System Identification

GianAndrea Müller

October 28, 2018

Contents

1 System Identification

2 Definitions

Definition 1. *A system is said to be **time invariant** if the response to a certain input is not depending on absolute time.*

Definition 2. *A system is said to be **linear** if its output response to a linear combination of inputs is the same as the linear combination of the output responses of the individual inputs.*

Definition 3. *A system is said to be **causal** if the output at a certain time depends on the input up to that time only.*

Definition 4. *A process is said to be **stationary** if it does not depend on time.*

3 Frequency Domain Methods

3.1 Sampling Operation

$$\boxed{y(k) = y(t)|_{t=kT, k=0,1,2,\ldots}} \text{ Sampling with period } T$$

3.2 Fourier Series of Periodic Signals

$$\boxed{X(e^{j\omega_m}) = \sum_{k=0}^{M-1} x(k) e^{-j\omega_m k}}$$

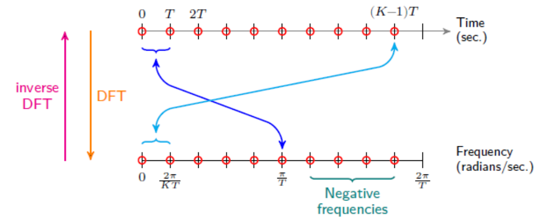
$$\omega_m = \frac{2\pi m}{M} = \omega_0$$

Non-negative frequencies are $m = 0$ to $m = M/s$.
They correspond to: $\omega_m = 0, \frac{2\pi}{M}, \frac{4\pi}{M}, \ldots, \frac{2\pi(M/2-1)}{M}, \pi$.

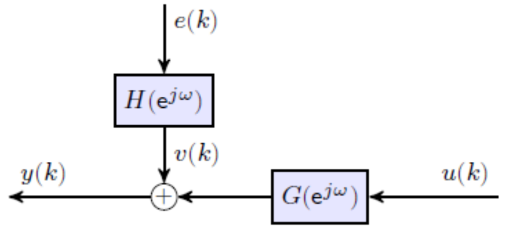
M	number of samples	
$\omega_0 = \frac{2\pi}{M}$	fundamental frequency ($y(k)$)	[rad]
T	sampling time	[s]
$\tau_p = MT$	period	[s]
$\omega_0 = \frac{2\pi}{\tau_p}$	fundamental frequency ($y(t)$)	[rad s^{-1}]

$$0, \underbrace{\frac{2\pi}{\tau_p}}_{\text{Fundamental frequency}}, \underbrace{2\left(\frac{2\pi}{\tau_p}\right), \ldots, \frac{M}{2}\left(\frac{2\pi}{\tau_p}\right)}_{\text{Harmonics}}$$

Definition 5. *The highest frequency $\omega_u = \omega_{M/2} = \frac{\pi}{T}$ is called the **Nyquist frequency**.*



4 Spectral Estimation



$$Y(j\omega) = G(j\omega)U(j\omega) \quad \text{Transfer function}$$

$$Y(e^{j\omega} = G(e^{j\omega}U(e^{j\omega}) \quad \text{Discrete time TF}$$

$$\frac{\omega_u}{2\pi} = \frac{r}{NT}$$

ω_u	input frequency	[rad s^{-1}]
N	calculation length	[ ]
T	experiment duration?	[s]
r	some integer	[ ]

$$u(k) = \alpha \cos(\omega_u k), \ k = 0, 1, \ldots, K-1 \ \text{ with } K \geq N \quad \text{Input}$$

$$y(k) = \alpha \left|G(e^{j\omega_u})\right| \cos(\omega_u k + \theta(\omega_u)) + v(k) + \text{transient} \quad \text{Output}$$

where $\theta(\omega_u) = arg(G(e^{j\omega_u})$

4.1 Sinusoidal correlation methods

Correlation functions:

$$I_c(N) = \frac{1}{N} \sum_{k=0}^{N-1} y(k) \cos(\omega_u k)$$

$$I_s(N) = \frac{1}{N} \sum_{k=0}^{N-1} y(k) \sin(\omega_u k)$$

To calculate those from the data:

$$I_c(N) = \frac{\alpha}{2}\left|G(e^{j\omega_u})\right|\cos(\theta(\omega_u) + \frac{\alpha}{2}\left|G(e^{j\omega_u})\right|\frac{1}{N}\sum_{k=0}^{N-1}\cos(2\omega_u k + \theta(\omega_u)) + \frac{1}{N}\sum_{k=0}^{N-1} v(k)\cos(\omega_u k)$$

If the noise, $v(k)$ is sufficiently uncorrelated then the variance satisfies,

$$\lim_{N\to\infty} \text{var}\left\{\frac{1}{N}\sum_{k=0}^{N-1} v(k)\cos(\omega_u k)\right\} = 0$$

with a convergence rate of $1/N$.
Thus in the limit $N \to \infty$,

$$E\{I_c(N)\} \to \frac{\alpha}{2}\left|G(e^{j\omega_u})\right|\cos(\theta(\omega_u))$$
$$E\{I_s(N)\} \to -\frac{\alpha}{2}\left|G(e^{j\omega_u})\right|\sin(\theta(\omega_u))$$

and since $\lim_{N\to\infty} \text{var}\{I_c(N)\} = 0, \quad \lim_{N\to\infty} \text{var}\{I_s(N)\} = 0$
The transfer function can be estimated via:

$$\boxed{\hat{G}_N(e^{j\omega_u}) = \frac{I_c(N) - jI_s(N)}{\alpha/2}}$$

- Advantages
  - Energy is concentrated at the frequencies of interest.
  - Amplitude of $u(k)$ can easily be tuned as a function of frequency.
  - Easy to avoid saturation and tune signal/noise (S/N) ratio.
- Disadvantages
  - A large amount of data is required.
  - Significant amount of time required for experiments.
  - Some processes won't allow sinusoidal inputs.

