Subspace State Space System Identification

SysIDEA
IIT Gandhinagar

October 28, 2014

Introduction

- Subspace system identification methods combine ***numerical linear algebra***, ***systems theory*** and ***geometry*** to obtain a (***state space***) representation for the input-output data gathered from experimentation.
- Subspace methods for system identification involve the following important operations:
  - Creation of block Hankel matrices
  - Projection of Hankel matrices
  - Singular Value Decomposition
  - Least Squares Solution

Problem Definition: Deterministic System

Available Information:

- Input data: $[u(0)\ u(1)\ u(2)\ \cdots\ ]$
- Output data: $[y(0)\ y(1)\ y(2)\ \cdots\ ]$

Dataset is infinitely large and *ergodic*

Assuming that the data could be fit into a state space model representation given by,

$$x_{k+1} = Ax_k + Hu_k + w_k$$

$$y_k = Cx_k + Gu_k + v_k$$

we must find:

- n - order of the system
- A, H, C, G matrices
- State of the system x

Ergodic: Statistical properties like mean and variance do not change erratically over entire range of data

The general idea of Subspace Identification

1. Generate ***'subspaces'*** from input-output data
2. Get ***orthogonal projections*** of certain subspaces
3. Projections result into ***extended observability matrix*** or ***state estimates***
4. ***Order*** of the system can be known by taking SVD of the extended observability matrix Γ_i
5. $\boldsymbol{A}$, $\boldsymbol{C}$ matrices can be deduced from extended observability matrix Γ_i
6. $\boldsymbol{H}$, $\boldsymbol{G}$ matrices can be calculated by least squares

Subspace structure of linear systems

- Consider the matrix equation,

$$Y_f = \Gamma_i X_i + M_i^d U_f + M_i^s M_f + N_f$$

where,

$$\Gamma_i \triangleq \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{i-1} \end{bmatrix} ; M_i^d \triangleq \begin{bmatrix} G & 0 & \cdots & 0 & 0 \\ CH & G & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ CA^{i-2}H & CA^{i-3}H & \cdots & G & 0 \\ CA^{i-1}H & CA^{i-2}H & \cdots & CH & G \end{bmatrix}$$

and

$$M_i^s \triangleq \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ CH & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ CA^{i-2}H & CA^{i-3}H & \cdots & 0 & 0 \\ CA^{i-1}H & CA^{i-2}H & \cdots & CH & 0 \end{bmatrix}$$

---

i is similar to r in the notation we generally use...

Subspace structure of linear systems

- The input and output block Hankel matrices are defined as:

$$U_{0|i-1} \triangleq \begin{bmatrix} u_0 & u_1 & \cdots & u_{j-1} \\ u_1 & u_2 & \cdots & u_j \\ \vdots & \vdots & \ldots & \vdots \\ u_{i-1} & u_i & \cdots & u_{i+j-2} \end{bmatrix}; Y_{0|i-1} \triangleq \begin{bmatrix} y_0 & y_1 & \cdots & y_{j-1} \\ y_1 & y_2 & \cdots & y_j \\ \vdots & \vdots & \ldots & \vdots \\ y_{i-1} & y_i & \cdots & y_{i+j-2} \end{bmatrix}$$

- $U_p \triangleq U_{0|i-1}$ and $U_f \triangleq U_{i|2i-1}$
- $Y_p \triangleq Y_{0|i-1}$ and $Y_f \triangleq Y_{i|2i-1}$
- The subscripts f and p denote ***future*** and the ***past*** respectively.
- $W_p \triangleq \begin{bmatrix} Y_p \\ U_p \end{bmatrix}$
- Stochastically, $j \rightarrow \infty$, but for practical purposes, $j = n_y - 2i + 1$

Orthogonal Projection

Consider two matrices $\alpha \in \mathbb{R}^{p \times j}$ and $\beta \in \mathbb{R}^{q \times j}$. Then,

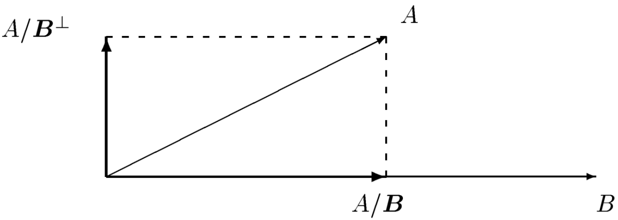
- The orthogonal projection of row space of α to the row space of β is given by,

$$\alpha/\beta = \alpha\beta^{\dagger}\beta = \alpha\beta^{T}(\beta\beta^{T})^{-1}\beta$$

