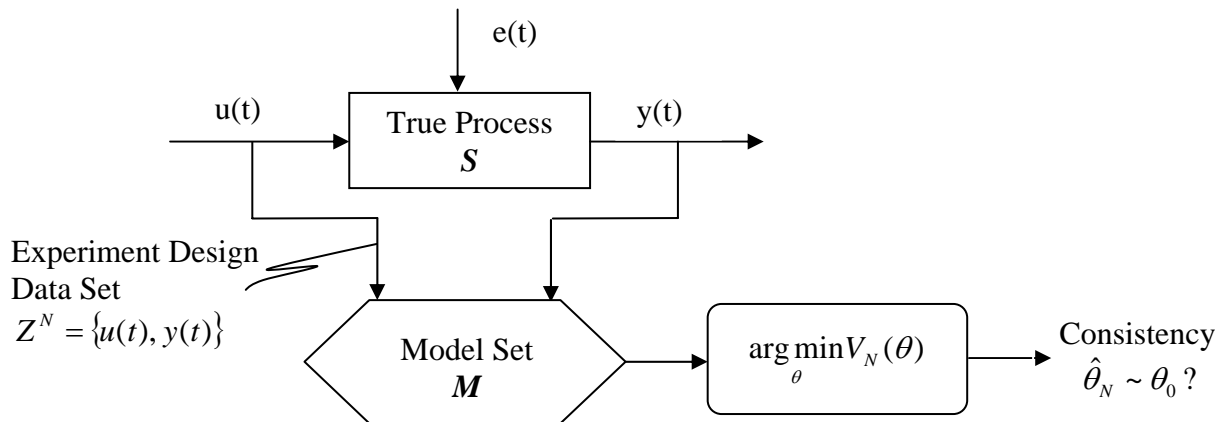


2.160 System Identification, Estimation, and Learning
Lecture Notes No. 15
 April 12, 2006

Part 3 System Identification

Perspective of System Identification Theory



Key Questions:

Q1: Is a given data set *informative enough* to uniquely determine a model from a given model set?

Does Z contain sufficient information to distinguish any two models in M ?

Q2: Is merely minimizing $V_N(\theta)$ good enough to obtain the true (unbiased) model?

What if the true model is not involved in the model set?

How is the model-data fitting influenced by noise characteristics and input properties?

Q3: How accurate is the estimated model?

How much variance, expected error, etc.?

How much data needed?

How to design experiments

Key Results

- Informative experiment and persistent excitation
- Consistent (unbiased) estimate
- Signal to noise ration
- Asymptotic variance
- Input design: Pseudo Random Binary signal
- Accuracy-variance trade-off
- System order estimate: Model selection

Mathematical tools for Part 3 system identification

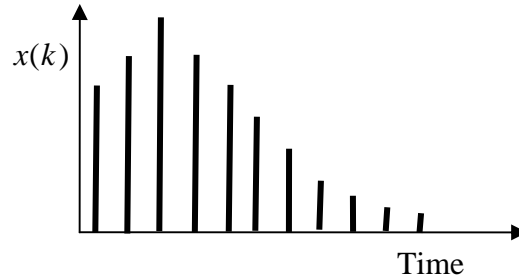
- Discrete Fourier transform and spectral analysis
- Central limit theorems
- Random processes: wide-sense stationary process, ergodic process, etc.

10 Frequency Domain Analysis

10.1 Discrete Fourier Transform and Power Spectrum

Discrete Fourier transform of a sampled-data system: $x(k)$,

$$X(\omega) = \sum_{k=-\infty}^{+\infty} x(k)e^{-i\omega k} \quad (1)$$



Note that $X(\omega)$ is a 2π -periodic function:

$$X(\omega + 2\pi n) = \sum_{k=-\infty}^{\infty} x(k)e^{-i(\omega+2\pi n)k} = \underbrace{\sum_{k=-\infty}^{\infty} x(k)e^{-i\omega k}}_{X(\omega)} \underbrace{e^{-i2\pi nk}}_1 = X(\omega) \quad (2)$$

$$\therefore X(\omega + 2\pi n) = X(\omega) \quad (3)$$

A periodic function can be expanded to a Fourier series expansion. Therefore, we can write

$$\int_{-\pi}^{\pi} X(\omega)e^{i\omega\ell} d\omega = \int_{-\pi}^{\pi} \sum_{k=-\infty}^{\infty} x(k)e^{-i\omega k} e^{i\omega\ell} d\omega = \sum_{k=-\infty}^{\infty} x(k) \int_{-\pi}^{\pi} e^{i\omega(\ell-k)} d\omega = \begin{cases} 2\pi x(\ell) & : k = \ell \\ 0 & : k \neq \ell \end{cases} \quad (4)$$

