Lecture 3

- Nonparametric Methods (Ch. 3)
- Input Signals (Ch. 5)
- Model Parameterizations (Ch. 6)

System Identification

Obtain a model of a system from measured inputs and outputs.

Type of model depends on application and system. Often we assume that the true system can be described as a LTI (linear time-invariant) system:

$$y(t) = G_0(q)u(t) + v(t) \tag{1a}$$

or, equivalently,

$$y(t) = \sum_{k=1}^{\infty} g_0(k)u(t-k) + v(t) \tag{1b}$$

Question: How do we determine the model $G_0(q)$ or $\{g_0(k)\}$?

Parametric models:

Postulate a model $G(q, \boldsymbol{\theta})$ parameterized by $\boldsymbol{\theta}$.

- Easy to use for simulation, control design, etc.
- Often accurate models.
- Requires some work...
- Example: FIR model

$$y(t) = u(t) + b_1 u(t-1) + b_2 u(t-2)$$
$$\Rightarrow G(q^{-1}, \boldsymbol{\theta}) = 1 + b_1 q^{-1} + b_2 q^{-2}, \quad \boldsymbol{\theta} = [b_1\, b_2]^T$$

Question: Can we determine $G_0(q)$ or $\{g_0(k)\}$ without postulating a parameterized model?

Nonparametric Identification

Nonparametric models:

Determine G_0 or $\{g_0(k)\}$ without parameterizing.

- Simple to obtain.
- Results often in graphs or tables which can not easily be used for simulation, etc.
- Often used to validate parametric models.
- Transient analysis, correlation analysis, frequency analysis, spectral analysis.

Transient Analysis

Impulse response analysis: Applying the input

$$u(t) = \begin{cases} k, & t = 0 \\ 0, & t \neq 0 \end{cases}$$

to (1b) gives the output

$$y(t) = k g_0(t) + v(t)$$

which motivates the impulse response estimate

$$\hat{g}(t) = \frac{y(t)}{k}$$

Step-response analysis Applying the input

$$u(t) = \begin{cases} k, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

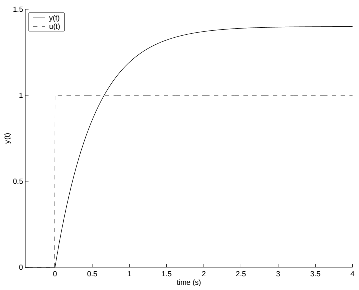
gives the output

$$y(t) = k \sum_{k=1}^{t} g_0(k) + v(t)$$

which motivates the impulse response estimate

$$\hat{g}(t) = \frac{y(t) - y(t-1)}{k}$$

Ex: Step-response (true – solid, measured – $\times$)



Transient analysis

- Input taken as impulse or step.
- Model consists of recorded output, or an estimate of $g_0(k)$.
- Convenient for deriving crude models. Gives estimates of dominating time constants, time delays and static gain.
- Sensitive to noise.
- Poor excitation.

Correlation Analysis

System:

$$y(t) = \sum_{k=1}^{\infty} g_0(k)u(t-k) + v(t)$$

where $u(t)$ is a stochastic process which is independent of $v(t)$.
Multiplying by $u(t-\tau)$ and taking expectation yields

$$r_{yu}(\tau) = \sum_{k=1}^{\infty} g_0(k) r_u(\tau - k)$$

which is known as the Wiener-Hopf equation.

In practice, truncate the sum and solve the resulting system of eq.

$$\hat{r}_{yu}(\tau) = \sum_{k=1}^{M} \hat{g}(k)\hat{r}_u(\tau - k)$$

Estimates of the covariance functions.

