Lecture 21

Monday, April 11, 2022 8:22 AM

Admin

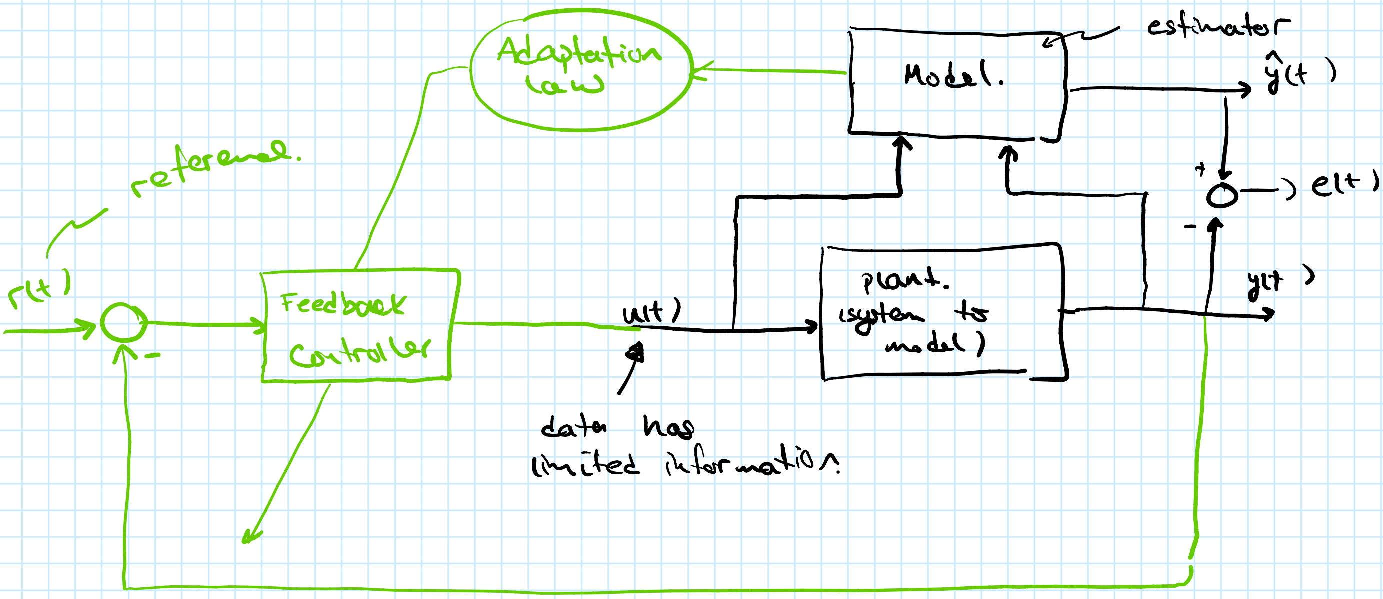
1. HW set 4 → due Friday.
2. HW set 5 → due 04/22
3. HW set 6 → free.
4. HW set 7 → survey.

Informative Datasets, Persistence of Excitation, and Asymptotic Distribution of Parameter Estimates

<u>Informative datasets & Persistence of Excitation</u>

- Informative dataset are closely related to "persistence of excitation"

$\underbrace{\text{"persistence of excitation"}}_{\text{auto-covariance matrix is invertible}}$



- The central question is if the input-output data is informative enough to identify the model.

  → often questionable since controller is trying to set the output to a reference.

<u>Definition</u>: "Persistence of excitation"

A quasi-stationary signal $\{u(t)\}$, w/ spectrum $\Phi_u(\omega)$, is said to be persistently exciting of order n if the condition

$$|M_n(e^{i\omega})|^2 \, \Phi_u(\omega) \equiv 0$$

implies that $M_n(e^{i\omega}) \equiv 0$

implies that $M_n(e^{\cdot}) = 0$

where $M_n(z) = m_1 z^{-1} + m_2 z^{-2} + m_3 z^{-3} + \dots + m_n z^{-n}$
(arbitrary linear filter)

Note: $|M_n(e^{i\omega})|^2 \Phi_u(\omega)$ is the power spectrum of $v(t) = M(z)u(t)$

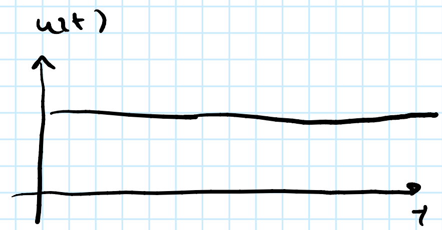
$\Rightarrow$ Therefore a signal $u(t)$ that is persistently exciting of order n cannot be filtered to zero by any $(n-1)^{st}$ order moving average filter.

