Non-parametric Identification

Jonas Sjöberg

April 14, 2020
Non-parametric identification

Time domain

- Transient analysis
- Correlation analysis
Non-parametric identification

Time domain
- Transient analysis
- Correlation analysis

Frequency domain
- Frequency analysis
- The empirical transfer function estimate
- Spectral analysis
Consider the system described by

\[ y(t) = G_0(q)u(t) + v(t) \]

or, equivalently,

\[ y(t) = \sum_{k=1}^{\infty} g_0(k)u(t - k) + v(t) \]
Consider the system described by

\[ y(t) = G_0(q)u(t) + v(t) \]

or, equivalently,

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

**Question:** Can we determine \( G_0(q) \) or \( \{g_0(k)\} \) without parameterizing in \( \theta \)?
Impulse response analysis

Applying the input

\[ u(t) = \begin{cases} 
M, & t = 0 \\
0, & t \neq 0 
\end{cases} \]

gives the output

\[ y(t) = M \cdot g_0(t) + v(t) \]
Impulse response analysis

Applying the input

\[ u(t) = \begin{cases} 
  M, & t = 0 \\
  0, & t \neq 0 
\end{cases} \]

gives the output

\[ y(t) = M \cdot g_0(t) + v(t) \]

which motivates the impulse response estimate

\[ \hat{g}(t) = \frac{y(t)}{M} \]
Step input: \( u(t) = \begin{cases} M, & t \geq 0 \\ 0, & t < 0 \end{cases} \)

gives the output \( y(t) = M \sum_{\tau=1}^{t} g_0(\tau) + v(t) \)

so that \( \hat{g}(t) = \frac{y(t) - y(t - 1)}{M} \)
Step input: \( u(t) = \begin{cases} M, & t \geq 0 \\ 0, & t < 0 \end{cases} \)

gives the output \( y(t) = M \sum_{\tau=1}^{t} g_0(\tau) + v(t) \)

so that \( \hat{g}(t) = \frac{y(t) - y(t-1)}{M} \)

Problems with transient analysis techniques:

- excitation
- disturbances
- nonlinearities
Assume that $u$ is quasi-stationary and $u$ and $v$ are uncorrelated. Estimate correlation $y(t)$ with $u(t)$:

$$R_{yu}(\tau) = \sum_{k=1}^{\infty} g_0(k) R_u(\tau - k), \quad R_{yu}(\tau) = \frac{1}{N} \sum_{k=-\infty}^{\infty} E[y(k)u(\tau + k)]$$

since $R_{uv}(\tau) = 0 \ \forall \ \tau$. 
Assume that $u$ is quasi-stationary and $u$ and $v$ are uncorrelated. Estimate correlation $y(t)$ with $u(t)$:

$$R_{yu}(\tau) = \sum_{k=1}^{\infty} g_0(k) R_u(\tau - k), \quad R_{yu}(\tau) = \frac{1}{N} \sum_{k=-\infty}^{\infty} E[y(k)u(\tau + k)$$

since $R_{uv}(\tau) = 0 \ \forall \ \tau$.

Transform of this gives:

$$\Phi_{yu}(\omega) = G_0(e^{i\omega}) \Phi_u(\omega)$$

where

$$\Phi_u(\omega) = \sum_{\tau=-\infty}^{\infty} R_{uu}(\tau) e^{-i\tau\omega}$$
Estimate the correlations

Truncate the sum and solve the resulting system of equations for $\hat{g}(k)$:

$$\hat{R}^N_{yu}(\tau) = \sum_{k=1}^{M} \hat{g}(k) \hat{R}^N_u(\tau - k)$$

The covariance functions must be estimated:

$$\hat{R}^N_u(\tau) = \frac{1}{N} \sum_{k=1}^{N} u(k)u(k - \tau)$$
Estimate the correlations

Truncate the sum and solve the resulting system of equations for \( \hat{g}(k) \):

\[
\hat{R}_{yu}^N (\tau) = \sum_{k=1}^{M} \hat{g}(k) \hat{R}_{u}^N (\tau - k)
\]