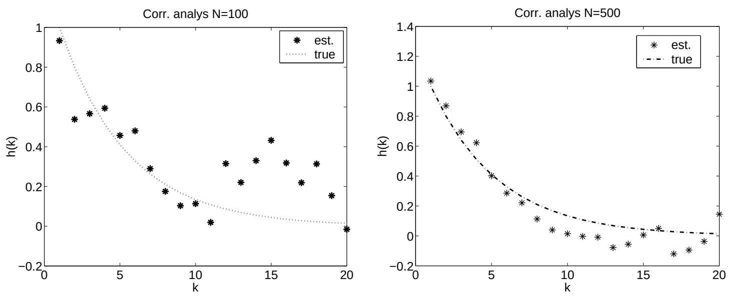
- First choice:

$$\hat{r}_{yu}(\tau) = \frac{1}{N}\sum_{k=1}^{N-\tau} y(k+\tau)u(k) \quad (\tau \geq 0)$$

- Second choice:

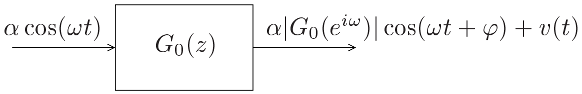
$$\hat{r}_{yu}(\tau) = \frac{1}{N-\tau}\sum_{k=1}^{N-\tau} y(k+\tau)u(k) \quad (\tau \geq 0)$$

Which one to prefer?



Frequency Analysis

Estimate $G_0(e^{i\omega})$!



- Repeat experiment for different ω $(t = 1, \ldots, N)$.
- Determine the phase shift and the amplitude of the output.
- Results in a Bode plot $(|G_0(e^{i\omega})|$ and $\arg G_0(e^{i\omega}))$.
- Sensitive to noise. Require long experiments.
- Gives basic information about the system.

Spectral Analysis

- The correspondence of the Wiener-Hopf equation in the frequency domain is given by:

$$\Phi_{yu}(\omega) = G(e^{-i\omega})\Phi_u(\omega)$$

- An estimate of the transfer function can be obtained as:

$$\hat{G}(e^{-i\omega}) = \hat{\Phi}_{yu}(\omega)/\hat{\Phi}_u(\omega)$$

- Use estimates of the spectral densities, *e.g.*,

$$\hat{\Phi}_{yu}(\omega) = \frac{1}{2\pi N}\sum_{\tau=-N}^{N} \hat{r}_{yu}(\tau)e^{-i\tau\omega}$$

- Errors in $\hat{r}_{yu}(\tau)$ are summed together $\Rightarrow$ not consistent!
  - N large $\Rightarrow$ total (square) error is large even if the error in $\hat{r}_{yu}(\tau)$ is small for all τ.
  - $\hat{r}_{yu}(\tau)$ decreases slowly $\Rightarrow$ poor estimate of $\hat{r}_{yu}(\tau)$ for large values of τ.
- Better estimates are obtained if a lag window, $w(t)$, is used:

$$\hat{\Phi}_{yu}(\omega) = \frac{1}{2\pi}\sum_{\tau=-N}^{N} \hat{r}_{yu}(\tau)w(\tau)e^{-i\tau\omega}$$

- Length of lag window (M) - compromise between bias and variance (high resolution and reducing erratic fluctuations).

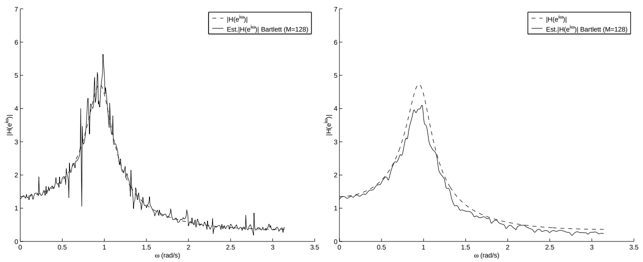


Figure 1: Spectral analysis, $N = 256$: Left: Periodogram. Right: Bartlett window $M = 128$.

Summary - Nonparametric Methods

- Results often in graph or table (step response, weighting function, transfer function etc.).
- Transient analysis (step-response, impulse response).
- Frequency analysis (sinusoidal input).
- Correlation analysis (weighting function estimate).
- Spectral analysis (transfer function estimate).
- Useful for obtaining crude estimates of time constants, cut-off frequencies etc. or for model validation.

