Filtering and Identification

Day 4 - Lecture 2:
Subspace Identification
Consistency Analysis

Michel Verhaegen

Division of Automatic Control - EE

Overview

- **Recap of the deterministic solution**
- Formulating the noise perturbation cases
- Consistency analysis: Additive white noise disturbance to the output (Case 0)
- Consistency other cases: Use of Instrumental Variables

The Deterministic Case

SGM: $$\begin{cases} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k) \end{cases} \quad x(k) \in \mathbb{R}^n$$

Data equation $Y_{i,s,N} = \mathcal{O}_s X_{i,N} + \mathcal{T}_s U_{i,s,N}$

Recovery column space of extended observability matrix $\mathcal{O}_s$:

The Deterministic Case

SGM: $\begin{cases} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k) \end{cases} \quad x(k) \in \mathbb{R}^n$

Data equation $Y_{i,s,N} = \mathcal{O}_s X_{i,N} + \mathcal{T}_s U_{i,s,N}$

Recovery column space of extended observability matrix $\mathcal{O}_s$:

Let, $\Pi^{\perp}_{U_{i,s,N}} = \left(I - U^T_{i,s,N} \left(U_{i,s,N} U^T_{i,s,N} \right)^{-1} U_{i,s,N} \right)$, then:

$$Y_{i,s,N} \Pi^{\perp}_{U_{i,s,N}} = \mathcal{O}_s X_{i,N} \Pi^{\perp}_{U_{i,s,N}} \Rightarrow \text{range}(Y_{i,s,N} \Pi^{\perp}_{U_{i,s,N}}) \stackrel{?}{=} \text{range}(\mathcal{O}_s)$$

The Deterministic Case

SGM: $\begin{cases} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k) \end{cases} \quad x(k) \in \mathbb{R}^n$

Data equation $Y_{i,s,N} = \mathcal{O}_s X_{i,N} + \mathcal{T}_s U_{i,s,N}$

Recovery column space of extended observability matrix $\mathcal{O}_s$:

Let, $\Pi^{\perp}_{U_{i,s,N}} = \left(I - U^T_{i,s,N} \left(U_{i,s,N} U^T_{i,s,N} \right)^{-1} U_{i,s,N} \right)$, then:

$$Y_{i,s,N}\Pi^{\perp}_{U_{i,s,N}} = \mathcal{O}_s X_{i,N} \Pi^{\perp}_{U_{i,s,N}} \Rightarrow \text{range}(Y_{i,s,N}\Pi^{\perp}_{U_{i,s,N}}) \overset{?}{=} \text{range}(\mathcal{O}_s)$$

The RQ factorization provides an efficient solution $\Rightarrow$:

$$\begin{bmatrix} U_{i,s,N} \\ Y_{i,s,N} \end{bmatrix} = \begin{bmatrix} R_{11} & 0 \\ R_{21} & R_{22} \end{bmatrix} \underbrace{\begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix}}_{Q} \quad QQ^T = I \quad \boxed{\Rightarrow Y_{i,s,N}\Pi^{\perp}_{U_{i,s,N}} = R_{22} Q_2}$$

The 5-line matlab solution:Basic MOESP

GIVEN: The i/o data sequences $\{u(k), y(k)\}_{k=1}^{N}$ and the integer s to specify number of block rows of the Hankel matrices $U_{i,s,N}, Y_{i,s,N}$ ($s > n$?).

THEN DO:

- Construct Hankel matrices $U_{i,s,N}, Y_{i,s,N}$ (`U,Y`)
- RQ factorization: `r = triu(qr([U ; Y]'))';`
- Extract `R22`
- Range (column space) calculation + order detection:

```
[uu,ss,vv]=svd(R22);
semilogy(diag(ss),'xr');
```