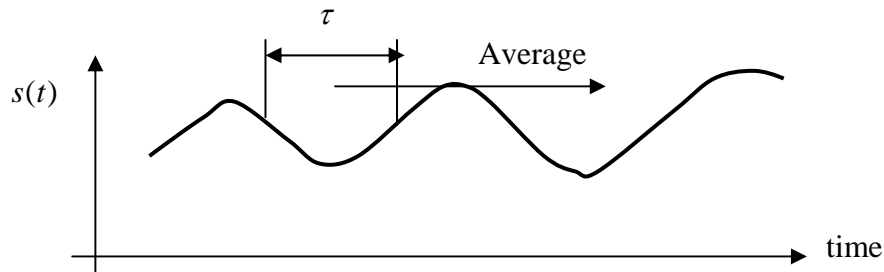
which means that the inverse transform of $X(\omega)$ exists:

$$x(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(\omega)e^{i\omega n} d\omega \quad (5)$$

Power Spectrum

Consider a deterministic, bounded sequence $\{s(t)\}$ for which the following limit exists:

$$R_s(\tau) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N s(t)s(t+\tau) : \quad \text{Covariance Function} \quad (6)$$



The power spectrum of $\{s(t)\}$ is defined as the Fourier transform of auto-covariance function $R_s(\tau)$

$$\Phi_s(\omega) = \sum_{\tau=-\infty}^{\infty} R_s(\tau)e^{-i\tau\omega} \quad (7)$$

Inverse Transform:

$$R_s(\tau) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_s(\omega)e^{i\tau\omega} d\omega \quad (8)$$

A special case is:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N s^2(t) = R_s(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_s(\omega)e^{i\tau\omega} d\omega \quad (9)$$

Important! We will often use this formula to obtain the mean of a squared signal $s(t)$, a special case of $\tau = 0$.

White Noise

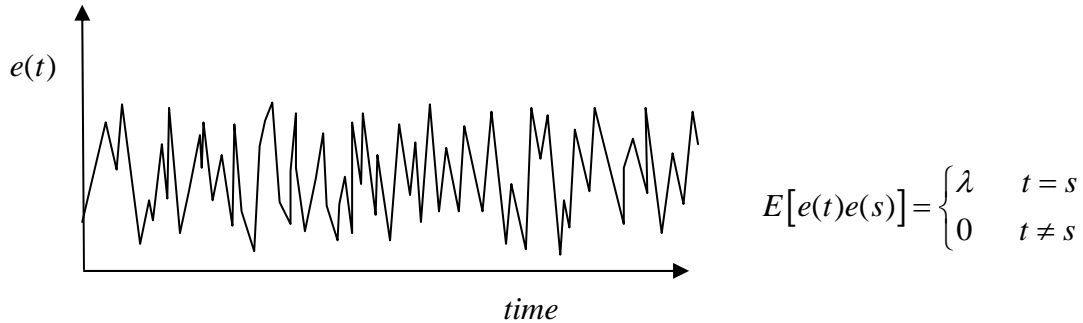
We have seen “White Noise” in many sections of the previous lectures.

We defined $\{e(t)\}$ as a sequence of independent random variables with zero mean values and covariance λ :

$$E[e(t)e(s)] = \begin{cases} \lambda & t = s \\ 0 & t \neq s \end{cases} \quad (10)$$

Now its characteristics are formally defined using Power Spectrum.

The auto-covariance of the random process $e(t)$ is given by



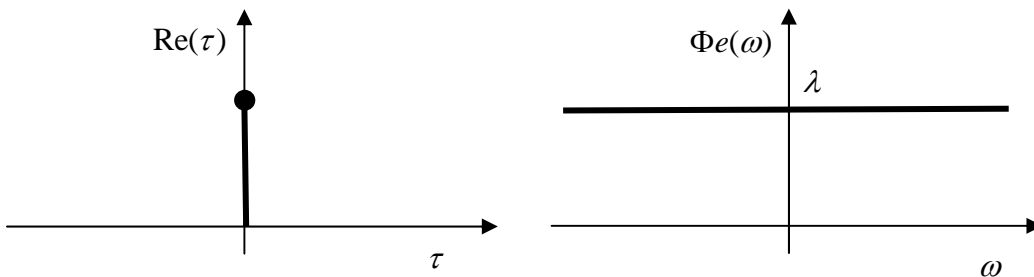
$$R_s(\tau) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N e(t)e(t+\tau) \Rightarrow E[e(t)e(t+\tau)] = \lambda \delta(\tau) \quad (11)$$

The left hand side of the above expression is a time average, while the right hand side is an ensemble average. If these two averages are the same, the process is called **ergodic**. We assume this ergodicity for most of the processes. See the discussion at the end of this lecture notes.

For the above equation the Power spectrum is given by

$$\Phi_e(\omega) = \sum_{\tau=-\infty}^{\infty} R_e(\tau)e^{-i\tau\omega} = \sum_{\tau=-\infty}^{\infty} \lambda \delta(\tau)e^{-i\tau\omega} = \lambda \quad (12)$$

The figures below show the plots of the auto-covariance $Re(\tau)$ against time and the corresponding power spectrum $\Phi_e(\omega)$ against frequency. Note that the power spectrum plot is constant for the entire frequency. In optics, this means that the light has a uniform distribution over the entire wave length, that is, “White”. This is why the random process $e(t)$ is called “White Noise”.