5 Frequency Domain Methods

- Autocorrelation
- Crosscorrelation
- Frequency domain representation
- Spectral density (energy or power)

$$x(k),\ k = -\infty, \dots, \infty \quad \text{Discrete-time domain signal}$$

$$\boxed{X(e^{j\omega}) = \sum_{k=-\infty}^{\infty} x(k)e^{-j\omega k}} \text{ Fourier Transform}$$

$$\boxed{x(k) = \frac{1}{2\pi}\int_{-pi}\pi X(e^{j\omega})e^{j\omega k}d\omega} \text{ Inverse Fourier Transform}$$

where $k = -\infty, \dots, \infty$

5.1 Finite Energy Signal

5.1.1 Energy Spectral Density (Finite Energy Signal)

If $x(k)$ is a finite energy signal,

$$||x(k)||_2^2 = \sum_{k=-\infty}^{\infty} |x(k)|^2 < \infty$$

$$\boxed{S_x(e^{j\omega}) = |X(e^{j\omega})|^2} \text{ Energy Spectral Density}$$

For all following calculations of the energy spectral density finiteness is assumed.

5.1.2 Autocorrelation (Finite Energy Signal)

$$\boxed{R_x(\tau) = \sum_{k=-\infty}^{\infty} x(k)x(k-\tau), \quad \tau = -\infty, \dots, 0, \dots, \infty}$$

The spectral density is the Fourier Transform of the autocorrelation:

$$\sum_{\tau=-\infty}^{\infty} R_x(\tau)e^{-j\omega\tau} = S_x(e^{j\omega})$$

$\boxed{\text{xcorr(u)}}$ Autocorrelation

$\boxed{\text{xcorr(u,v)}}$ Crosscorrelation

5.2 Discrete Periodic Signal

$$x(k) = x(k+M), \quad \forall\ k \in \{-\infty, \infty\} \quad \text{Periodic signal}$$

$$\omega_0 = \frac{2\pi}{M} \quad \text{Fundamental frequency}$$

- There are only M unique harmonics of the sinusoid $e^{j\omega_0}$.
- The non-negative harmonic frequencies are,

$$e^{jn\omega_0},\ n = 0, 1, \dots, M/2$$

5.2.1 Discrete Fourier Series (Discrete Periodic Signal)

$$\boxed{X(e^{j\omega_n}) = \sum_{k=0}^{N-1} x(k)e^{-j\omega_n k}, \text{ where } \omega_n = \frac{2\pi n}{N} = n\omega_0}$$

$$\boxed{x(k) = \frac{1}{N}\sum_{k=0}^{N-1} X(e^{j\omega_n})e^{j\omega_n k}} \text{ Inverse Transform}$$

5.2.2 Autocorrelation (Discrete Periodic Signal)

$$\boxed{R_x(\tau) = \frac{1}{N}\sum_{k=0}^{N-1} x(k)x(k-\tau)}$$

The Fourier transform of $R_x(\tau)$ is now defined as the **power spectral density**, since it is normalized with the signal length.

$$\boxed{\phi_x(e^{j\omega_n}) = \sum_{\tau=0}^{N-1} R_x(\tau)e^{-j\omega_n\tau} = \frac{1}{N}|X(e^{j\omega_n})|^2}$$

The energy in a single period is:

$$\sum_{k=0}^{N-1} |x(k)|^2 = \sum_{n=0}^{N-1} \phi_x(e^{j\omega_n})$$

5.2.3 Cross-Correlation (Discrete Periodic Signal)

$$\boxed{R_{yu}(\tau) = \frac{1}{N}\sum_{k=0}^{N-1} y(k)u(k-\tau)}$$

The Fourier transform of $R_{yu}(\tau)$ is now defined as the **cross-spectral density**.

$$\boxed{\phi_{yu}(e^{j\omega_n}) = \sum_{\tau=0}^{N-1} R_{yu}(\tau)e^{-j\omega_n\tau}) = \frac{1}{N}Y(e^{j\omega_n})U^*(e^{j\omega_n})}$$

5.3 Random Signal

Normally distributed noise:

$$\boxed{e(k) \in \mathcal{N}(0,\lambda) \Rightarrow \begin{cases} \mathbb{E}\left[e(k)\right] = 0 \text{ (zero mean)} \\ \mathbb{E}\left[|e(k)|^2\right] = \lambda \text{ (varianve)} \end{cases}}$$

The $e(k)$ are independent and identically distributed (i.i.d.).

5.3.1 Autocovariance (Random Signal)

$$\begin{aligned} R_x(\tau) &= \mathbb{E}\left[x(k)x(k-\tau)\right] \\ &= \mathbb{E}\left[x(k)x^*(k-\tau)\right] \text{ (in the complex case)} \\ &= \mathbb{E}\left[x(k)x^*(x-\tau)\right] \text{ (in the multivariable case)} \end{aligned}$$

General (non-stationary, non-zero mean) case:

$$\begin{aligned} R_x(s,t) &= \mathbb{E}\left[(x(s) - \mathbb{E}\left[x\right])(x(t) - \mathbb{E}\left[E\right])\right] \\ &= \mathbb{E}\left[x(s)x(t)\right] \text{ (if zero mean)} \\ &= R_x(s-t) \text{ (if stationary)} \end{aligned}$$

Further properties are

- $R_x(-\tau) = R_x^*(\tau)$
- $R_x(0) \geq |R_x(\tau)|\ \forall \tau > 0$

5.3.2 Power Spectral Density (Random Signal)

$$\boxed{\phi_x(e^{j\omega}) := \sum_{\tau=-\infty}^{\infty} R_x(\tau)e^{-j\omega\tau} \text{ where } \omega \in [-\pi,\pi)}$$

For a zero-mean random signal:

$$\lim_{N\to\infty} \frac{1}{N}\sum_{k=0}^{N-1} |x(k)|^2 = \operatorname{Var}\left(x(k)\right) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \phi_x(e^{j\omega})d\omega$$

Further properties are

- $\phi_x(e^{j\omega}) \in \mathbb{R}$
- $\phi_x(e^{j\omega}) \geq 0\ \forall\ \omega$
- $\phi_x(e^{j\omega}) = \phi_x(e^{-j\omega})$ for all real-valued $x(k)$

5.3.3 Cross-Covariance (Random Signal)

$$\boxed{R_{yu}(\tau) = \mathbb{E}\left[(y(k) - \mathbb{E}\left[y(k)\right])(u(k-\tau) - \mathbb{E}\left[u(k)\right]\right]}$$

For zero mean signals:

$$R_{yu}(\tau) = \mathbb{E}\left[y(k)u(k-\tau)\right]$$

Joint stationarity is required to make the definition dependent on τ only.
If $R_{yu}(\tau) = 0$ for all τ then $y(k)$ and $u(k)$ are uncorrelated.