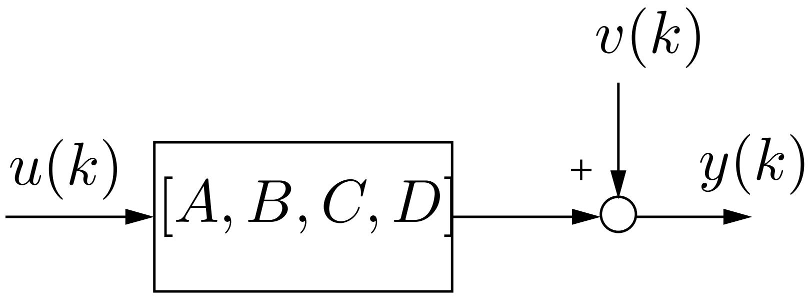
- Estimate A_T, C_T

subid_basic.m

Overview

- Recap of the deterministic solution
- **Formulating the noise perturbation cases**
- Consistency analysis: Additive white noise disturbance to the output (Case 0)
- Consistency other cases: Use of Instrumental Variables

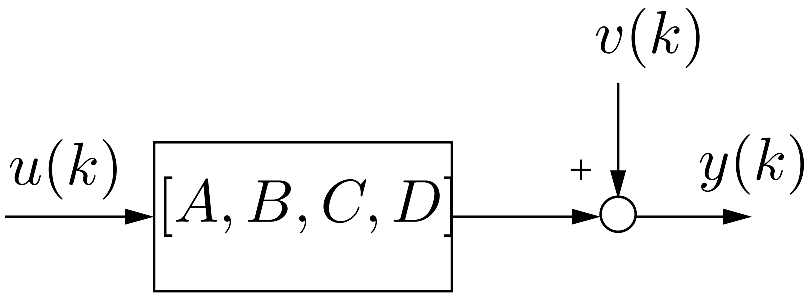
Case 0: Additive White Noise



$$
\begin{aligned}
x(k+1) &= Ax(k) + Bu(k) \\
y(k) &= Cx(k) + Du(k) + v(k)
\end{aligned}
$$

Properties about the additive disturbance $v(k)$:

Case 0: Additive White Noise

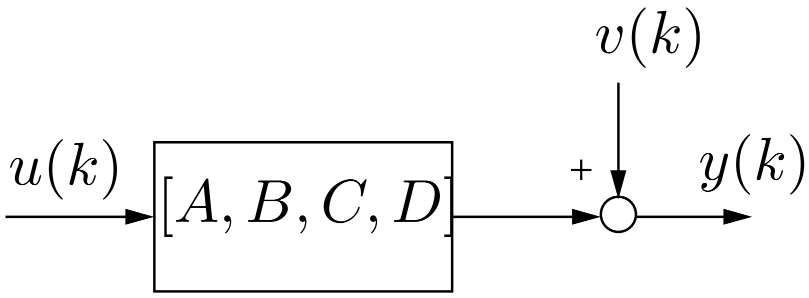


$$
\begin{aligned}
x(k+1) &= Ax(k) + Bu(k) \\
y(k) &= Cx(k) + Du(k) + v(k)
\end{aligned}
$$

Properties about the additive disturbance $v(k)$:

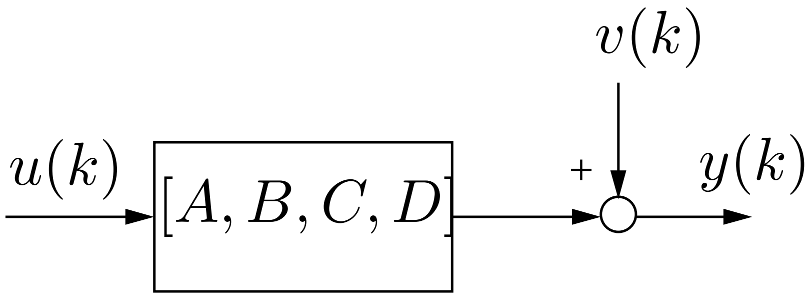
- zero-mean
- WSS stochastic process — white Spectrum
- uncorrelated with the input $u(k)$

Case 1: Output-Error Model



$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k) + v(k) \end{aligned}$$

Case 1: Output-Error Model



$$
\begin{aligned}
x(k+1) &= Ax(k) + Bu(k) \\
y(k) &= Cx(k) + Du(k) + v(k)
\end{aligned}
$$

Properties about the additive disturbance $v(k)$:

- zero-mean
- WSS stochastic process — Rational Spectrum
- uncorrelated with the input $u(k)$