The covariance functions must be estimated:

\[
\hat{R}_{u}^N (\tau) = \frac{1}{N} \sum_{k=1}^{N} u(k)u(k - \tau)
\]

But less data available for the estimate for large \( \tau \). This makes the estimate less accurate for large \( \tau \) (more variance). Introduce a weightening window \( W(\tau) \), reduce variance but get some bias (towards zero)

\[
\hat{R}_{u}^N (\tau) = \frac{1}{N} \sum_{k=1}^{N} W(\tau)u(k)u(k - \tau)
\]
If \( u(\cdot) \) is white noise the estimate becomes trivial (diagonal matrix)

\[
\hat{g}(k) = \frac{\hat{R}_{yu}(k)}{\sigma_u^2}
\]
So far we have estimated $g_0(k)$ — what about estimating $G_0(e^{i\omega})$?

\[ \alpha \cos \omega t \rightarrow G \rightarrow \alpha |G| \cos(\omega t + \phi) + v(t) \]
So far we have estimated $g_0(k)$ — what about estimating $G_0(e^{i\omega})$?

\[
\begin{align*}
\alpha \cos \omega t & \quad \rightarrow \quad \alpha |G| \cos(\omega t + \phi) + v(t) \\
\end{align*}
\]

Measure amplitude and phase using correlation:
\[
(\cos(\omega t + \phi) = \cos(\omega t) \cos(\phi) - \sin(\omega t) \sin(\phi))
\]

\[
I_c(N) = \frac{1}{N} \sum_{t} y(t) \cos \omega t \approx \frac{\alpha}{2} |G_0(e^{i\omega})| \cos \phi
\]

\[
I_s(N) = \frac{1}{N} \sum_{t} y(t) \sin \omega t \approx -\frac{\alpha}{2} |G_0(e^{i\omega})| \sin \phi
\]
implying

\[ I_c(N) - i \cdot I_s(N) \approx \frac{\alpha}{2} |G_0(e^{i\omega})|(\cos \phi + i \sin \phi) = \frac{\alpha}{2} |G_0(e^{i\omega})|e^{i\phi} \]

so that (for \(\omega\))

\[ \hat{G}_N = \frac{2}{\alpha} (I_c(N) - i \cdot I_s(N)) \]
Interpretation of frequency response analysis by the *correlation method*:
Recall that

\[
Y_N(\omega) = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} y(t) e^{-i\omega t} = \sqrt{N}(I_c(N) - i \cdot I_s(N))
\]

so that

\[
\hat{G}_N = \frac{2}{\alpha} (I_c(N) - i \cdot I_s(N)) = \frac{2}{\alpha \sqrt{N}} \cdot Y_N(\omega) = \frac{Y_N(\omega)}{U_N(\omega)}
\]

where the last equality holds for \( \omega = \frac{2\pi}{N} k \). Drawback, time consuming one frequency at the time.
Interpretation of frequency response analysis by the correlation method:

Recall that

\[
Y_N(\omega) = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} y(t)e^{-i\omega t} = \sqrt{N} (I_c(N) - i \cdot I_s(N))
\]

so that

\[
\hat{G}_N = \frac{2}{\alpha} (I_c(N) - i \cdot I_s(N)) = \frac{2}{\alpha \sqrt{N}} \cdot Y_N(\omega) = \frac{Y_N(\omega)}{U_N(\omega)}
\]

where the last equality holds for \( \omega = \frac{2\pi}{N} k \). Drawback, time consuming one frequency at the time.