5.3.4 Cross Power Spectral Density (Random Signal)

$$\boxed{\phi_{yu}(e^{j\omega}) = \sum_{\tau=-\infty}^{\infty} R_{yu}(\tau)e^{-j\omega\tau},\ \omega \in [-\pi, \pi)}$$

The inverse is,

$$R_{yu}(\tau) = \tfrac{1}{2\pi}\int_{-\pi}^{\pi} \phi_{yu}(e^{j\omega})e^{j\omega\tau}d\omega$$

5.4 Finite Length Signal

5.4.1 Discrete-Fourier Transform (Finite Length Signal)

$$\boxed{X_N(e^{j\omega_n}) = \sum_{k=0}^{N-1} x(k)e^{-j\omega_n k}, \text{ where } \omega_n = \tfrac{2\pi n}{N}}$$

The inverse DFT is

$$x(k) = \tfrac{1}{N}\sum_{n=0}^{N-1} X_N(e^{j\omega_n})e^{j\omega_n k}, \quad k = 0, \ldots, N-1$$

5.4.2 Periodogram (Finite Length Signal)

$$\boxed{\tfrac{1}{N}\left|V_N(e^{j\omega})\right|^2}$$

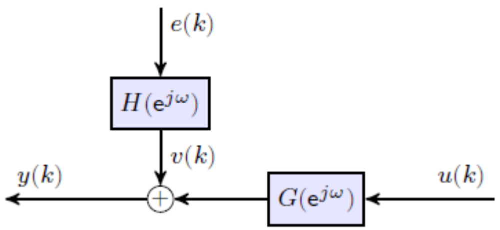
An asymptotically unbiased estimator of the spectrum is

$$\lim_{N\to\infty} \mathbb{E}\left[\tfrac{1}{N}|V_N(e^{j\omega}|^2\right] = \phi_v(\omega)$$

This assumes that the autocorrelation decays quickly enough:

$$\lim_{N\to\infty} \tfrac{1}{N}\sum_{\tau=-N}^{N} |\tau R_v(\tau)| = 0$$

6 ETFE



Linear, time-invariant system, $g(l)$:

$$y(k) = \sum_{l=0}^{\infty} g(l)u(k-l) + v(k), \quad k = 0, 1, \ldots$$

Assumptions:

1. causal system: $g(l) = 0, \forall l < 0$
2. noise: $E\{v(k)\} = 0$, zero mean, stationary

Given $\{u(k), y(k)\}$ find an estimate $\hat{G}(e^{j\omega})$ such that it fits the $G(e^{j\omega})$.

$$\boxed{\text{Bias}(\hat{G}))G - E\{\hat{G}\}} \text{ Bias}$$

$$\boxed{\text{var}((\hat{G}) = E\left\{|\hat{G} - E\{\hat{G}\}|^2\right\}} \text{ Variance}$$

$$\boxed{\text{MSE}(\hat{G}) = E\left\{|G - \hat{G}|^2\right\}} \text{ Mean-square error}$$

Note that MSE$(\hat{G})$ =var$(\hat{G})$+Bias$^2(\hat{G})$.

6.1 Input-output relationship

For finite energy signals:

$$y(k) = \sum_{l=0}^{\infty} g(l)u(k-l) + v(k)$$

$$Y(e^{j\omega}) = G(e^{j\omega})U(e^{j\omega}) + V(e^{j\omega})$$

which in the idealized case leads to:

$$\boxed{\tfrac{Y(e^{j\omega})}{U(e^{j\omega})} = G(e^{j\omega}) + \tfrac{V(e^{j}\omega)}{U(e^{j\omega})} \approx G(e^{j\omega})}$$

In reality we only have N samples:

$$\underbrace{Y_N(e^{j\omega_n})}_{\text{length-N DFT}} = \sum_{k=0}^{N-1} y(k)e^{-j\omega_n k} \approx \sum_{k=-\infty}^{\infty} y(k)e^{-j\omega_n k} = Y(e^{j\omega_n}$$

$$\underbrace{U_N(e^{j\omega_n})}_{\text{length-N DFT}} = \sum_{k=0}^{N-1} u(k)e^{-j\omega_n k} \approx \sum_{k=-\infty}^{\infty} u(k)e^{-j\omega_n k} = U(e^{j\omega_n})$$

$$\boxed{\hat{G}_N(e^{j\omega_n}) := \tfrac{Y_N(e^{j\omega_n})}{U_N(e^{j\omega_n})}} \text{ ETFE}$$

6.2 Periodic input case

Period M inputs: $u(k) = u(k+M)$
If $sM = N$ for an integer s, the fourier series over N samples is equal to the real fourier series!

$$U_N(e^{j\omega_n}) = U(e^{j\omega_n}) \forall \omega_n = \frac{2\pi n}{N},\ n = 0, \ldots, N-1$$

Then

$$Y_N(e^{j\omega_n}) = G(e^{j\omega_n}) U_N(e^{j\omega_n}) + V_N(e^{j\omega_n})$$

$$\hat{G}_N(e^{j\omega_n}) = G(e^{j\omega_n}) + \frac{V_N(e^{j\omega_n})}{U_N(e^{j\omega_n})}$$