Example: Output-Error case

Artificial Human Joints:

Example: Output-Error case

Artificial Human Joints:

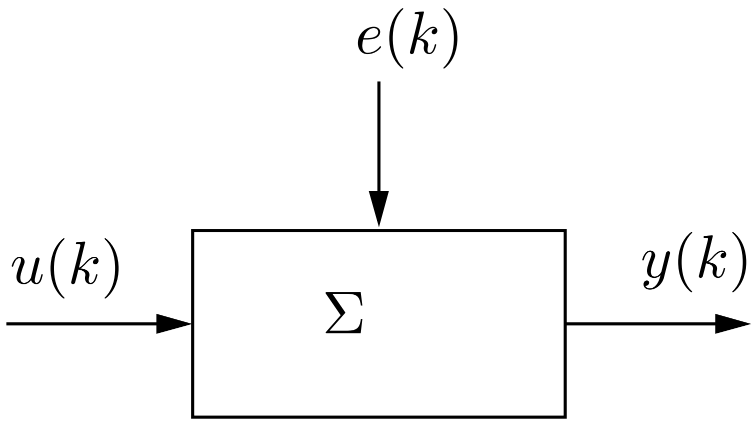
Noise $v(k)$ represents:

1. measurement noise
2. unwanted dynamical phenomena in the data not of interest in the modeling of the relationship “torque” — “angular rotation” of the joint

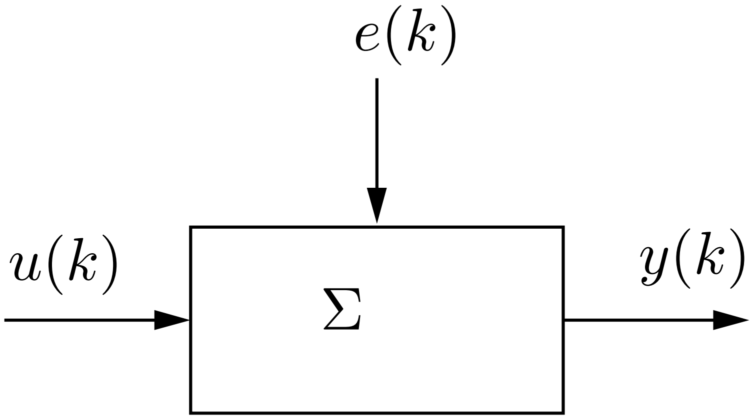
Lab Testing:

Courtesy RekLab Prof. R. Kearney - McGill

Case 2: Innovation Model



Case 2: Innovation Model

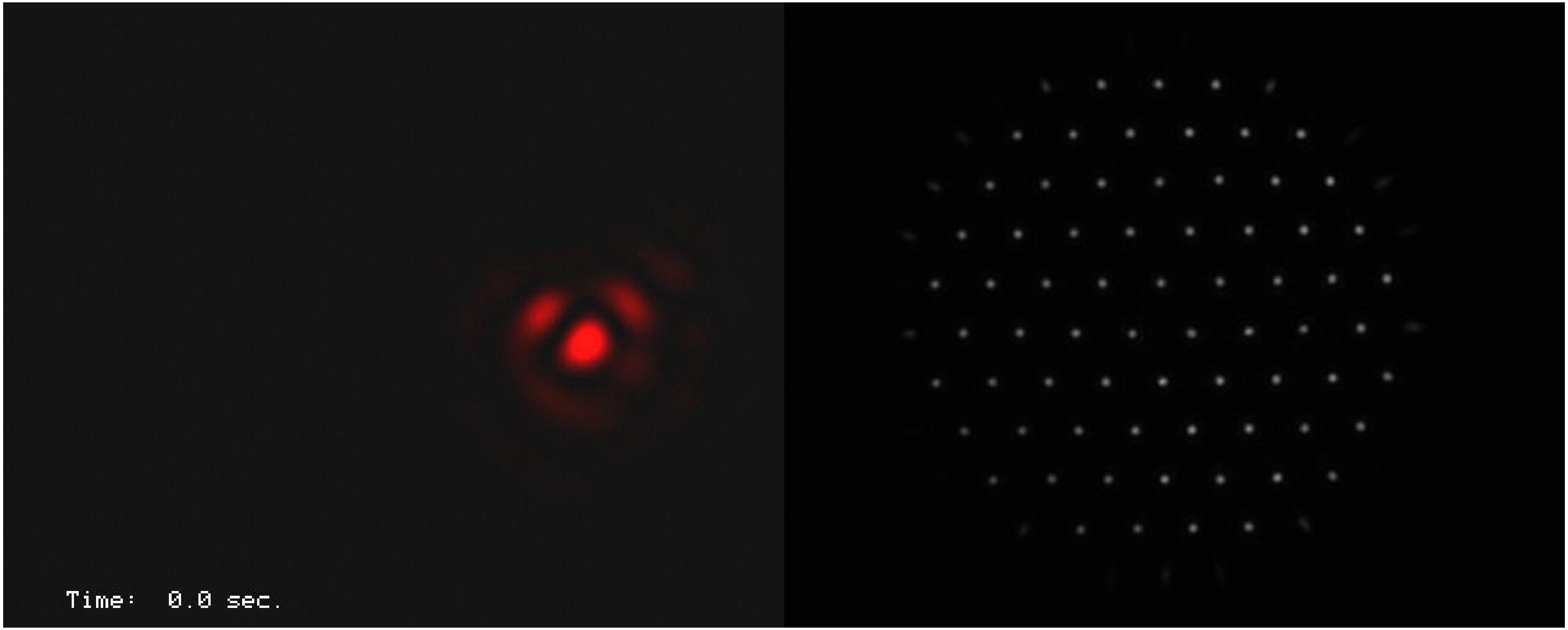


$$
\begin{aligned}
x(k+1) &= Ax(k) + Bu(k) + Ke(k) \\
y(k) &= Cx(k) + \Big(Du(k)\Big) + e(k)
\end{aligned}
$$

Properties about the additional input disturbance $e(k)$:

- zero-mean
- WSS stochastic process — white Spectrum
- uncorrelated with the input $u(k)$

Case 2: $v(k)$ is the “only” unmeasurable input exciting dynamics of interest



Overview

- Recap of the deterministic solution
- Formulating the noise perturbation cases
- **Consistency analysis: Additive white noise disturbance to the output (Case 0)**
- Consistency other cases: Use of Instrumental Variables

Additive zero-mean, white noise Problem

LTI System:

$$\begin{cases} x(k+1) & = & Ax(k) + Bu(k) \\ y(k) & = & Cx(k) + Du(k) + v(k) \end{cases}$$

Data equation

$$Y_{i,s,N} = \mathcal{O}_s X_{i,N} + \mathcal{T}_s U_{i,s,N} + V_{i,s,N}$$

Additive zero-mean, white noise Problem

LTI System:

$$\begin{cases} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k) + v(k) \end{cases}$$

Data equation

$$Y_{i,s,N} = \mathcal{O}_s X_{i,N} + \mathcal{T}_s U_{i,s,N} + V_{i,s,N}$$

Problem: Can we find $\text{span}_{\text{col}}\left(\mathcal{O}_s\right)$

Additive zero-mean, white noise Problem

LTI System:

$$\begin{cases} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k) + v(k) \end{cases}$$

Data equation

$$Y_{i,s,N} = \mathcal{O}_s X_{i,N} + \mathcal{T}_s U_{i,s,N} + V_{i,s,N}$$

Problem: Can we find $\text{span}_{\text{col}}\left(\mathcal{O}_s\right)$ *consistently*?

Additive zero-mean, white noise case

LTI System:

$$\begin{cases} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k) + v(k) \end{cases}$$

Data equation

$$Y_{i,s,N} = \mathcal{O}_s X_{i,N} + \mathcal{T}_s U_{i,s,N} + V_{i,s,N}$$

Additive zero-mean, white noise case

LTI System:

$$\begin{cases} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k) + v(k) \end{cases}$$

Data equation

$$Y_{i,s,N} = \mathcal{O}_s X_{i,N} + \mathcal{T}_s U_{i,s,N} + V_{i,s,N}$$

What do the properties of $v(k)$ and $u(\ell)$ (that they are uncorrelated $\forall k, \ell$, etc.) mean for operations with this equation?