Input Signals (Ch. 5)

The quality of the model is dependent on an appropriate choice of input signal.

Examples of useful input signals are:

- Step function.
- Pseudorandom binary sequence (PRBS).
- Autoregressive moving average process (ARMA).
- Sum of sinusoids.

Most often the input signal is characterized by its first and second order moments:

$$m = Eu(t)$$
$$r(\tau) = E(u(t+\tau) - m)(u(t) - m)^T$$

and/or its spectral density:

$$\Phi(\omega) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} r(\tau) e^{-i\tau\omega}$$

Rem: Deterministic signals

$$Eu(t) = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} u(t)$$

Step Function

$$u(t) = \begin{cases} k, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

Properties

- Mostly used for transient analysis: overshoot, static gain, major time constants.
- Limited usability for parametric modeling.

PseudoRandom Binary Sequence (PRBS)

A PRBS $u(t)$ is a periodic, deterministic signal with white-noise-like properties.

$$u(t) = \text{rem}(A(q)u(t), 2) = \text{rem}(a_1 u(t-1) + \cdots + a_n u(t-n), 2)$$

Properties

- The signal shifts between two levels in a certain fashion depending on $A(q)$.
- Spectral characteristics is determined by $A(q)$ and, in particular, by the period length $M = 2^n - 1$.
- Deterministic sequence behaving as noise (reproducibility).

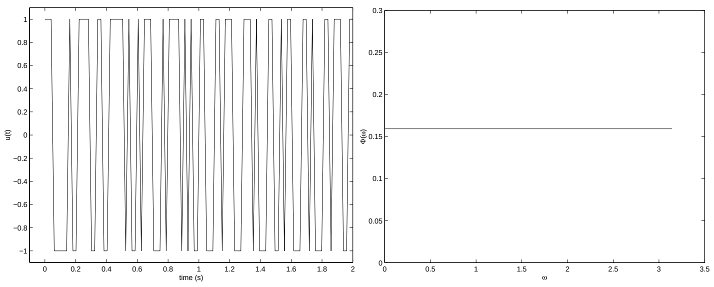


Figure 2: PRBS sequence, p=0.5, $M = \infty$. Left: Example of realization. Right: Spectral density.

ARMA Process

$$A(q^{-1})y(t) = C(q^{-1})e(t)$$

where $e(t)$ is white noise with $Ee(t) = 0$ and $Ee^2(t) = \lambda^2$.

Properties

- The signal $y(t)$ can be obtained by filtering white noise.
- The filters can be chosen to obtain almost any desired frequency characteristics.
- The spectral density of an ARMA process $y(t)$ is given by:

$$\Phi_y(\omega) = \frac{\lambda^2}{2\pi}\left|\frac{C(e^{i\omega})}{A(e^{i\omega})}\right|^2$$

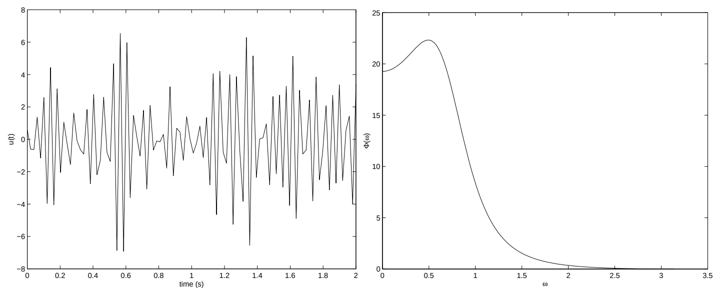


Figure 3: ARMA(2,2) process. Left: Example of realization. Right: Spectral density.

