Lecture 14

Monday, February 28, 2022 9:23 AM

Modeling of Linear Dynamical Systems Part II

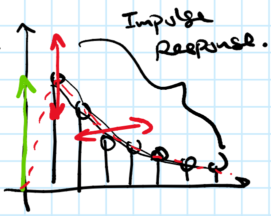
<u>Admin</u>

1. Graded HW 1 → references. (format)
2. Class cancelled weds.
3. OH: 9am weds, Friday. — Stacy tired. Tues 9am ?
4. HW 02 due this week.
5. project signups soon.

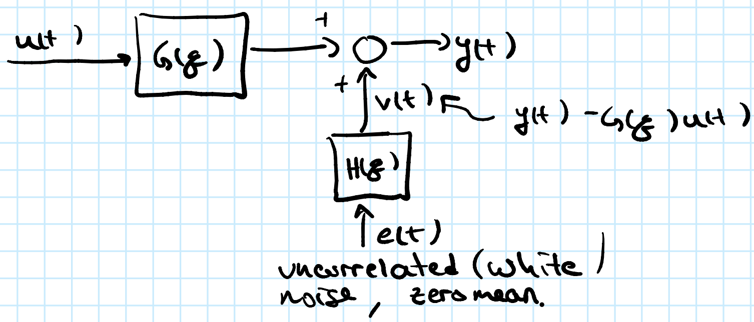
<u>last time</u>

$u(t) \rightarrow \boxed{G(q)} \rightarrow y(t)$

- FIR ← considers the history of inputs
- ARX ← considers the time history of inputs & outputs.
- ARMAX ← ARX plus parameters to account for noise.



<u>Parameter Estimation for Prediction</u>



$v(t) = H(q)\, e(t)$

$\Rightarrow e(t) = H^{-1}(q)\, v(t)$

$= H^{-1}(q)\left[y(t) - G(q)u(t)\right]$

$y(t) = G(q)\underset{\sim}{u(t)} + H(q)e(t)$

$\Rightarrow$ can we predict the future, $\hat{y}(t|t-1)$?

- can show that we can predict the most probable next output : (maximum likelihood)

$$\hat{y}(t|t-1) = \underline{G(q)\,u(t)} + \hat{v}(t|t-1)$$

$$+\underbrace{\left[H(q) - 1\right]}_{\text{history dynamics}} e(t)$$

$1 + h_1 q^{-1} + \cdots$

$v(t) = H(q)e(t)$

$v(t-1) = H(q)\, e(t-1)$

$= [1 + h_1 q^{-1} + \cdots]\, e(t-1)$

$= [1 + h_1 q^{-1} + \cdots]\, e(t) - e(t)$

$$\hat{y}(t|t-1) = G(q)u(t) + \left[1 - H^{-1}(q)\right]\left[y(t) - G(q)u(t)\right]$$

Algebra.

$$\boxed{\hat{y}(t|t-1) = H^{-1}(q)\, G(q)u(t) + \left[1 - H^{-1}(q)\right] y(t)}$$

can show prediction error is

$\underset{\sim}{e(t)} = y(t) - \hat{y}(t|t-1)$

$$\hat{y}(t|t-1) = H^{-1}(q)G(q)u(t) + \underbrace{[1-H^{-1}(q)]y(t)}$$

predictor

predict $y(t)$ given all info: $\{u(t), u(t-1), \ldots, y(t-1), y(t-2), \ldots\}$

time history of $y(t-1)$, $y(t-2)$, ...

$e(t) = y(t) - \hat{y}(t|t-1)$

Part of $y(t)$ that cannot be predicted.

$1 - \{1 + h_1 q^{-1} + h_2 q^{-2}\}$

$\Rightarrow [h_1 q^{-1} + h_2 q^{-2}]y(t)$

$\Rightarrow h_1 y(t-1) + h_2 y(t-2)$.

Pseudo linear Regression of ARMAX coeffs for prediction.

- ARMAX: $A(q)y(t) = B(q)u(t) + C(q)e(t)$

$$\Rightarrow G(q,\overset{\text{parameters}}{\theta}) = \frac{B(q)}{A(q)} \; ; \quad H(q,\theta) = \frac{C(q)}{A(q)}$$

- The one-step ahead predictor for ARMAX is:

$$\hat{y}(t,\theta) = \underbrace{G(q)H^{-1}(q)}_{\frac{B(q) - [1+b_1 q^{-1}+\cdots]}{C(q) - [1+c_1 q^{-1}+\cdots]}}u(t) + [1-\underbrace{H^{-1}(q)}_{\frac{A(q)}{C(q)}}]y(t)$$