Bridge between Statistics and Linear Algebra?

The Answer is: Ergodicity!

Condition 1: $v(k)$ and $u(\ell)$ uncorrelated

Statistically:

$$E[v(k)u^T(\ell)] = 0 \quad \forall k, \ell$$

Condition 1: $v(k)$ and $u(\ell)$ uncorrelated

Statistically:

$$E[v(k)u^T(\ell)] = 0 \quad \forall k, \ell$$

Algebraically (by ergodicity):

$$\Leftrightarrow \lim_{N\to\infty} \frac{1}{N} \sum_{j=0}^{N-1} v(k+j)u^T(\ell+j) = 0 \quad \forall k, \ell$$

Condition 1: $v(k)$ and $u(\ell)$ uncorrelated

Statistically:

$$E[v(k)u^T(\ell)] = 0 \quad \forall k, \ell$$

Algebraically (by ergodicity):

$$\Leftrightarrow \lim_{N\to\infty} \frac{1}{N} \sum_{j=0}^{N-1} v(k+j)u^T(\ell+j) = 0 \quad \forall k, \ell$$

Link to the data equation:

$$\Leftrightarrow \frac{1}{N} \begin{bmatrix} v(k) & v(k+1) & \cdots & v(k+N-1) \end{bmatrix} \begin{bmatrix} u^T(\ell) \\ u^T(\ell+1) \\ \vdots \\ u^T(\ell+N-1) \end{bmatrix} = O_N$$

with $\lim_{N\to\infty} O_N = 0$

Use of Condition 1 in SI

Consider the data equation: $\boxed{Y_{(i,s,)N} = \mathcal{O}_s X_{(i,)N} + \mathcal{T}_s U_{(i,s,)N} + V_{(i,s,)N}}$

$$\Leftrightarrow \frac{1}{N} \begin{bmatrix} v(i) & v(i+1) & \cdots & v(i+N-1) \\ v(i+1) & & & \\ \vdots & & \ddots & \\ v(i+s-1) & & & \end{bmatrix} \begin{bmatrix} u^T(i) & u^T(i+1) & \cdots \\ u^T(i+1) & & \\ \vdots & & \ddots \\ u^T(i+N-1) & & \end{bmatrix} = O_N$$

Use of Condition 1 in SI

Consider the data equation: $\boxed{Y_{(i,s,)N} = \mathcal{O}_s X_{(i,)N} + \mathcal{T}_s U_{(i,s,)N} + V_{(i,s,)N}}$

$$\Leftrightarrow \frac{1}{N}\begin{bmatrix} v(i) & v(i+1) & \cdots & v(i+N-1) \\ v(i+1) & & & \\ \vdots & & \ddots & \\ v(i+s-1) & & & \end{bmatrix}\begin{bmatrix} u^T(i) & u^T(i+1) & \cdots \\ u^T(i+1) & & \\ \vdots & & \ddots \\ u^T(i+N-1) & & \end{bmatrix} = O_N$$

$$\Leftrightarrow \frac{1}{N} V_N U_N^T = O_N$$

Use of Condition 1 in SI

Consider the data equation: $\boxed{Y_{(i,s,)N} = \mathcal{O}_s X_{(i,)N} + \mathcal{T}_s U_{(i,s,)N} + V_{(i,s,)N}}$

$$\Leftrightarrow \frac{1}{N} \begin{bmatrix} v(i) & v(i+1) & \cdots & v(i+N-1) \\ v(i+1) & & & \\ \vdots & & \ddots & \\ v(i+s-1) & & & \end{bmatrix} \begin{bmatrix} u^T(i) & u^T(i+1) & \cdots \\ u^T(i+1) & & \\ \vdots & & \ddots \\ u^T(i+N-1) & & \end{bmatrix} = O_N$$

$$\Leftrightarrow \frac{1}{N} V_N U_N^T = O_N$$

$$\begin{aligned} R_{22}Q_2 &= Y_N \Pi_{U_N}^{\perp} \\ &= \left(\mathcal{O}_s X_N + V_N\right)\left(I_N - U_N^T (U_N U_N^T)^{-1} U_N\right) \\ &= \left(\mathcal{O}_s X_N \Pi_{U_N}^{\perp} + V_N + O_N\right) \end{aligned}$$