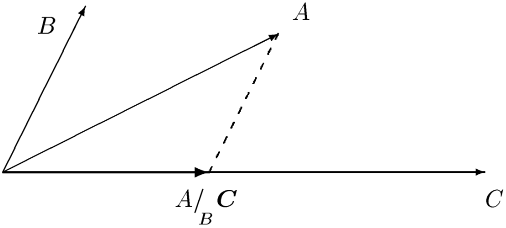
  † indicates the Moore-Penrose generalized pseudo-inverse.

- Also, $\alpha/\beta^{\perp} = \alpha - \alpha/\beta$, is the orthogonal projection of row space of α onto the perpendicular (orthogonal complement) of row space β

- It must be noted that, $\beta/\beta^{\perp} = 0$
- From other viewpoint, α/β means the ***contribution*** of β in α.



(a) Orthogonal Projection



(b) The oblique projection of rowspace of A along the rowspace of B on the row space of C

Figure 1: Projection of vectors

Finding Extended Observability matrix Γ_i and the states X_i

- Consider the matrix equation,

$$Y_f = \Gamma_i X_i + M_i^d U_f$$

- Subspace methods focus on recovering the $\Gamma_i X_i$ term from the above equation as $\Gamma_i X_i$ is a rank deficient term, which has rank n .
- Performing SVD on $\Gamma_i X_i$ provides the order n of the system.
- $\Gamma_i X_i$ can be extracted from the matrix equation by orthogonally projecting the row space of Y_f on to row space of $U_f^{\perp}$ such that:

$$Y_f/U_f^{\perp} = \Gamma_i X_i/U_f^{\perp} + M_i^d U_f/U_f^{\perp}$$

$$Y_f/U_f^{\perp} = \Gamma_i X_i/U_f^{\perp}$$

Finding Extended Observability matrix Γ_i and the states X_i

- Pre-multiplying and post-multiplying the previous equation by W_1 and W_2 weighting matrices respectively we have,

$$W_1 Y_f / U_f^{\perp} W_2 = W_1 \Gamma_i X_i / U_f^{\perp} W_2$$

  with the following conditions:

  1. $rank(W_1 \Gamma_i) = rank(\Gamma_i)$
  2. $rank(X_i / U_f^{\perp} W_2) = rank(X_i)$
  3. $rank[W_1 (M_i^s M_f + N_f) W_2] = 0$

- First and second conditions imply that $rank(\Gamma_i X_i) = n$ is preserved after the orthogonal projection.
- Third condition means that the noises w_k and v_k are uncorrelated with W_2

Finding Extended Observability matrix Γ_i and the states X_i

- Assuming W_1 and W_2 follow the three conditions stated previously, let

$$\mathcal{O}_i \triangleq W_1 Y_f / U_f^{\perp} W_2 = W_1 \Gamma_i X_i / U_f^{\perp} W_2$$

- Then $\mathcal{O}_i$ has the following SVD:

$$\mathcal{O}_i = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} S_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}$$

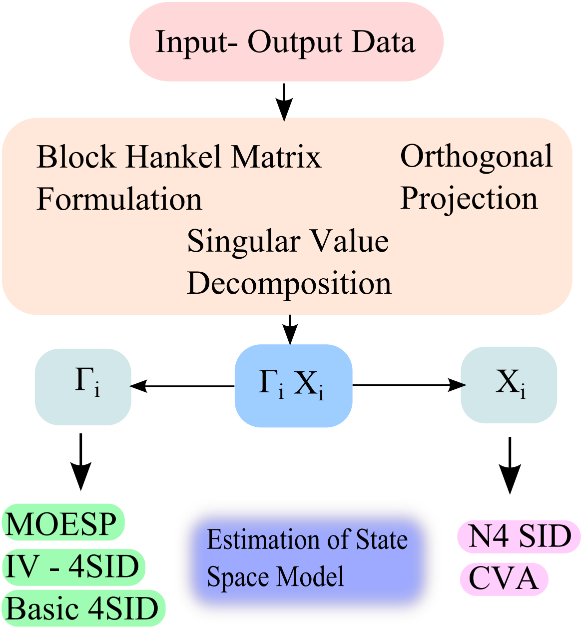
- Following properties then exist,
  1. $rank(\mathcal{O}_i) = n$
  2. $W_1 \Gamma_i = U_1 S_1^{\frac{1}{2}}$
  3. $X_i / U_f^{\perp} W_2 = S_1^{\frac{1}{2}} V_2^T$
- Also,

$$\tilde{X}_i \triangleq X_i / U_f W_2 = S_1^{\frac{1}{2}} V_2^T$$

What do weights W_1 and W_2 mean?