---

persistence of excitation.

Can filter to zero with $n=?$

say $m_1 = 1. \Rightarrow m_2 = -1$



$$u(t) \longrightarrow \boxed{M_n(z)} \longrightarrow v(t)$$

order of model

$$v(t) = \varphi^T(t)\,\theta \;;\; \varphi(t) = [u(t-1) \;\; u(t-2) \; \cdots \; u(t-n)]^T$$

$$\theta = [m_1 \;\; m_2 \; \cdots \; m_n]^T$$

$$\underbrace{\left\{ \sum_{t=1}^{N} \varphi(t)\,\varphi^T(t) \right\}}_{} \hat{\theta} = \sum_{t=1}^{N} \varphi(t)\,y(t)$$

$$\approx E \begin{bmatrix} u(t)u(t) & u(t)u(t-1) & \cdots & u(t)u(t-n) \\ & \ddots & & \vdots \\ u(t)u(t-n) & & & u(t)u(t) \end{bmatrix}$$

Theorem

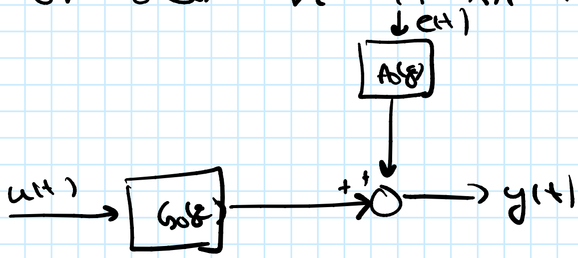
- Let $u(t)$ be a quasi-stationary signal with.

$$\bar{R}_n = \begin{bmatrix} R_u(0) & \cdots & R_u(n-1) \\ \vdots & & \vdots \\ R_u(n-1) & \cdots & \end{bmatrix} \leftarrow \mathbb{R}^{n \times n}.$$

$$\bar{R}_n = \begin{bmatrix} \vdots & & \vdots \\ R_u(n-1) & \cdots & R_u(0) \end{bmatrix} \leftarrow \mathbb{R}^{n\times n}.$$

Then, $u(t)$ is persistently exciting of order n if $\bar{R}_n$ is non-singular.

Proof: will be posted.



- Signal to noise ratio.

 - Consider true system: $y(t) = G_0(q)u(t) + H_0(q)e(t)$
 - The prediction error is:

$$\varepsilon(t,\theta) = y(t) - \hat{y}(t|t-1)$$

$$= G_0(q)u(t) + H_0(q)e(t) - \left\{ H_\theta^{-1}(q)G_\theta(q)u(t) + \left[1 - H_\theta^{-1}(q)\right]y(t) \right\}$$

(Algebra) $$= H_\theta^{-1}\left[(G_0 - G_\theta)u(t) + H_0 e(t)\right]$$

$$= H_\theta^{-1}(q)\left[(G_0 - G_\theta)u(t) + \underbrace{(H_0 - H_\theta)}_{\underbrace{1 + h_1 q^{-1} + h_2 q^{-2} + \cdots}_{\text{does not depend on } e(t)}} e(t)\right] + e(t)$$

uncorrelated

 - All three terms are uncorrelated.

$$\Phi_\varepsilon(\omega,\theta) = \frac{|G_0 - G_\theta|^2}{|H_\theta|^2}\Phi_u(\omega) + \frac{|H_0 - H_\theta|^2}{|H_\theta|^2}\lambda + \lambda$$

$$\hat{\theta} = \underset{\theta}{\arg\min} \int_{-\pi}^{\pi} \left[\qquad \right] d\omega$$