Condition 2: $v(k)$ is zero-mean white noise

Statistically:

$$E[v(k)v^T(\ell)] = \sigma_v^2 \Delta(k - \ell)$$

Condition 2: $v(k)$ is zero-mean white noise

Statistically:

$$E[v(k)v^T(\ell)] = \sigma_v^2 \Delta(k-\ell)$$

Algebraically (by ergodicity):

$$\Leftrightarrow \lim_{N\to\infty} \frac{1}{N} \sum_{j=0}^{N-1} v(k+j)v^T(\ell+j) = \begin{cases} \sigma_v^2 I & \text{for} \quad k = \ell \\ 0 & \text{for} \quad k \neq \ell \end{cases}$$

Condition 2: $v(k)$ is zero-mean white noise

Statistically:

$$E[v(k)v^T(\ell)] = \sigma_v^2 \Delta(k-\ell)$$

Algebraically (by ergodicity):

$$\Leftrightarrow \lim_{N\to\infty} \frac{1}{N} \sum_{j=0}^{N-1} v(k+j)v^T(\ell+j) = \begin{cases} \sigma_v^2 I & \text{for} \quad k=\ell \\ 0 & \text{for} \quad k \neq \ell \end{cases}$$

Link to the data equation:

$$\Leftrightarrow \frac{1}{N} \begin{bmatrix} v(k) & v(k+1) & \cdots & v(k+N-1) \end{bmatrix} \begin{bmatrix} v^T(\ell) \\ v^T(\ell+1) \\ \vdots \\ v^T(\ell+N-1) \end{bmatrix} = \begin{cases} \sigma_v^2 I + O_N & \text{for} \quad k=\ell \\ 0 + O_N & \text{for} \quad k \neq \ell \end{cases}$$

Use of Condition 2 in SI

Consider the data equation: $\boxed{Y_{(i,s,)N} = \mathcal{O}_s X_{(i,)N} + \mathcal{T}_s U_{(i,s,)N} + V_{(i,s,)N}}$

$$\Leftrightarrow \frac{1}{N} \begin{bmatrix} v(i) & v(i+1) & \cdots & v(i+N-1) \\ v(i+1) & & & \\ \vdots & & \ddots & \\ v(i+s-1) & & & \end{bmatrix} \begin{bmatrix} v^T(i) & v^T(i+1) & \cdots \\ v^T(i+1) & & \\ \vdots & & \ddots \\ v^T(i+N-1) & & \end{bmatrix} = \sigma_v^2 I + O_N$$

Use of Condition 2 in SI

Consider the data equation: $\boxed{Y_{(i,s,)N} = \mathcal{O}_s X_{(i,)N} + \mathcal{T}_s U_{(i,s,)N} + V_{(i,s,)N}}$

$$\Leftrightarrow \frac{1}{N}\begin{bmatrix} v(i) & v(i+1) & \cdots & v(i+N-1) \\ v(i+1) & & & \\ \vdots & & \ddots & \\ v(i+s-1) & & & \end{bmatrix}\begin{bmatrix} v^T(i) & v^T(i+1) & \cdots \\ v^T(i+1) & & \\ \vdots & & \ddots \\ v^T(i+N-1) & & \end{bmatrix} = \sigma_v^2 I + O_N$$

$$\begin{aligned}
\frac{1}{N}R_{22}R_{22}^T &= \frac{1}{N}\Big(R_{22}Q_2\Big)\Big(Q_2^T R_{22}^T\Big) \\
&= \frac{1}{N}\Big(Y_N \Pi_{U_N}^{\perp}\Big)\Big(\Pi_{U_N}^{\perp} Y_N^T\Big) \\
&= \frac{1}{N}\Big(\mathcal{O}_s X_N \Pi_{U_N}^{\perp} + V_N + O_N\Big)\Big(\mathcal{O}_s X_N \Pi_{U_N}^{\perp} + V_N + O_N\Big)^T \\
&= \frac{1}{N}\mathcal{O}_s X_N \Pi_{U_N}^{\perp} X_N^T {\mathcal{O}_s}^T + \frac{1}{N}\mathcal{O}_s X_N \Pi_{U_N}^{\perp} V_N^T + (\bullet)^T + \frac{1}{N}V_N V_N^T + O_N \\
&= \frac{1}{N}\mathcal{O}_s X_N \Pi_{U_N}^{\perp} X_N^T {\mathcal{O}_s}^T + \frac{1}{N}\mathcal{O}_s X_N \Pi_{U_N}^{\perp} V_N^T + (\bullet)^T + \sigma_v^2 I + O_N
\end{aligned}$$