Bias:

$$E\{\hat{G}_N(e^{j\omega_n})\} = G(e^{j\omega_n}) + E\left\{\frac{V_N(e^{j\omega_n})}{U_N(e^{j\omega_n})}\right\} = G(e^{j\omega_n})$$

when assuming zero mean noise. Thus for periodic inputs with N being an integer number of periods, **the ETFE is unbiased**.
Variance:
For the unbiased case:

$$E\left\{|\hat{G}_N(e^{j\omega_n}) - G(e^{j\omega_n})|^2\right\} = \frac{\phi_v(e^{j\omega_n}) + \frac{2}{N}c}{\frac{1}{N}|U_N(e^{j\omega_n})|^2}$$

where $|c| \leq C = \sum_{\tau=1}^{\infty} |\tau R_v(\tau)|$ is assumed to be finite.
For estimates at different frequencies ($\omega_n \neq \omega_i$):

$$E\left\{(\hat{G}_N(e^{j\omega_n}) - G(e^{j\omega_n}))(\hat{G}_N(e^{-j\omega_i}) - G(e^{-j\omega_i}))\right\} = 0$$

Transient responses:
Initial transient corrupts the measurement

$$y(k) = G(u_{periodic}(k) W_{[0,N-1]}(k)) + v(k)$$

with the window function:

$$W_{[0,N-1]}(k) = \begin{cases} 1 & \text{if} 0 \leq 0 < N \\ 0 & \text{otherwise} \end{cases}$$

For all outputs up to time $k = N-1$

$$y(k) = Gu_{periodic}(k) - \underbrace{G(u_{periodic} W_{(-\infty,-1)})}_{r(k)} + v(k)$$

$$Y_N(e^{j\omega_n}) = G(e^{j\omega_n}) U_N(e^{j\omega_n}) + R_N(e^{j\omega_n}) + V_N(e^{j\omega_n})$$

The input in negative time, which is present in a ideal periodic input, and missing in a real periodic input, has an influence on positive time, which is described by $r(k)$.

When using a periodic signal multiple times the resulting DFt does not contain more information, since in a periodic signal there are only a certain number of frequencies contained, but the energy in those frequencies increases!
Transient bias error:

$$\hat{G}e^{(j\omega_n)} = \frac{Y_N e^{(j\omega_n)}}{U_N e^{(j\omega_n)}} = Ge^{(j\omega_n)} + \frac{R_N e^{(j\omega_n)}}{U_N e^{(j\omega_n)}} + \frac{V_N e^{(j\omega_n)}}{U_N e^{(j\omega_n)}}$$

For periodic $u(k)$
As $N = mM, m \to \infty$

$$|U_N e^{(j\omega_n)}| = m|U_M e^{(j\omega_n)}|$$

For random $u(k)$
As $N \to \infty$

$$E\{|U_N e^{(j\omega_n)}|\} \to \sqrt{N}\sqrt{\phi_u e^{(j\omega_n)}}$$

Thus

$$\boxed{\left|\frac{R_N e^{(j\omega_n)}}{U_N e^{(j\omega_n)}}\right| \to 0 \text{ with rate } \begin{cases} \frac{1}{N} & \text{for periodic input} \\ \frac{1}{\sqrt{N}} & \text{for random inputs} \end{cases}}$$

A fix for getting rid of the influence of the transient response: Get rid of the first period.

6.3 Spectral Transformations

If $v(k) = 0$

$$\phi_y e^{(j\omega_n)} = Ge^{(j\omega_n)} \phi_u e^{(j\omega_n)} G^T e^{(j\omega_n)}$$

where $G^T e^{(j\omega_n)}$ is the complex conjugate of $Ge^{(j\omega_n)}$.
If $v(k) \neq 0$ and uncorrelated

$$\phi_y e^{(j\omega_n)} = |Ge^{(j\omega_n)}|^2 \phi_u e^{(j\omega_n)} + |He^{(j\omega_n)}|^2$$

But this approach has no more phase information. For that reason use the cross spectrum:

$$\phi_{yu} e^{(j\omega_n)} = Ge^{(j\omega_n)} \phi_u e^{(j\omega_n)} + \phi_{uv} e^{(j\omega_n)} = Ge^{(j\omega_n)} \phi_u e^{(j\omega_n)}$$

if $u(k)$ and $v(k)$ are uncorrelated.

$$\boxed{\frac{\hat{\phi}_{yu} e^{(j\omega_n)}}{\hat{\phi}_u e^{(j\omega_n)}}} \text{ Spectral estimation}$$

where

$$\phi_y e^{(j\omega_n)} = |Ge^{(j\omega_n)}|^2 \phi_u e^{(j\omega_n)} + \phi_v e^{(j\omega_n)}$$

$$\phi_{yu} e^{(j\omega_n)} = G e^{(j\omega_n)} \phi_u e^{(j\omega_n)}$$

The periodogram is an asymptotically unbiased estimator of the spectrum given $\lim_{n\to\infty} \frac{1}{N} \sum_{\tau=-\mathrm{N}}^{N} |\tau R_v(\tau)| = 0$

$$\frac{1}{N} |V_N e^{(j\omega_n)}|^2 \quad \text{periodogram}$$

$$\boxed{\lim_{N\to\infty} E\left\{\frac{1}{N}|V_N e^{(j\omega_n)}|^2\right\} = \phi_v e^{(j\omega_n)}}$$

The autocovariance of the noise for stochastic $v(k)$ is described as:

$$\hat{R}_v(\tau) = \begin{cases} \frac{1}{N-|\tau|} \sum_{k=\tau}^{N_1} v(k)v(k-\tau), & \text{for } \tau \geq 0 \\ \frac{1}{N-|\tau|} \sum_{k=0}^{N+\tau-1} v(k)v(k-\tau), & \text{for } \tau < 0 \end{cases}$$