Method	W_1	W_2
N4SID	I_{li}	$(W_p/U_f^{\perp})^{\dagger} W_p$
CVA	$[(Y_f/U_f^{\perp})(Y_f/U_f^{\perp})^T]^{-\frac{1}{2}}$	$(W_p/U_f^{\perp})^{\dagger}(W_p/U_f^{\perp})$
MOESP	I_{li}	$(W_p/U_f^{\perp})^{\dagger}(W_p/U_f^{\perp})$
Basic **4SID**	I_{li}	I_j
IV 4SID	I_{li}	Φ

- **N4SID** - Numerical Subspace State Space Identification
- **CVA** - Continuous Variate Analysis
- **MOESP** - Multivariable Output Error State Space
- **IV**-4SID - **Instrumental Variable** Subspace State Space Identification



Input- Output Data
Block Hankel Matrix Formulation
Orthogonal Projection
Singular Value Decomposition
Γ_i
$\Gamma_i X_i$
X_i
MOESP
IV - 4SID
Basic 4SID
Estimation of State Space Model
N4 SID
CVA

Estimation of State Space Model
Using X_i

- The state space matrices A, H, C and G can be found out by solving an overdetermined set of equations by least squares.

$$\begin{bmatrix} \tilde{X}_{i+1} \\ Y_{i|i} \end{bmatrix} = \begin{bmatrix} A & H \\ C & G \end{bmatrix} \begin{bmatrix} \tilde{X}_{i+1} \\ U_{i|i} \end{bmatrix}$$

Estimation of State Space Model

Using Γ_i

- Determining A and C:
  - $A = \underline{\Gamma}_i^{\dagger}\overline{\Gamma}_i$
  - The shift structure of the matrix Γ_i is used to compute A. $\underline{\Gamma}_i$ is matrix Γ_i without last l rows and $\overline{\Gamma}_i$ is matrix Γ_i without first l rows.
  - $C = \Gamma_i(1:l,:)$, where l is the number of outputs.
- Determining H and G:

$$\Gamma_i^{\perp} Y_f U_f^{\dagger} = \Gamma_i^{\perp} M_i^d$$

  - The above equation is linear in H and G with $\Gamma_i^{\perp}$, Y_f, U_f, A and C known.

- The equation $\Gamma_i^{\perp} Y_f U_j^{\dagger} = \Gamma_i^{\perp} M_i^d$ can be written as,

$$\begin{bmatrix} \mathcal{M}_1 & \mathcal{M}_2 & \cdots & \mathcal{M}_i \end{bmatrix} = \begin{bmatrix} \mathcal{L}_1 & \mathcal{L}_2 & \cdots & \mathcal{L}_i \end{bmatrix} \times \begin{bmatrix} G & 0 & \cdots & 0 & 0 \\ CH & G & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ CA^{i-2}H & CA^{i-3}H & \cdots & G & 0 \\ CA^{i-1}H & CA^{i-2}H & \cdots & CH & G \end{bmatrix}$$

- Which is of the form,

$$\begin{bmatrix} \mathcal{M}_1 \\ \mathcal{M}_2 \\ \vdots \\ \mathcal{M}_i \end{bmatrix} = \begin{bmatrix} \mathcal{L}_1 & \mathcal{L}_2 & \cdots & \mathcal{L}_{i-1} & \mathcal{L}_i \\ \mathcal{L}_2 & \mathcal{L}_3 & \cdots & \mathcal{L}_i & 0 \\ \mathcal{L}_3 & \mathcal{L}_4 & \vdots & 0 & 0 \\ \vdots & \vdots & \cdots & 0 & 0 \\ \mathcal{L}_i & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} I_l & 0 \\ 0 & \underline{\Gamma}_i \end{bmatrix} \begin{bmatrix} G \\ H \end{bmatrix}$$