Condition 3: $x(k)$ and $v(\ell)$ are uncorrelated

Statistically: **Lemma:** Let the following state space model be given:

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \quad \text{with} \quad E[x(0)v^T(\ell)] = 0 \ \ \forall \ell \\ y(k) &= Cx(k) + v(k) \end{aligned}$$

with $E[v(k)u^T(\ell)] = 0$ ($\forall k, \ell$), then

$$E[x(k)v^T(\ell)] = 0 \quad \forall k, \ell$$

Condition 3: $x(k)$ and $v(\ell)$ are uncorrelated

Statistically: **Lemma:** Let the following state space model be given:

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \quad \text{with} \quad E[x(0)v^T(\ell)] = 0 \ \ \forall \ell \\ y(k) &= Cx(k) + v(k) \end{aligned}$$

with $E[v(k)u^T(\ell)] = 0$ ($\forall k, \ell$), then

$$E[x(k)v^T(\ell)] = 0 \quad \forall k, \ell$$

Proof: The state sequence equals:

$$x(k) = A^k x(0) + \sum_{j=0}^{k-1} A^j B u(k-j-1)$$

Therefor,

$$E[x(k)v^T(\ell)] = A^k E[x(0)v^T(\ell)] + \sum_{j=0}^{k-1} A^j B E[u(k-j-1)v^T(\ell)] = 0$$

Condition 3: $x(k)$ and $v(\ell)$ are uncorrelated (C'td)

Algebraically (by ergodicity):

$$\Leftrightarrow \lim_{N\to\infty} \frac{1}{N} X_N V_N^T = 0$$

Use of Condition 3 in SI

$$\frac{1}{N} R_{22} R_{22}^T = \frac{1}{N} \mathcal{O}_s X_N \Pi_{U_N}^{\perp} X_N^T \mathcal{O}_s{}^T + \frac{1}{N} \mathcal{O}_s X_N \Pi_{U_N}^{\perp} V_N^T + (\bullet)^T + \sigma_v^2 I + O_N$$

$$\frac{1}{N} R_{22} R_{22}^T = \underbrace{\frac{1}{N} \mathcal{O}_s X_N \Pi_{U_N}^{\perp} X_N^T \mathcal{O}_s{}^T} + \sigma_v^2 I + O_N$$

Use of Condition 3 in SI

$$\frac{1}{N}R_{22}R_{22}^T = \frac{1}{N}\mathcal{O}_s X_N \Pi_{U_N}^{\perp} X_N^T \mathcal{O}_s{}^T + \frac{1}{N}\mathcal{O}_s X_N \Pi_{U_N}^{\perp} V_N^T + (\bullet)^T + \sigma_v^2 I + O_N$$

$$\frac{1}{N}R_{22}R_{22}^T = \underbrace{\frac{1}{N}\mathcal{O}_s X_N \Pi_{U_N}^{\perp} X_N^T \mathcal{O}_s{}^T} + \sigma_v^2 I + O_N$$

Let the SVD of the underbraced term be given as:

$$\begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_n & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} U_1^T \\ U_2^T \end{bmatrix}$$

then,

$$\begin{aligned} \frac{1}{N}R_{22}R_{22}^T &= \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_n & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} U_1^T \\ U_2^T \end{bmatrix} + \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \sigma_v^2 I & 0 \\ 0 & \sigma_v^2 I \end{bmatrix} \begin{bmatrix} U_1^T \\ U_2^T \end{bmatrix} + O_N \\ &= \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_n + \sigma_v^2 I & 0 \\ 0 & \sigma_v^2 I \end{bmatrix} \begin{bmatrix} U_1^T \\ U_2^T \end{bmatrix} + O_N \end{aligned}$$