This is an unbiased estimator of $R_v(\tau)$: $E\{\hat{R}_v(\tau) = R_v(\tau)\}$

$$\boxed{\hat{\phi}_v e^{(j\omega_n)} = \sum_{\tau=-N+1}^{N-1} \hat{R}_v(\tau) e^{-j\omega\tau}}$$

Thus the spectral estimate is:

$$\hat{\phi}_v(e^{j\omega}) = \sum_{\tau=-N+1}^{N-1} \hat{R}_v(\tau) e^{-j\omega\tau}$$

6.3.1 Spectral estimation for periodic signals

Periodic signal $x(k)$ with period M, $N = mM$ for some integer m

$$R_x(\tau) = \frac{1}{M} \sum_{k=0}^{M-1} x(k)x(k-\tau)$$

The power spectral density can be calculated and is equal to the periodogram:

$$\boxed{\phi_x e^{(j\omega_n)} = \sum_{\tau=0}^{M-1} R_x(\tau) e^{-j\omega_n\tau} = \frac{1}{M}\left|X_M e^{(j\omega_n)}\right|^2}$$

6.3.2 Spectral estimation (more general case)

Alternative autocorrelation estimate:

$$\hat{R}_x(\tau) = \begin{cases} \frac{1}{N} \sum_{k=\tau}^{N-1} x(k)x(k-\tau), & \text{for } \tau \geq 0 \\ \frac{1}{N} \sum_{k=0}^{N+\tau-1} x(k)x(k-\tau), & \text{for } \tau < 0 \end{cases}$$

Periodic $x(k)$: unbiased (exact) if $N = mM$

Random $x(k)$ biased $\mathrm{E}\left\{\hat{R}_x(\tau)\right\} = \frac{N-|\tau|}{N} R_x(\tau)$.
asymptotically biased as $N \to \infty, \tau/N \to 0$

7 Averaging and Smoothing

Multiple experiments $u_r(k), y_r(k), r = 1\ldots, R, k = 0, \ldots, K-1$

$$\hat{G} e^{(j\omega_n)} = \sum_{r=1}^{R} \alpha_r \hat{G}_r e^{(j\omega_n)}$$

where $\sum_{r=1}^{G} \alpha_r = 1$ and for calculating the average $\alpha_r = \frac{1}{R}$.

The averaging can be optimized by selecting α_r such that the variance $\sigma_r^2 e^{(j\omega_n)}$ is minimized.

$$\mathrm{Var}\left(\hat{G} e^{(j\omega_n)}\right) = \mathrm{Var}\left(\sum_{r=1}^{R} \alpha_r e^{(j\omega_n)} \hat{G}_r e^{(j\omega_n)}\right) = \sum_{r=1}^{R} \alpha_r^2 \sigma_r^2 e^{(j\omega_n)}$$

This is minimized by

$$\boxed{\alpha_r e^{(j\omega_n)} = \frac{1/\sigma_r^2 e^{(j\omega_n)}}{\sum_{r=1}^{T} 1/\sigma_r^2 e^{(j\omega_n)}}}$$

Thus the signal is weighted inversely proportional to the variance.

Thus if $\mathrm{Var}\left(\hat{G}_r e^{(j\omega_n)}\right) = \frac{\phi_v e^{(j\omega_n)}}{\frac{1}{N}|U_r e^{(j\omega_n)}|^2}$ then $\alpha_r e^{(j\omega_n)} = \frac{|U_r e^{(j\omega_n)}|^2}{\sum_{r=1}^{R} |U_r e^{(j\omega_n)}|^2}$.

The best result is obtained if the input is the same for all r, which will lead to a reduction of the variance as follows:

$$\mathrm{Var}\left(\hat{G} e^{(j\omega_n)}\right) = \frac{\mathrm{Var}\left(\hat{G}_r e^{(j\omega_n)}\right)}{R}$$

Biased estimates will reduce the improvement in variance.

- Since we are adding complex numbers the magnitude of the average is not equal to the average of the magnitudes r_i.

7.1 Bias-variance trade-offs in data record splitting

Divide a data record into smaller parts for averaging:

$$\{u(k), y(k)\}, k = 0, \ldots, K-1$$

Choose R records and calculation length N, such that $NR \leq K$:

$$u_r(n) = u(rN + n)$$

And average the resulting estimates:

$$\boxed{\hat{G}e^{(j\omega_n)} = \frac{1}{R}\sum_{r=0}^{R-1}\hat{G}_r e^{(j\omega_n)} = \frac{1}{R}\sum_{r=0}^{R-1}\frac{\hat{Y}_r e^{(j\omega_n)}}{\hat{U}_r e^{(j\omega_n)}}}$$

As R increases:

- The number of points calculated, N decreases.
- The variance decreases (by up to $1/R$).
- The bias increases (due to non-periodicity transients).

Mean-square error

- Transient bias grows linearly with the number of data splits.
- Variance decays with a rate of up to 1/(number of averages).

What if there is no option of running periodic input experiments? $\rightarrow$ exploit the assumed smoothness of the underlying system.

7.2 Smoothing the ETFE

Assume the true system to be close to constant for a range of frequencies: $G(e^{j\omega_{n+r}}) \approx G(e^{j\omega_n})$ for $r = 0, \pm 1, \ldots, \pm r$.
The minimum variance smoothed estimate is:

$$\tilde{G}_N e^{(j\omega_n)} = \frac{\sum\limits_{r=-R}^{R}\alpha_r \hat{G}_N(e^{j\omega_{n+r}})}{\sum\limits_{r=-R}^{R}\alpha_r}, \qquad \alpha_r = \frac{\frac{1}{N}|U_N(e^{j\omega_{r+n}})|^2}{\phi_v(e^{j\omega_{n+r}})}$$