Hence, the estimate obtained using frequency analysis by the correlation method is simply the ratio between the DFT of the output and the input. This can be generalized to general input signals . . .
Consider the open-loop system

\[ y(t) = G_o(q)u(t) + v(t) \]

\[ Y_N(\omega) = G_o(e^{i\omega})U_N(\omega) + V_N(\omega) + R_N(\omega) \]

where \( R_N(\omega) \propto 1/\sqrt{N} \) (or \( = 0 \) if \( u \) is periodic).
Consider the open-loop system

\[ y(t) = G_o(q)u(t) + v(t) \]
\[ Y_N(\omega) = G_o(e^{i\omega})U_N(\omega) + V_N(\omega) + R_N(\omega) \]

where \( R_N(\omega) \propto 1/\sqrt{N} \) (or = 0 if \( u \) is periodic).

The empirical transfer function estimate (ETF) is:

\[ \hat{G}(e^{i\omega}) = \frac{Y_N(\omega)}{U_N(\omega)}, \quad \omega = k \frac{2\pi}{N} \]
Properties:

\[
E \hat{G}_N(e^{i\omega}) = G_0(e^{i\omega}) + \frac{R_N(\omega)}{U_N(\omega)}
\]

\[
\text{Var}\hat{G}_N(e^{i\omega}) \approx \frac{\Phi_u(\omega)}{|U_N(\omega)|^2}
\]

When \( N \rightarrow \infty \):

\begin{align*}
\text{u periodic} & \quad \text{u stochastic process} \\
\text{number of } \omega \text{ fixed} & \quad \text{number of } \omega \text{ increases} \\
\text{unbiased} & \quad \text{asymptotically unbiased} \\
\text{variance } \propto 1/N & \quad \text{variance } \propto (\text{SNR})^{-1} \text{ as. uncorrelated}
\end{align*}
One can do better than accepting bias which only disappear asymptotically, in all practical cases, \( N \) is finite. The bias is due to initial condition, that it is not equal to the condition at the end, due to that finite time-window is equivalent to assuming one period of a periodic function.

\[
\hat{E}G_N(e^{i\omega}) = G_0(e^{i\omega}) + \frac{R_N(\omega)}{U_N(\omega)}
\]

The effect of \( R_N(\omega) \) can be removed by estimating the transient due to irrenous initial conditions. See [Tomas McKelvey and Guillaume Guérin, *Non-parametric frequency response estimation using a local rational model*, IFAC Symposium on System Identification, 2012].
The ETFE is a continuous function in $\omega$ and based on $N$ real measurements. It doesn’t make sense to evaluate $\hat{G}(e^{i\omega})$ for $\omega$ giving more than $N$ real values. Typically $\omega = \frac{2\pi k}{T N}$ for $k = 0, 1, \ldots N/2 - 1$. 
Spectral analysis - in figures

The ETFE is a continuous function in $\omega$ and based on $N$ real measurements. It doesn’t make sense to evaluate $\hat{G}(e^{i\omega})$ for $\omega$ giving more than $N$ real values. Typically $\omega = \frac{2\pi k}{TN}$ for $k = 0, 1, \ldots N/2 - 1$.

For larger $N$: denser values of $\hat{G}(e^{i\omega})$ on the interval $\omega = \{0, \frac{\pi}{T}\}$, but equally noisy $N$:

ETFIE for increasing $N$

A smoothing window necessary to give an estimate $\hat{G}(e^{i\omega})$. 

Close up:

Smoothing window, more data, narrower window, still more data inside the window.
The smoothing window should have the properties, when $N \to \infty$:

1. It becomes narrower, to allow correct estimation of peaks.
2. It contains more frequency data in its support to decrease the variance of the estimate.
The smoothing window should have the properties, when $N \to \infty$:

1. It becomes narrower, to allow correct estimation of peaks.
2. It contains more frequency data in its support to decrease the variance of the estimate.

ETFE for increasing $N$  

True trsf, ETFE, and smoothed estimate, which miss the peak
Each period contain the same frequencies, ie, with more periods, better estimates for these particular frequencies are obtained. No information inbetween.