Consistency of the deterministic subspa

Theorem: Let $u(k)$ be an ergodic sequence such that

$$\text{rank}\left(\begin{bmatrix} X_{0,N} \\ U_{0,s,N} \end{bmatrix}\right) = n + sm$$

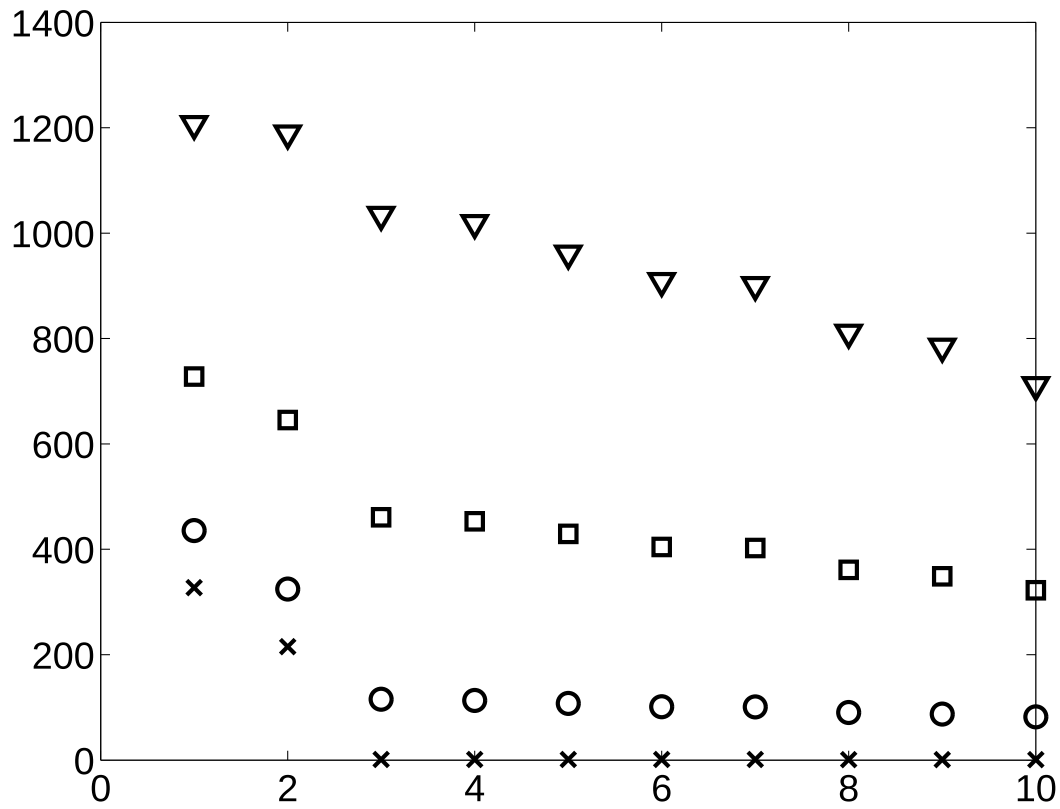
If $v(k)$ is an ergodic **white noise sequence** that is uncorrelated with $u(k)$, we have

$$\text{range}\left(\lim_{N\to\infty} \frac{1}{\sqrt{N}} Y_{0,s,N} \Pi^{\perp}_{U_{0,s,N}}\right) = \text{range}(\mathcal{O}_s)$$

subid_basic.m

Noisy singular values

Singular values of R_{22} for different values of σ:
$\sigma = 1$ (crosses)
$\sigma = 10$ (circles)
$\sigma = 20$ (squares)
$\sigma = 30$ (triangles).



Determine the system order!

Overview

- Recap of the deterministic solution
- Formulating the noise perturbation cases
- Consistency analysis: Additive white noise disturbance to the output (Case 0)
- **Consistency other cases: Use of Instrumental Variables (Tomorrow Final day)**

Preparation for this afternoon

Preparation:

1. Study Chapter 9 (9.1 - 9.3)

Download Homework 4