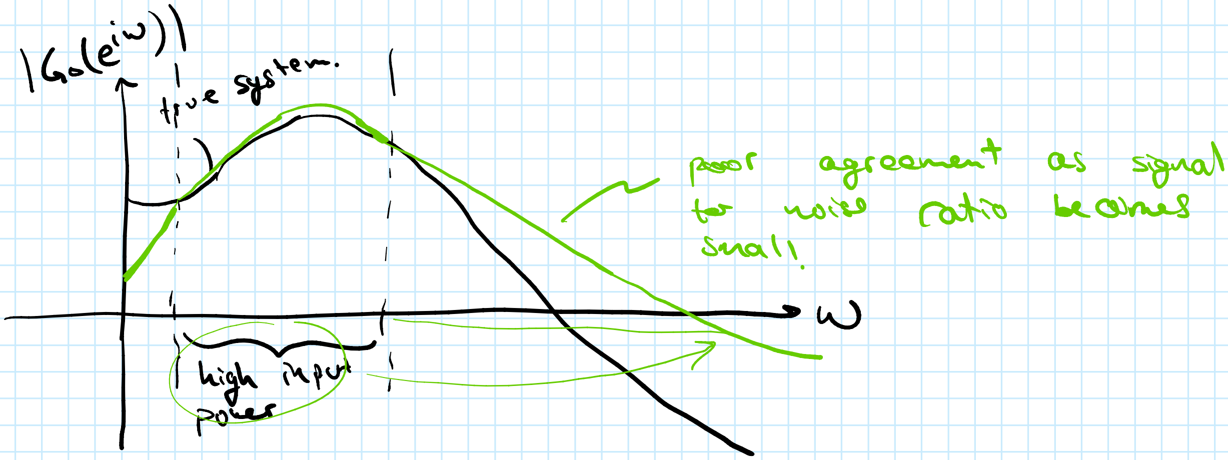
Let's simplify: $H_\theta(q) = H^*(q)$, Then:

$$\hat{\theta} = \arg\min \int^{\pi} \left| \;\; (e^{i\omega}) \;\; (e^{i\omega}) \right|^2 \Phi_u(\omega) \,.$$

$$\hat{\theta} = \underset{\theta}{\arg\min} \int_{-\pi}^{\pi} \left| G_0(e^{i\omega}) - G_\theta(e^{i\omega}, \theta) \right|^2 \frac{\Phi_u(\omega)}{|H^*(e^{i\omega})|^2} d\omega$$

- The model, G_θ, is pushed toward the true system $G_0(q)$ in a manner that is weighted by the power spectrum of the input @ each frequency.

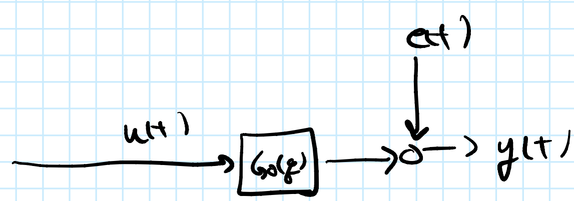
-> The weight $\dfrac{\Phi_u(\omega)}{|H^*(e^{i\omega})|^2}$ is the ratio of the input power spectrum to the noise



- Asymptotic analysis (Ljung, chap 9).

For $n > 1$, $N >> 1$.

$$\text{var}\left(\hat{G}(e^{i\omega})\right) = \frac{n}{N} \frac{\Phi_{ee}(\omega)}{\Phi_{uu}(\omega)}$$



Reduce variance

- $n \downarrow$
- $N \uparrow$
- $\Phi_{uu}(\omega) \uparrow$
- $\Phi_{ee}(\omega) \downarrow$

Asymptotic Analysis of Parameter Estimates

Overview:

- If convergence is guaranteed, $\hat{\theta}_N \to \theta^*$ ← vs. θ_0 ← true parameters.
- But how quickly does the estimate $\hat{\theta}_N \to \theta^*$?
  - → How many data points do we need?
  - → Asymptotic analysis.
- Go back to estimating the slope of a line:

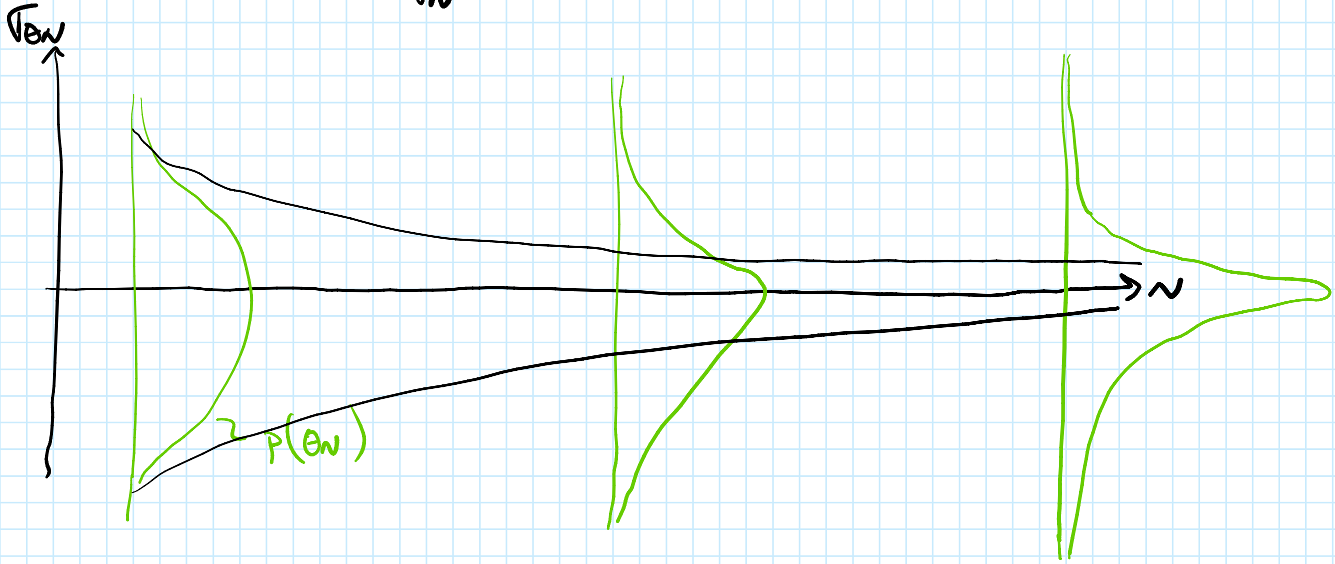
$$\hat{Y} = \beta_1 x + \beta_0 + e_0 \quad \lambda_0$$