ETFET based on periodic input

Close up.
The ETFE is comprised of *uncorrelated* estimates at different $\omega$ — but we know that the frequency response is smooth in reality! Idea:
The ETFE is comprised of *uncorrelated* estimates at different $\omega$ — but we know that the frequency response is smooth in reality! Idea:

(1) Compute $\hat{G}(e^{i\omega})$ as a weighted sum of the ETFE $\hat{G}$ around $\omega = \omega_0$:

$$
\hat{G}(e^{i\omega_0}) = \frac{\sum_{k=k_1}^{k_2} \alpha_k \hat{G}(e^{i\omega_k})}{\sum_{k=k_1}^{k_2} \alpha_k} \tag{1}
$$

where $2\pi k_1/N = \omega_0 - \Delta$ and $2\pi k_2/N = \omega_0 + \Delta$
The ETFE is comprised of *uncorrelated* estimates at different $\omega$ — but we know that the frequency response is smooth in reality! Idea:

(1) Compute $\hat{G}(e^{i\omega})$ as a weighted sum of the ETFE $\hat{G}$ around $\omega = \omega_0$:

$$
\hat{G}(e^{i\omega_0}) = \frac{\sum_{k=k_1}^{k_2} \alpha_k \hat{G}(e^{i\omega_k})}{\sum_{k=k_1}^{k_2} \alpha_k}
$$

where $2\pi k_1/N = \omega_0 - \Delta$ and $2\pi k_2/N = \omega_0 + \Delta$

(2) Now choose $\{\alpha_k\}$ inversely proportional to the variance of $\hat{G}(e^{i\omega_k})$:

$$
\hat{G}(e^{i\omega_0}) \approx \frac{\sum \frac{|U_N(\omega_k)|^2}{\Phi_v(\omega_k)} \cdot \hat{G}(e^{i\omega_k})}{\sum \frac{|U_N(\omega_k)|^2}{\Phi_v(\omega_k)}} \approx \frac{\int_{\omega_0 - \Delta}^{\omega_0 + \Delta} \frac{|U_N(\omega)|^2}{\Phi_v(\omega)} \cdot \hat{G}(e^{i\omega}) d\omega}{\int_{\omega_0 - \Delta}^{\omega_0 + \Delta} \frac{|U_N(\omega)|^2}{\Phi_v(\omega)} d\omega}
$$
(3) Put weight according to distance from $\omega_0$: $\hat{G}(e^{i\omega_0}) =$

$$\frac{\int_{-\pi}^{\pi} W_\gamma(\omega - \omega_0) \frac{|U_N(\omega)|^2}{\Phi_v(\omega)} \cdot \hat{G}(e^{i\omega}) d\omega}{\int_{-\pi}^{\pi} W_\gamma(\omega - \omega_0) \frac{|U_N(\omega)|^2}{\Phi_v(\omega)} d\omega} \approx \frac{\int_{-\pi}^{\pi} W_\gamma(\omega - \omega_0) |U_N(\omega)|^2 \hat{G}(e^{i\omega}) d\omega}{\int_{-\pi}^{\pi} W_\gamma(\omega - \omega_0) |U_N(\omega)|^2 d\omega}
= \frac{\int_{-\pi}^{\pi} W_\gamma(\omega - \omega_0) Y_N(\omega) \overline{U_N(\omega)} d\omega}{\int_{-\pi}^{\pi} W_\gamma(\omega - \omega_0) |U_N(\omega)|^2 d\omega} = \frac{\hat{\Phi}_{yu}(\omega_0)}{\hat{\Phi}_u(\omega_0)}$$

(2)
(3) Put weight according to distance from $\omega_0$: $\hat{G}(e^{i\omega_0}) =$

$$\frac{\int_{-\pi}^{\pi} W_\gamma(\omega - \omega_0) \frac{|U_N(\omega)|^2}{\Phi_u(\omega)} \cdot \hat{G}(e^{i\omega}) d\omega}{\int_{-\pi}^{\pi} W_\gamma(\omega - \omega_0) \frac{|U_N(\omega)|^2}{\Phi_u(\omega)} d\omega} \approx \frac{\int_{-\pi}^{\pi} W_\gamma(\omega - \omega_0) |U_N(\omega)|^2 \hat{G}(e^{i\omega}) d\omega}{\int_{-\pi}^{\pi} W_\gamma(\omega - \omega_0) |U_N(\omega)|^2 d\omega} = \frac{\hat{\Phi}_{yu}(\omega_0)}{\hat{\Phi}_u(\omega_0)}$$

