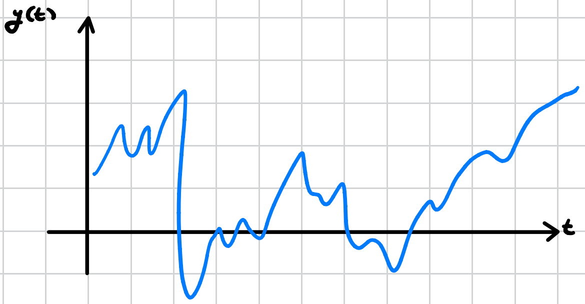
24103 - Nonparametric Identification

In Time series Domain have worked with stationary stochastic processes. Looking only 1 realization you cannot infer the other. We would like, starting from a portion of a single realization of SSP, we may want To infer some of its statistical properties

mean (m_y) covariance ($\gamma_y(\tau)$) spectrum ($\Gamma_y(\omega)$)



So far we have followed these steps To find them: Collect measurement, create a data set, identify the model $W(z)$, and compute the desired measure $m_y = E[y(t)] = W(1)E[e(t)]$ This a model based approach inference (not direct).

Sample based estimation

Sample Mean: we want To find a direct estimation of the expected value m_y of $y(t)$

$$D_N = \{y(1), \dots y(N)\} \to m_y = E[y(t)]$$

The most natural estimator is the mean over Time.

$$\hat{m}_N = 1/N \sum_{1}^{N} y(t)$$

Is This a "good" estimator for the real m_y? We need To define what "good" means:

> **DEF Property of correctness:** an estimator is correct if The expected value of The estimator is equal To The probabilistic property To be estimated

For The mean $E[\hat{m}_N] = m_y$. Is $\hat{m}_N$ correct? Yes!

$$E[\hat{m}_N] = E[\,1/N \sum y(t)\,] = 1/N \sum E[y(t)] = 1/N \sum^{N} m_y = \cancel{N}\tfrac{1}{\cancel{N}} m_y = m_y$$

Is correctness enough?

> **DEF: Consistency:** an estimator is consistent if The variance of The estimation error Tends To zero, as The number of data Tends To ∞

For The mean: $E[(\hat{m}_N - m_y)^2] \to 0$ for $N \to +\infty$. Is $\hat{m}_N$ consistent? For example, consider $y(t,s) = v(s)$ $\forall t$ with $v(s) \sim N(0,1)$

$$E[(\hat{m}_N - m_y)^2] = E[(1/N \sum y(t,s) - m_y)^2] = E[(1/N \sum v(s))^2]$$
$$= E[(N \cdot 1/N\, v(s))^2] = E[v^2(s)] = 1 \quad \forall N$$

The variance is constant To 1 for any N (more sample does not add any information). So This is not consistent!

Theorem: if $\gamma(\tau) \to 0$ as $|\tau| \to +\infty$, Then $\hat{m}_N$ is consistent

The condition $\gamma(\tau) \to 0$ for $|\tau| \to +\infty$ is True for all ARMA/ARMAX processes (for The ARMAX one, The exogenous part is always deterministic and fully known).

Covariance Function: we want To find a direct estimation

$$D_N = \{y(1), \dots, y(N)\} \to \gamma_y(\tau) = E[y(t)y(t-\tau)] \quad \forall \tau$$

For semplicity we assume $m_y = 0$ without loss of generality. One possible estimator is

$$\hat{\gamma}_N(\tau) = 1/N-\tau \sum_{t}^{N-\tau} y(t)y(t+\tau) \quad \left(\text{or} \quad \hat{\gamma}_N(\tau) = 1/N \sum_{1+\tau}^{N} y(t)y(t-\tau)\right)$$

Notice That:

- we cannot use any τ: $0 \le \tau \le N-1$ (Theoretically True, but we cannot compute close To These bounds)
- $\hat{\gamma}_N(0) = [y^2(1) + \dots + y^2(N)]\, 1/N$ average over N elements
- $\hat{\gamma}_N(1) = [y(1)y(2) + \dots + y(N-1)y(N)]\, 1/N-1$ average over N-1 elements
- $\hat{\gamma}_N(N-1) = [y(1)y(t+\tau=N)] \cdot 1/1$ average computed over one single pair of samples
- From a practical point of view, we can compute in a reliable way only for $\tau \ll N-1$

What about negative τ? We recall that $\gamma_y(\tau) = \gamma_y(-\tau)$, so we consider $|\tau|$ instead of simply τ; so we can write

$$\hat{\gamma}_N(\tau) = \frac{1}{N-|\tau|} \sum_{t=1}^{N-|\tau|} y(t)\, y(t+|\tau|)$$

- Is $\hat{\gamma}_N(\tau)$ correct? By applying the formula:

$$E[\hat{\gamma}_N(\tau)] = E\left[\frac{1}{N-\tau} \sum y(t) y(t+\tau)\right] = \frac{1}{N-\tau} \sum E[y(t) y(t+\tau)]$$
$$= \frac{1}{\cancel{N-\tau}} \cancel{(N-\tau)}\, \gamma_y(\tau) = \gamma_y(\tau) \quad \checkmark$$

- Is $\hat{\gamma}_N(\tau)$ consistent? We have to introduce a new theorem

Theorem: $\hat{\gamma}_N(\tau)$ is consistent if $\gamma(\tau) \to 0$ for $\tau \to +\infty$