Sum of Sinusoids

$$u(t) = \sum_{m=1}^{M} a_m \sin(\omega_m t + \varphi_m)$$

Properties

- User parameters: a_m, ω_m and φ_m.
- Covariance function given by:

$$r(\tau) = \sum_{m=1}^{M} \frac{a_m^2}{2} \cos(\omega_m t + \varphi_m)$$

- Spectral density given by:

$$\Phi(\omega) = \sum_{m=1}^{M} \frac{a_m^2}{4}[\delta(\omega - \omega_m) + \delta(\omega + \omega_m)]$$

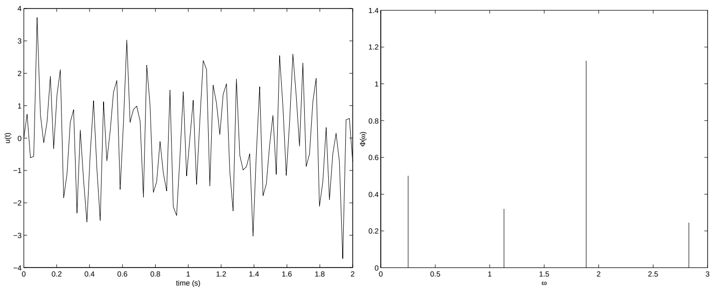


Figure 4: Sum of 4 sinusoids. Left: Signal. Right: Spectral density.

Persistent Excitation

To obtain estimates of a parametric model the input signal has to be "rich" enough to excite all modes of the system.

A input signal is said to be persistently exciting ($p.e.$) of order n if:

(i) The following limit exists:

$$r_u(\tau) = \lim_{N\to\infty} \frac{1}{N}\sum_{t=1}^{N} u(t+\tau)u^T(t)$$

Rem: $u(t)$ ergodic implies

$$r_u(\tau) = E\,u(t+\tau)u^T(t)$$

(ii) The matrix:

$$\mathbf{R}_u(n) = \begin{pmatrix} r_u(0) & r_u(1) & \cdots & r_u(n-1) \\ r_u(-1) & r_u(0) & & \vdots \\ \vdots & & \ddots & \\ r_u(1-n) & \cdots & & r_u(0) \end{pmatrix}$$

is positive definite.

- Another definition: $\det \mathbf{R}_u(n) \neq 0$.
- And another: $u(t)$ is $p.e.$ of order n if $\Phi_u(\omega) \neq 0$ on at least n points in the interval $-\pi < \omega \leq \pi$.

An input signal that is $p.e.$ of order $2n$ can be used to consistently estimate a parametric model of order $\leq n$.

- A step function is $p.e.$ of order 1.
- A PRBS with period M is $p.e.$ of order M.
- An ARMA process is $p.e.$ of any finite order.
- A sum of m sinusoids is $p.e.$ of order $2m$ (if $\omega_m \neq 0$ and $\omega_m \neq \pi$).

Another important observation!

A parametric model becomes more accurate in the frequency region where the input signal has the major part of its energy.

A physical process is often of low frequency character $\Rightarrow$ use low-pass filtered signal as input.

Summary - Input Signals

- The choice of input signal determines the quality of the final parametric model.
- The obtained parametric model is more accurate in frequency regions where the input signal contains much energy.
- An input signal has to be rich enough to excite all interesting modes of the system (persistently exciting of sufficiently high order).
- In practice there might be some restrictions on the input.

Model Parameterizations (Ch.6)

Mathematical models can be derived from:
- Physical modeling
- Identification

Classification of mathematical models
- SISO - MIMO
- Linear models - Nonlinear models
- Parametric models - Nonparametric models
- Time-invariant models - Time-varying models
- Time domain models - Frequency domain models
- Discrete-time models - Continuous-time models
- Deterministic models - Stochastic models

General model structure (SISO)

$$\begin{aligned} y(t) &= G(q^{-1},\boldsymbol{\theta})u(t) + H(q^{-1},\boldsymbol{\theta})e(t) \\ G(q^{-1},\boldsymbol{\theta}) &= \frac{B(q^{-1})}{A(q^{-1})} = \frac{b_1 q^{-n_k} + b_2 q^{-n_k-1} + \ldots + b_{n_b} q^{-n_k-n_b+1}}{1 + a_1 q^{-1} + \ldots + a_{n_a} q^{-n_a}} \\ H(q^{-1},\boldsymbol{\theta}) &= \frac{C(q^{-1})}{D(q^{-1})} = \frac{1 + c_1 q^{-1} + \ldots + c_{n_c} q^{-n_c}}{1 + d_1 q^{-1} + \ldots + d_{n_d} q^{-n_d}} \end{aligned}$$

where $e(t)$ is white noise with variance λ^2, and

$$\boldsymbol{\theta} = \begin{bmatrix} a_1 & \ldots & a_{n_a} & b_1 & \ldots & b_{n_b} & c_1 & \ldots & c_{n_c} & d_1 & \ldots & d_{n_d} \end{bmatrix}^T$$