The summation above can then be approximated by an integral:

$$\approx \frac{\int_{\omega_{n-r}}^{\omega_{n+r}}\alpha(e^{j\zeta})\hat{G}_N(e^{j\zeta})d\zeta}{\int_{\omega_{n-r}}^{\omega_{n+r}}\alpha(e^{j\zeta})d\zeta}, \qquad \text{with } \alpha(e^{j\zeta}) = \frac{\frac{1}{N}|U_N(e^{j\zeta})|^2}{\phi_v(e^{j\zeta})}$$

Which can be reformulated using a smoothing window:

$$\tilde{G}_N e^{(j\omega_n)} = \frac{\frac{1}{2\pi}\int_{-\pi}^{\pi}W_\gamma(e^{j(\zeta-\omega_n)})\alpha(e^{j\zeta})\hat{G}_N(e^{j\zeta})d\zeta}{\frac{1}{2\pi}\int_{-\pi}^{\pi}W_\gamma(e^{j(\zeta-\omega_n)})\alpha(e^{j\zeta})d\zeta} \qquad \text{with } \alpha(e^{j\zeta}) = \frac{\frac{1}{N}|U_N(e^{j\zeta})|^2}{\phi_v(e^{j\zeta})}$$

7.2.1 Assumptions on $\phi_v(e^{j\omega})$

Assume $\phi_v(e^{j\omega})$ is also a smooth function of frequency.

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}W_\gamma(e^{j(\zeta-\omega_n)})\left|\frac{1}{\phi_v(e^{j\zeta})} - \frac{1}{\phi e^{(j\omega_n)}}\right|d\zeta \approx 0$$

Then use,

$$\alpha(e^{j\zeta}) = \frac{\frac{1}{N}|U_N(e^{j\zeta})|^2}{\phi_v e^{(j\omega_n)}}$$

to get

$$\boxed{\tilde{G}_N e^{(j\omega_n)} = \frac{\frac{1}{2\pi}\int_{-pi}^{\pi}W_\gamma(e^{-j(\zeta-\omega_n)})\frac{1}{N}|U_N(e^{j\zeta})|^2\hat{G}_N(e^{j\zeta})d\zeta}{\frac{1}{2\pi}\int_{-\pi}^{\pi}W_\gamma(e^{j(\zeta-\omega_n)})\frac{1}{N}|U_N(e^{j\zeta})^2 d\zeta}}$$

The wider the frequency window (decreasing γ)

- the more adjacent frequencies included in the smoothness estimate.
- the smoother the result.
- the lower the noise induced variance.
- the higher the bias.

7.2.2 Characteristic windows

$$\boxed{W_\gamma(e^{j\omega}) = \frac{1}{\gamma}\left(\frac{\sin\gamma\omega/2}{\sin\omega/2}\right)^2} \text{ Bartlett}$$

$$\boxed{W_\gamma(e^{j\omega}) = \frac{1}{2}D_\gamma(\omega) + \frac{1}{4}D_\gamma(\omega - \pi/\gamma) + \frac{1}{4}D_\gamma(\omega + \pi/\gamma)} \text{ Hann}$$

where

$$D_\gamma(\omega) = \frac{\sin\omega(\gamma+0.5)}{\sin\omega/2}$$

Properties of window functions:

- $\frac{1}{2\pi}\int_{-\pi}^{\pi}W_\gamma(e^{j\zeta})d\zeta = 1$
- $\int_{-\pi}^{\pi}\zeta W_\gamma(e^{j\zeta})d\zeta = 0$
- $M(\gamma) := \int_{-\pi}^{\pi}\zeta^2 W_\gamma(e^{j\zeta})d\zeta$
- $\bar{W}(\gamma) := 2\pi\int_{-\pi}^{\pi}W_\gamma^2(e^{j\zeta})d\zeta$

Bartlett	$M(\gamma) = \frac{2.78}{\gamma}$,	$\bar{W}(\gamma) \approx 0.67\gamma$	(for $\gamma > 5$)
Hamming	$M(\gamma) = \frac{\pi^2}{2\gamma^2}$,	$\bar{W}(\gamma) \approx 0.75\gamma$	(for $\gamma > 5$)

- $M(\gamma)$ gives an idea of the bias effect.
- $\bar{W}(\gamma)$ gives an idea of the variance effect.

7.2.3 Asymptotic bias properties

$$\mathrm{E}\left\{\tilde{G}e^{(j\omega_n)} - \mathrm{E}\left\{Ge^{(j\omega_n)}\right\}\right\} = \mathrm{E}\left\{\tilde{G}e^{(j\omega_n)} - Ge^{(j\omega_n)}\right\} =$$

$$M(\gamma)\left(\frac{1}{2}\underbrace{G''e^{(j\omega_n)}}_{\text{curvature}} + \underbrace{G'e^{(j\omega_n)}}_{slope}\frac{\phi'_u e^{(j\omega_n)}}{\phi_u e^{(j\omega_n)}}\right) + H.O.T.$$

Increasing γ

- makes the frequency window smaller.
- averages over fewer frequency values.
- makes $M(\gamma)$ smaller
- reduces the bias of the smoothed estimate $\tilde{G}e^{(j\omega_n)}$

7.2.4 Asymptotic variance properties

$$\mathrm{E}\left\{(\tilde{G}e^{(j\omega_n)} - \mathrm{E}\left\{\tilde{G}e^{(j\omega_n)}\right\})^2\right\} = \frac{1}{N}\bar{W}(\gamma)\frac{\phi_v e^{(j\omega_n)}}{\phi_u e^{(j\omega_n)}} + H.O.T.$$

Increasing γ

- makes the frequency window narrower.
- averages over fewer frequency values.
- makes $\bar{W}_\gamma$ larger.
- increases the variance of the smoothed estimate $\tilde{G}e^{(j\omega_n)}$.