Spectral Density: we want to find a direct estimation

$$D_N = \{y(1), \dots, y(N)\} \Rightarrow \Gamma_y(\omega) = \sum_{\tau=-\infty}^{+\infty} \gamma(\tau) e^{-j\omega\tau} \quad \forall \omega$$

One possible estimator is

$$\hat{\Gamma}_N(\omega) = \sum_{\tau=-(N-1)}^{N-1} \hat{\gamma}_N(\tau) e^{-j\omega\tau}$$

we have a problem due to the approximation:

- sum boundaries: $(N-1)$ and $-(N-1)$
- $\hat{\gamma}_N(\tau)$ from previous approximation

This is again an indirect non parametric estimator as we need to estimate $\hat{\gamma}_N(\tau)$ first

Is $\hat{\Gamma}_N(\omega)$ correct?

$$E[\hat{\Gamma}_N(\omega)] = E\left[\sum \hat{\gamma}_N(\tau) e^{-j\omega\tau}\right]$$
$$= \sum \boxed{E[\hat{\gamma}_N(\tau)]} e^{-j\omega\tau} \quad \text{is correct so this coincide with } \gamma_y(\tau)$$
$$= \sum_{-(N-1)}^{N-1} \gamma_y(\tau) e^{-j\omega\tau} \neq \Gamma_y(\omega)$$

This is not correct (since $-(N-1) \neq -\infty$ and $N-1 \neq +\infty$) but it is asymptotically correct, so

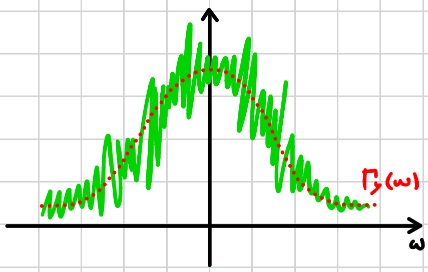
$$\lim_{N\to+\infty} E[\hat{\Gamma}_N(\omega)] = \Gamma_y(\omega) \quad \forall \omega$$

Is $\hat{\Gamma}_N(\omega)$ consistent? $E[(\hat{\Gamma}_N(\omega) - \Gamma_y(\omega))^2] \to \Gamma_y(\omega)$ for $N \to +\infty$

The error is large with large quantity

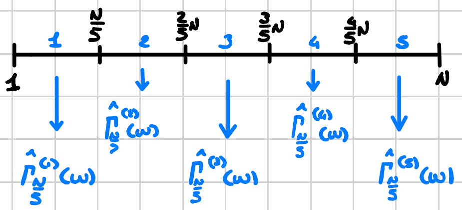
Moreover: $E[(\hat{\Gamma}_N(\omega_1) - \Gamma_y(\omega_1))(\hat{\Gamma}_N(\omega_2) - \Gamma_y(\omega_2))] \to 0$

no correlation between errors at ω_1 and ω_2.



There are several way to deal with that. One is AVERAGING (or "regularization"):

Take the dataset and divide it into fixed portion and compute the estimation of only one portion



If N is large also $\frac{N}{5}$ is good to be used for estimation

The average will compensate the fluctuation of each predictor

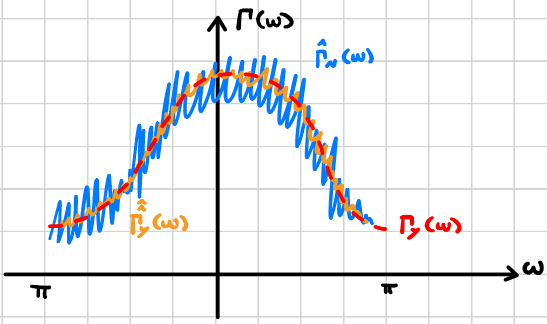
So:

$$\hat{\hat{\Gamma}}_N(\omega) = \frac{1}{5} \sum_{i=1}^{5} \hat{\Gamma}^{(i)}_{\frac{N}{4}}(\omega)$$

We could prove that $\frac{N}{5}$ is "large" then

$$E[(\hat{\hat{\Gamma}}_N(\omega) - \Gamma_y(\omega))^2] \approx E[(\hat{\Gamma}_N(\omega) - \Gamma_y(\omega))^2] \frac{1}{5}$$

we are introducing a really small bias (can be ignored) with N very large



The number of parts cannot be too large, otherwise the estimate would be biased (bias/variance trade off not convinient anymore!)

The sampled Spectrum is Tipically computationally more demanding than the other sample estimators. Let's Take a different estimator of $\gamma_y(\tau)$

$$\hat{\gamma}'_N(\tau) = \frac{1}{N}\sum_{t=1}^{N-\tau} y(t)\, y(t+\tau)$$ This is only asymptotically correct

IT can be proven That $\hat{\Gamma}'_N(\tau) = \sum \hat{\gamma}'_N(\tau) e^{-j\omega\tau}$ is equivalent To $\hat{\Gamma}'_N(\tau) = \frac{1}{N} \left| \sum_{t=1}^{N} y(t) e^{-j\omega t} \right|^2$

So you can see That $\hat{\Gamma}'_N$ can be computed directly from $y(t)$ and This is The Discrete Fourier Transformation DFT of $y(t)$ (we can compute This easily with a lot of Tool)