Rem: Often $\lambda^2 = \lambda^2(\boldsymbol{\theta})$.

Assumptions:

- Time-delay $n_k \geq 1 \Rightarrow G(0, \boldsymbol{\theta}) = 0$ (also $H(0, \boldsymbol{\theta}) = 1$)
- $H^{-1}(q^{-1}, \boldsymbol{\theta})$ and $H^{-1}(q^{-1}, \boldsymbol{\theta})G(q^{-1}, \boldsymbol{\theta})$ are asymptotically stable. Often $H(q^{-1}, \boldsymbol{\theta})$ also needs to be asym. stable.

Commonly used simplified models:

ARMAX

$$A(q^{-1})y(t) = B(q^{-1})u(t) + C(q^{-1})e(t)$$

Rem: $A(q^{-1})$ describes the dynamics: Here the input and the noise is governed by the same dynamics.

ARX

$$A(q^{-1})y(t) = B(q^{-1})u(t) + e(t)$$

FIR

$$y(t) = B(q^{-1})u(t) + e(t)$$

OE

$$y(t) = \frac{B(q^{-1})}{A(q^{-1})}u(t) + e(t)$$

Time series models (no input signal)

ARMA

$$A(q^{-1})y(t) = C(q^{-1})e(t)$$

AR

$$A(q^{-1})y(t) = e(t)$$

MA

$$y(t) = C(q^{-1})e(t)$$

Applications: Times-series modeling is useful in various disciplines, such as, economy, astrophysics, speech, etc.

Uniqueness and Identifiability

Uniqueness:

Let the true system $\mathcal{S}$ be described by G_0, H_0 and λ_0^2.

Introduce the set

$D_T = \{\boldsymbol{\theta} | G_0 = G(q^{-1}, \boldsymbol{\theta}), H_0 = H(q^{-1}, \boldsymbol{\theta}), \lambda_0^2 = \lambda^2(\boldsymbol{\theta})\}$.

- D_T empty $\Rightarrow$ underparameterized model structure.
- D_T contains several points $\Rightarrow$ overparameterized model structure. Numerical problems are likely to occur.
- D_T contains one point $\Rightarrow$ Ideal case! The system is **uniquely** described by the model structure ($\boldsymbol{\theta} = \boldsymbol{\theta}_0$).

Identifiability:

- System Identifiable (SI): D_T is nonempty, and $\hat{\boldsymbol{\theta}} \to D_T$ as $N \to \infty$.
- Parameter Identifiable (PI): If the system is SI and D_T contains one point ($\hat{\boldsymbol{\theta}} \to \boldsymbol{\theta}_0$).

In other words, if the choice of *model*, *input signal* and *identification method* makes the estimated parameter vector, $\hat{\boldsymbol{\theta}}$, converge (with probability one, as $N \to \infty$)) to a parameter vector that perfectly describes the system as the number of data points tend to infinity then the system is said to be **system identifiable**. If the system is uniquely described by the model *and* system identifiable then the system is said to be **parameter identifiable**.

Summary - Model Parameterizations

- It is essential that the model structure suits the actual system.
- Many standard model structures are available with different approaches of how to model the influence of the input and the disturbances.
- Finding the correct, or best, model structure and model order is normally an iterative procedure (Ch. 11).
- A model should ideally be unique and the complete experimental set-up should be such that the system is parameter identifiable.
- Not included: Ex 6.3, 6.4, 6.6, continuous-time models. "Kursivt": Ex 6.5.