7.2.5 Asymptotic MSE properties

$$\mathrm{E}\left\{|\tilde{G}e^{(j\omega_n)} - Ge^{(j\omega_n)}|^2\right\} \approx M^2(\gamma)|Fe^{(j\omega_n)}|^2 + \frac{1}{N}\bar{W}(\gamma)\frac{\phi_v e^{(j\omega_n)}}{\phi_u e^{(j\omega_n)}}$$

where

$$Fe^{(j\omega_n)} = \frac{1}{2}G''e^{(j\omega_n)} + G'e^{(j\omega_n)}\frac{\phi'_u e^{(j\omega_n)}}{\phi_u e^{(j\omega_n)}}$$

If $M(\gamma) = M/\gamma^2$ and $\bar{W}(\gamma) = \bar{W}\gamma$ then MSE is minised by:

$$\boxed{\gamma_{optimal} = \left(\frac{4M^2|Fe^{(j\omega_n)}|^2\phi_u e^{(j\omega_n)}}{\bar{W}\phi_v e^{(j\omega_n)}}\right)^{1/5} N^{1/5}}$$

and

$$\text{MSE at } \gamma_{optimal} \approx CN^{-4/5}$$

8 Windowing and Input Signals

$$\phi_{yu}(e^{j\omega}) = G(e^{j\omega})\phi_u(e^{j\omega})$$

$$\boxed{\hat{G}e^{(j\omega_n)} = \frac{\hat{\phi}_{yu}e^{(j\omega_n)}}{\hat{\phi}_u e^{(j\omega_n)}}}$$

Recall that the smoothed ETFE is:

$$\tilde{G}_N e^{(j\omega_n)} = \frac{\frac{1}{2\pi}\int_{-\pi}^{\pi} W_\gamma(e^{j(\zeta-\omega_n)})\frac{1}{N}|U_N(e^{j\zeta})|^2\hat{G}_N(e^{j\zeta})d\zeta}{\frac{1}{2\pi}\int_{-\pi}^{\pi} W_\gamma(e^{j(\zeta-\omega_n)})\frac{1}{N}|U_n(e^{j\zeta})|^2 d\zeta}$$

The denominator term approaches $\frac{1}{2\pi}\int_{-\pi}^{\pi} W_\gamma(e^{j(\zeta-\omega_n)})\phi e^{(j\omega_n)}d\zeta$ as $N \to \infty$.

If in addition $W_\gamma(e^{j\omega})$ is concentrated around $\zeta = 0$ (i.e. $\gamma/N \to 0$) then the denominator term approaches $\phi_u e^{(j\omega_n)}$ as $N \to \infty$.

This motivates the smoothed spectral estimate:

$$\boxed{\tilde{\phi}_u e^{(j\omega_n)} = \frac{1}{2\pi}\int_{-\pi}^{\pi} W_\gamma(e^{j(\zeta-\omega_n)})\frac{1}{N}|U_N(e^{j\omega})|^2 d\zeta}$$

Similarly the numerator approaches ϕ_{yu} as $N \to \infty$:

$$\boxed{\tilde{\phi}_{yu}e^{(j\omega_n)} = \frac{1}{2\pi}\int_{-\pi}^{\pi} W_\gamma(e^{j(\zeta-\omega_n)})\frac{1}{N}|U_N(e^{j\omega})|^2\hat{G}_N(e^{j\omega})d\zeta}$$

For this reason the smoothed ETFE is equal to the smoothed spectral estimate for $N \to \infty$.

8.1 Time domain windows

Define, via the inverse Fourier transform a time domain window:

$$\omega_\gamma(\tau) = \frac{1}{2\pi}\int_{-\pi}^{\pi} W_\gamma(e^{j\omega})e^{j\zeta\tau}d\zeta$$

Then the smoothed input spectral estimate $\tilde{\phi}_u e^{(j\omega_n)}$ is:

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} W_\gamma(e^{j(\zeta-\omega_n)})\frac{1}{N}|U_N(e^{j\omega})|^2 d\zeta \approx \sum_{\tau-\infty}^{\infty}\omega_\gamma(\tau)\hat{R}_u(\tau)e^{-j\tau\omega_n}$$

where

$$\omega_\gamma = \begin{cases} 0 & \text{for } \tau < -\gamma \\ > 0 & \text{for } -\gamma \leq \tau \leq \gamma \\ 0 & \text{for } \tau > \gamma \end{cases}$$

where often $\gamma << N$, which enables the faster calculated redefinition:

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} W_\gamma(e^{j(\zeta-\omega_n)})\frac{1}{N}|U_N(e^{j\omega})|^2 d\zeta \approx \sum_{\tau-\gamma}^{\gamma}\omega_\gamma(\tau)\hat{R}_u(\tau)e^{-j\tau\omega_n}$$

The cross spectral estimate can also be formulated as a convolution in the frequency domain which leads to the analogous formulation to the spectral estimate of u:

$$\tilde{\phi}_u e^{(j\omega_n)} = \sum_{\tau=-\gamma}^{\gamma}\omega_\gamma(\tau)\hat{R}_u(\tau)e^{-j\tau\omega_n}$$

$$\tilde{\phi}_{yu} e^{(j\omega_n)} = \sum_{\tau=-\gamma}^{\gamma}\omega_\gamma(\tau)\hat{R}_{yu}(\tau)e^{-j\tau\omega_n}$$

8.1.1 Window characteristics

Decreasing γ: narrower $\omega_\gamma(\tau)$, wider $W_\gamma(e^{j\omega})$

- the more frequencies, $\hat{G}e^{(j\omega_n)}$ included in the smoothing.
- the fewer $\hat{R}(\tau)$ estimates included in the smoothing.
- the smoother the result.
- the lower the noise induced variance.
- the **higher the bias**.