(2)

This is the estimate obtained by spectral analysis according to Blackman-Tukey. The estimate is obtained by a natural replacement of the spectra in

$$\Phi_{yu}(\omega) = G_0(e^{i\omega})\Phi_u(\omega)$$

(3)

by the corresponding smoothed periodograms. This is the freq. equivalence of compensating for uncertainty of $\hat{R}_u^N(\tau)$ for large $\tau$. 
Consider the spectral estimate

\[ \hat{\Phi}_u(\omega) = \int_{-\pi}^{\pi} W_\gamma(\omega_0 - \omega)|U_N(\omega_0)|^2 d\omega_0 \]

where the frequency window, \( W_\gamma \), is narrow for large \( \gamma \). Take inv. Fourier trsf:

\[ \hat{R}_u(\tau) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{\Phi}_u(\omega)e^{i\omega \tau} d\omega = \frac{1}{2\pi} \int_{-\pi}^{\pi} W_\gamma(\omega)e^{i\omega \tau} d\omega \frac{1}{2\pi} \int_{-\pi}^{\pi} |U_N(\omega_0)|^2 e^{i\omega \tau} d\omega = \]

\[ = w_\gamma(\tau) \frac{1}{N} \sum_{t=1}^{N} u(t)u(t - \tau) \]
That is, the convolution above is transformed into multiplication in the time domain, so that

\[ \hat{\Phi}_u(\omega) = \sum_{\tau=-\infty}^{\infty} w_\gamma(\tau) \hat{R}_u(\tau) e^{-i\omega \tau} \]

where

\[ w_\gamma(\tau) = \frac{1}{2\pi} \int_{-\pi}^{\pi} W_\gamma(\omega) e^{i\omega \tau} d\omega \]

\[ \frac{1}{2\pi} \int_{-\pi}^{\pi} |U_N(\omega)|^2 e^{i\omega \tau} d\omega = \frac{1}{N} \sum_{N} u(t)u(t - \tau) \]

Note: The time (lag) window \( w_\gamma(\tau) \) is wide when the frequency window \( W_\gamma(\omega) \) is narrow, and vice versa.
When $\gamma$ increases (narrow frequency window), the bias decreases but the variance increases. This can be expressed in terms of the functions

$$M(\gamma) = \int_{-\pi}^{\pi} \omega^2 W_\gamma(\omega) d\omega \to 0$$  \hspace{1cm} (4)$$

$$\bar{W}(\gamma) = 2\pi \int_{-\pi}^{\pi} W_\gamma^2(\omega) d\omega \to \infty$$  \hspace{1cm} (5)$$

when $\gamma \to \infty$. Also, $\bar{W}(\gamma) \propto \gamma$. 
Theorem: Assume that

\[ y(t) = G_0(q)u(t) + v(t) \]  

where \( v(\cdot) \) is a stochastic process, independent of \( u(\cdot) \), which is quasi-stationary.

Then the following asymptotic results hold when \( \gamma \rightarrow \infty, \ N \rightarrow \infty, \ \gamma/N \rightarrow 0 \):

- \( E \hat{G}_N(e^{i\omega}) - G_0(e^{i\omega}) \approx M(\gamma) \left[ \frac{1}{2} G_0'' + G_0' \cdot \frac{\Phi_u'}{\Phi_u} \right] \)
- \( E|\hat{G}_N - E\hat{G}_N|^2 \approx \frac{1}{N} \bar{W}_\gamma \cdot \frac{\Phi_v}{\Phi_u} \)
- \( \text{Re}\hat{G}_N, \ \text{Im}\hat{G}_N \) are as. uncorrelated
- Estimates at different \( \omega \) are as. uncorrelated