtained from estimates $\hat{G}_c(q)$ and $\hat{G}_{\tilde{r}u}(q)$ as

$$\hat{G}(q) = \frac{\hat{G}_c(q)}{\hat{G}_{\tilde{r}u}(q)}.$$

The Joint Input-output Approach...

The classic method for **spectral analysis** in closed-loop settings can be viewed as a joint input-output approach. In this method the open-loop system is estimated as

$$\hat{G}_N(e^{i\omega}) = \frac{\hat{\Phi}_{yr}^N(\omega)}{\hat{\Phi}_{ur}^N(\omega)}$$

(Here, $\hat{\Phi}_{yr}^N(\omega)$ and $\hat{\Phi}_{ur}^N(\omega)$ are smoothed estimates of the cross-spectra.)

The Joint Input-output Approach...

The **two-stage method** :

- Estimate the sensitivity function $S(q, \hat{\eta})$ from $\tilde{r}(t)$ and $u(t)$ (an open-loop problem) and use it to construct a "new" input signal:

$$\hat{u}(t) = S(q, \hat{\eta})\tilde{r}(t)$$

- Estimate a model of G_0 from $\hat{u}(t)$ to $y(t)$ (an open-loop problem).

The Joint Input-output Approach...

Underlying idea of the two-stage method:

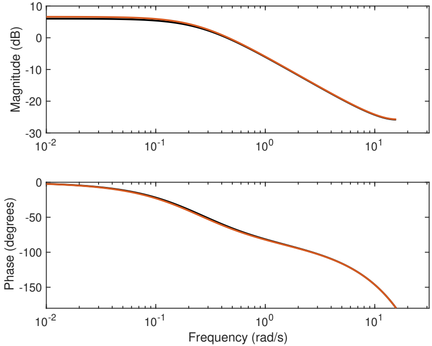
- If the estimation of the sensitivity model is successful (high enough model order, sufficient data record from an informative experiment), the remaining residual $\tilde{u}(t) = u(t) - \hat{u}(t)$ will be uncorrelated with $\tilde{r}(t)$.
- The output can be written

$$y(t) = G_0(q)\hat{u}(t) + v(t) + G_0(q)\tilde{u}(t).$$

Since the signal $\hat{u}(t)$ has been constructed from $\tilde{r}(t)$, it is uncorrelated with both $v(t)$ and $\tilde{u}(t)$. Hence, the estimation of $G_0(q)$ from $\hat{u}(t)$ to $y(t)$ can be viewed as an open-loop problem.

Closed-Loop Example: Two-Stage Method

Using the two-stage approach where the sensitivity function is modeled from r and u using an OE structure with $n_b = 3$, $n_f = 2$ and $n_k = 0$ and an OE model with correct orders ($n_b = n_f = n_k = 1$) is used to the describe how y depends on $\hat{u}$ results in an accurate estimate



The Joint Input-output Approach…

Advantages:

- Any open-loop method can be used, including spectral analysis, instrumental variables and standard subspace methods.
- Consistent estimates of the system dynamics can be obtained also with a fixed (simplified) noise model.
- Frequency weighting can be applied in a straightforward manner (in particular for the two-stage method).

Disadvantages:

- The reference signal has to be measured
- The accuracy is typically worse than for the direct PEM approach (higher variance). For the two-stage method, this is due to the extra "noise" $G_0(q)\tilde{u}(t)$.

Summary

- Closed-loop data cause correlation analysis, spectral analysis and subspace methods to give biased model estimates
- Approaches to closed-loop identification: The direct, indirect and joint input-output approach
- Direct PEM: Remember to use a flexible noise model
- Two-stage method: Estimate the sensitivity function and use it to generate a new input $\hat{u}$. Proceed as in open loop with $\hat{u}$ and y as input and output.

www.liu.se

LINKÖPING UNIVERSITY