- This cannot be written as a linear regression

  - multiply through by $C(q)$, add $[1-C(q)]\hat{y}(t,\theta)$

Algebra

$$\Rightarrow \underset{\text{parameters}}{\hat{y}(t,\theta)} = \underset{1+b_1 q^{-1}+b_2 q^{-2}+\cdots}{B(q)}u(t) + [1-\underbrace{A(q)}_{1+a_1 q^{-1}+a_2 q^{-2}+\cdots}]y(t) + [\underbrace{C(q)-1}_{1+c_1 q^{-1}+\cdots}]\underbrace{\{y(t)-\hat{y}(t,\theta)\}}_{\underbrace{\varepsilon(t|\theta)}_{\text{prediction error given params, }\theta.}}$$

$\rightarrow$ Then, $\varphi(t,\theta) = [u(t-1), \ldots, u(t-n_b), y(t-1), \ldots y(t-n_a), \underbrace{\varepsilon(t-1|\theta), \ldots \varepsilon(t-n_c|\theta)}]^T$

$\Rightarrow \hat{y}(t|t-1,\theta) = \varphi^T(t|\theta)\theta$

where

$$\theta = [b_1, \ldots b_{n_b}, a_1, \ldots, a_{n_a}, c_1, \ldots c_{n_c}]$$

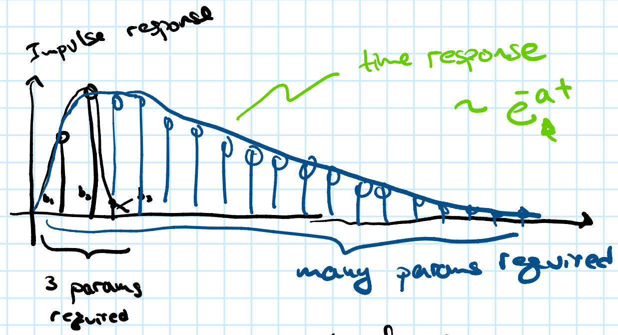
- Since $\varepsilon(t|\theta)$ includes θ, φ depends on θ, so not strictly a linear function of θ

  $\Rightarrow$ A pseudo linear regression.

Time series "data compression"

Finite impulse response model.

$\Rightarrow$



$$\text{FIR: } y(t) = b_1 u(t-1) + b_2 u(t-2) + \cdots + \underbrace{b_m}_{\#\text{ of params}} u(t-m).$$

$$y(t) = \varphi^T(t)\underbrace{\theta}_{[b_1 \cdots b_m]^T}$$

pros & cons of FIR modeling

pros

- least squared error gives a consistent estimate

$$\lim_{N\to\infty} \hat{\theta}_N^{LS} = \underbrace{\theta_0}_{\text{true values of parameters}}$$

as long as the input sequence $\{u(t)\}$ is uncorrelated w/ the noise $e(t)$.

cons

- Two failure scenarios:
  1. The impulse response has a slow decaying mode
  2. The sampling rate is high.

$\Rightarrow$ In either case, the number of parameters to estimate, m, becomes large.

$$\Rightarrow \Phi\Phi^T = \sum_{t=1}^{N} \varphi(t)\varphi^T(t) \text{ is not full rank.}$$

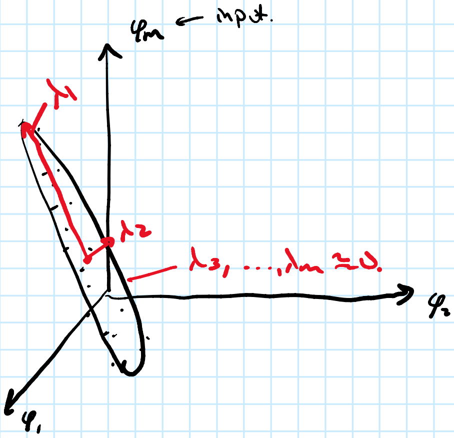
Then the eigenvalues become:

$$\lambda_1 \geq \lambda_2 \geq \underbrace{\cdots\cdots, \lambda_m}_{\simeq 0}$$

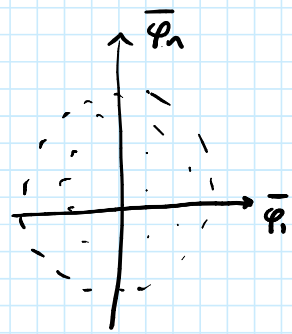
$\Rightarrow$ Since $\varphi(t)$ depends on $u(t)$, difficult to have a sufficiently complex input sequence such that all of the eigenvalues are nonzero.

$\Rightarrow$ "The input is <u>not</u> persistently exciting" $\leftarrow$

let's consider a coord. transformation:



Transform the data
⇒
to a lower dimensional space



$R^m \to R^n \; (n < m)$

$(n << m)$