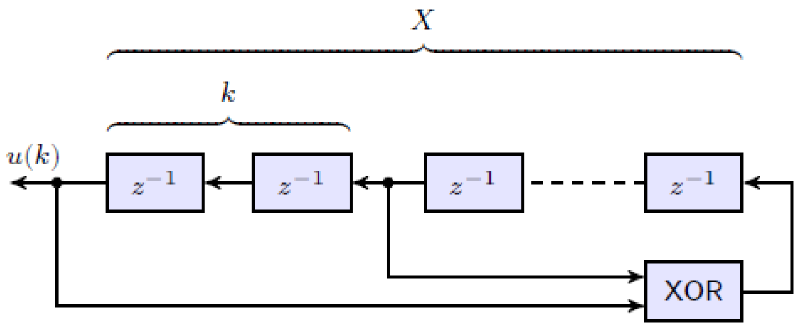
8.2 Input Signals

- Steps
- Doublet
- Sinusiods, Chirpts, Multi-Sines
- Filtered white noise
- Pseudo-Random Binary Signals (PRBS)

8.2.1 PRBS

$$u(k) = a \text{ or } -a$$

Shift-register generation



X	k
3	1 or 2
4	1 or 3
5	2 or 3
6	1 or 5
7	1 or 6
8	—
9	4 or 5
10	3 or 7

Periodicity

Periodic with period equal to at most $M = 2^X - 1$.

$$R_u(\tau) = \frac{1}{N}\sum_{k=0}^{N-1}u(k)u(k-\tau) = \begin{cases} a^2 & \text{if } \tau = 0 \\ \frac{-a^2}{2^X-1} & \text{if } \tau \neq 0 \end{cases}$$

Definition 6. ***Run length*** *defines how long the signal stays high.*

The run length distribution of $u(k)$ is then:

1/2 runs of length 1

1/4 runs of length 2

1/8 runs of length 3

$\vdots$

Other properties:

- Equal energy at all frequencies.
- Autocorrelation zero except for $\tau = 0$.

8.2.2 Multi-sinusoidal signals

$$u(k) = \sum_{s=1}^{S}\sqrt{2\alpha_s}\cos(\omega_s kT + phi_s)$$

where T is the sampling period, $\omega_s = \frac{2\pi}{T_P}$, $\frac{T_P}{T} = N$, $S \leq \frac{N}{2}$.

Choose N to be a power of 2 for efficient FFT calculations.

$$\sum_{s=1}^{S} \alpha_s = 1 \quad \text{Total signal power}$$

Schroeder phasing
Select the phases ϕ_s such that the minimize the peak amplitude:

$$\phi_s = 2\pi \sum_{j=1}^{s} j\alpha_s.$$

for equal power in each sinusoids:

$$\alpha_s = 1/S \text{ and } \phi_s = \frac{pi(s^2+s)}{S}$$

9 Residual Spectra, Coherency, Aperiodicty, Offsets and Drifts

9.1 Residual Spectrum

9.1.1 Estimating $\phi_v e^{(j\omega_n)}$

$$v(k) = y(k) - G(e^{j\omega})u(k)$$

$$\boxed{\tilde{\phi}_v e^{(j\omega_n)} \approx \frac{1}{N}\frac{1}{2\pi}\int_{-\pi}^{\pi} W_\gamma(e^{j(\zeta-\omega_n)})\left|Y_N(e^{j\omega}) - \tilde{G}(e^{j\omega})U_N(e^{j\omega})\right|^2 d\zeta \approx \tilde{\phi} e^{(j\omega_n)} - \frac{|\tilde{\phi}_{yu}e^{(j\omega_n)}|^2}{\tilde{\phi}_u e^{(j\omega_n)}}}$$

How much energy is accounted for by the model? How much by noise?

$$\phi_v = \phi_y(1 - \frac{|\phi_{yu}|}{\phi_y\phi_u}$$

$$\boxed{\hat{\kappa}_{yu}e^{(j\omega_n)} = \sqrt{\frac{|\hat{\phi}_{yu}e^{(j\omega_n)}|^2}{\hat{\phi}_y e^{(j\omega_n)}\hat{\phi}_u e^{(j\omega_n)}}}} \quad \text{Coherency Spectrum}$$

- If all of the energy in the output is due to the model for a frequency ω_n then $\hat{\kappa}_{yu}e^{(j\omega_n)} = 1$.
- This can be used as a measure of effectiveness of the modelling at a particular frequency.
- Theoretically, $0 \leq \hat{\kappa}_{yu}e^{(j\omega_n)} \leq 1$. One should aim to keep the coherency spectrum as high as possible. It can be adjusted by adjusting the smoothing.

9.2 Time-domain data windowing

Putting a time domain window directly on the data.

$$U_w e^{(j\omega_n)} = \sum_{k=0}^{N-1} w_{data}(k)u(k)e^{-jk\omega_n}$$

often with $w_{data}(k) = w_\gamma(k - N/2)$ (shifted to middle). Typically $\gamma = N/2$ such that all of the data is used.

9.2.1 Welch's Method

1. Split the data record into L overlapping segments of length N.
2. $U_l e^{(j\omega_n)} = \sum_{k=0}^{N-1} w_{data}(k)u_l(k)e^{j\omega_n k}$
3. $\tilde{\phi}_u e^{(j\omega_n)} = \frac{1}{NLE_{scl}} \sum_{l=1}^{L} \left|U_l e^{(j\omega_n)}\right|^2$

- Advantages
  - Windowing can reduce transient response effects.
  - Noise reduction from averaging and windowing.
  - Variance error can be reduced.
  - Windowing can cause energy leakage to adjacent frequencies.
  - Frequency resolution deteriorates.
  - Bias error can be increased.
  - Noise on $u_l(k)$ and u_{l+1} is not uncorrelated.
- Tips
  - Do not use `welch()`, since it does not fit the definition here.

10 Frequency Domain Subspace ID

11 Closed-Loop ID

12 Time-Domain Correlation Method

13 Prediction Error Methods

14 Parameter estimation statistics

15 Nomenclature

$y(k) = Gu(k)$	output signal	[ ]
$u(k)$	input signal	[ ]
G	plant	[ ]
$\hat{G} = \frac{y}{u}$	estimated plant	[ ]
$Y(e^{j\omega})$	output spectrum	[ ]
$U(e^{j\omega})$	input spectrum	[ ]
ZOH	zero order hold	
DAC	digital analog converter	
ADC	analog digital converter	