A colored random signal can be created with the White noise going through a dynamical process. The following theorem plays a major role in many of the analyses involved in system identification.

Theorem

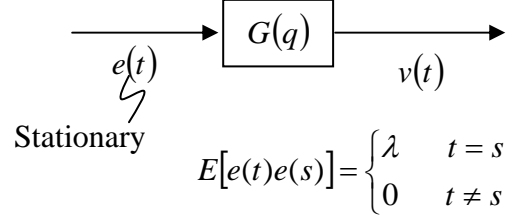
Let $H(q)$ be the transfer function of a (BIBO) stable process with a White noise input $e(t)$ of variance λ .

$$v(t) = H(q)e(t) \quad (13)$$

The power spectrum of $v(t)$ is given by

$$\Phi_v(\omega) = \lambda |H(e^{i\omega})|^2 \quad (14)$$

where $|\cdot|$ is the norm of a complex number.



Proof $v(k) = \sum_{l=0}^{\infty} h(l)e(k-l)$

$$= v(k-\tau) = \sum_{m=0}^{\infty} h(m)e(k-\tau-m)$$

combining these two

$$(15) \quad R_v(\tau) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N v(k)v(k-\tau)$$

$$= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \sum_{l=0}^{\infty} h(l)e(k-l) \sum_{m=0}^{\infty} h(m)e(k-\tau-m)$$

$$R_v(\tau) = \sum_{l=0}^{\infty} \sum_{m=0}^{\infty} h(l)h(m) \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N e(k-l)h(m)e(k-\tau-m)$$

||

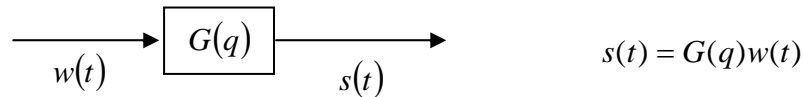
$$\lambda \delta(l-\tau-m) = \begin{cases} \lambda & l = \tau + m \\ 0 & \text{otherwise} \end{cases}$$

$$\therefore R_v(\tau) = \lambda \sum_{l=\max(0,\tau)}^{\infty} h(l)h(l-\tau) \quad : \quad h(l) = 0 \quad \text{for } l < 0$$

The power spectrum is then given by

$$\begin{aligned}
\Phi_v(\omega) &= \sum_{\tau=-\infty}^{\infty} R_v(\tau) e^{-i\omega\tau} \\
&= \lambda \sum_{\tau=-\infty}^{\infty} \sum_{l=\max(0,\tau)}^{\infty} h(l) e^{-i\omega l} h(l-\tau) e^{-i\omega(l-\tau)} \\
(16) \quad &\text{Re placing } l - \tau \text{ by } s \\
&= \lambda \sum_{l=0}^{\infty} h(l) e^{-il\omega} \sum_{s=0}^{\infty} h(s) e^{i\omega s} \\
&= \lambda H(e^{i\omega}) \cdot H(e^{-i\omega}) = \lambda |H(e^{i\omega})|^2
\end{aligned}$$

For a stationary signal with spectrum $\{\omega(t)\}$ with spectrum $\Phi_w(\omega)$



The power spectrum and cross spectrum are given by

$$(17) \quad \Phi_s(\omega) = |G(e^{i\omega})|^2 \Phi_w(\omega)$$

$$(18) \quad \Phi_{sw}(\omega) = G(e^{i\omega}) \Phi_w(\omega)$$

10.2 Applying spectral Analysis to System Identification

$$(19) \quad y(t) = G(q)u(t) + H(q)e(t)$$

Deterministic

Stochastic

Two issues to clarify for mathematical rigor

- 1) Strictly speaking, the process is not stationary; input $u(t)$ drives the system.

Therefore, the covariance function $R_s(\tau)$ cannot be defined in general.

$$R_s(\tau) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N s(t)s(t-\tau)$$

The stationary property that we need is the existence of this covariance. Then, we extend the definition to the one assuming the existence of $R_s(\tau)$ Quasi-Stationary (wide sense stationary)

Important theories and techniques of system identification are dependent upon the spectra of the involved signals, i.e. only the second-order properties, and not on any higher-order properties.

Deterministic and stochastic processes are mixed. $\{e(t)\}, \{v(t)\}$ are stochastic processes. The covariance function for this type of variable must be given an ensemble mean:

$$(20) \quad R_s(\tau) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N E[s(t)s(t-\tau)]$$

If this is equivalent to:

$$(21) \quad R_s(\tau) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N s(t)s(t-\tau)$$

treatment will be very convenient, since we need to consider only one realization of the stochastic process, rather than considering the whole collection of ensemble average.

This is an ergodicity problem. For dynamical systems,

this ergodicity holds for signals generated through uniformly stable filters $G(q)$

$$(22) \quad s(t) = G(q)e(t)$$

Under this assumption, the theoretical boundary between stochastic and deterministic processes is low.