$$\Rightarrow Var(\hat{\beta}_1) = \frac{Var(Y)}{\sum_{i=1}^{N} (x_i - \mu_x)^2} \quad \lambda_0$$

inputs → analogous to $\varphi(t)$

$$\Rightarrow Var(\beta_1) \sim \frac{1}{N} \sigma_Y^2$$

$$\Rightarrow \hat{\sigma}_{\beta_1} = \sim \frac{\sigma_Y}{\sqrt{N}}$$



Main points:

- The estimate $\hat{\theta}_N$ converges to θ^* at a rate proportional to $\frac{1}{\sqrt{N}}$
- The distribution over the parameters converges to a Gaussian $N(\theta^*, \sigma)$
- The identified parameters have $cov(\hat{\theta}_N)$, which defines a performance metric for the model!

Central Limit Theorem

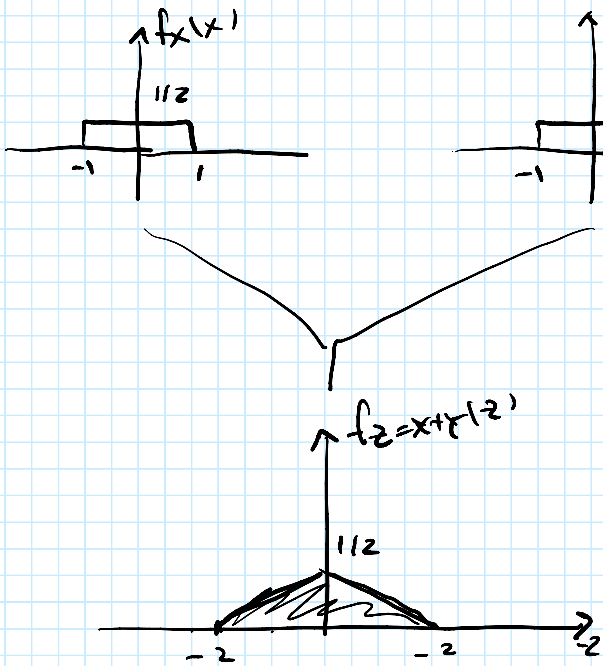
Central Limit Theorem

- Consider two indep. random variables X & Y, with PDF $f_X(x)$ & $f_Y(y)$, define

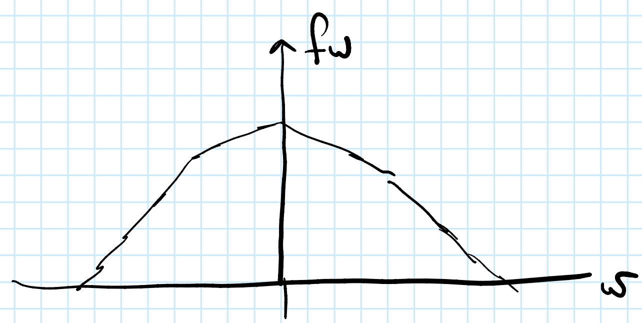
$\Rightarrow Z_1 = X + Y$

$$pr(z \le Z_1 \le z + \Delta z) = \iint f_X(x)\, f_Y(\overset{z-x}{y})\, dy$$

$\Rightarrow$



- Further consider $W = X + Y + Z \Rightarrow$



$\Rightarrow$ Resultant PDF starts to look Gaussian.

- In general

$$\sum_{i=1}^{N} X_i$$

approaches a Gaussian.