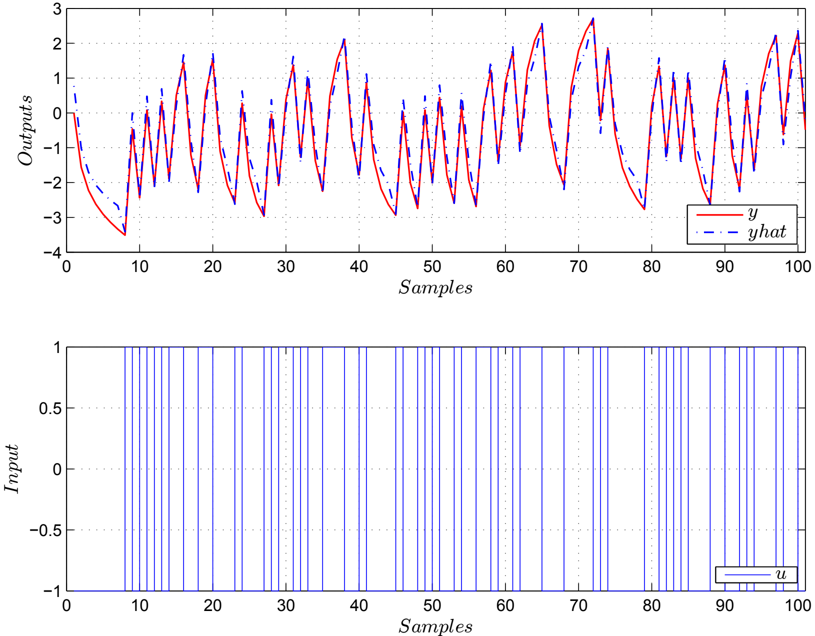
Example 1

Given System:

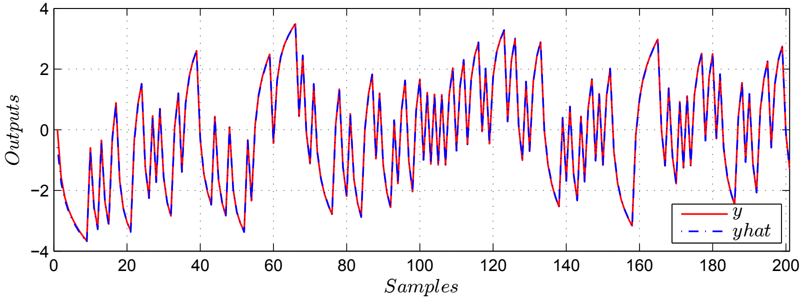
$$A = \begin{bmatrix} 0.9 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 0.2 \end{bmatrix}; H = \begin{bmatrix} 0.3 \\ 0.5 \\ 0.8 \end{bmatrix}; \; C = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}; \; G = [0]$$

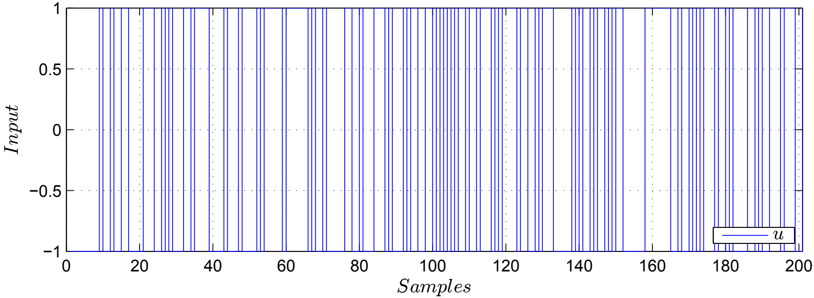
- Eigen values of the A matrix are : 0.9, 0.5 and 0.2
- Initial Conditions: $x_0 = [0.1 \; -0.1 \; 0]^T$

Result: Example 1



Result: Example 1 (More Samples)





Result: Example 1

Identified System:

$$\hat{A} = \begin{bmatrix} 0.5971 & 0.3291 & -0.0271 \\ 0.2630 & 0.5822 & 0.1939 \\ -0.0416 & 0.1343 & 0.4207 \end{bmatrix}; \hat{H} = \begin{bmatrix} -0.3206 \\ 0.0963 \\ -0.0072 \end{bmatrix};$$

$$\hat{C} = \begin{bmatrix} -4.5567 & 1.3686 & -0.1022 \end{bmatrix}; \hat{G} = [-0.0167]$$

- Eigen values of $\hat{A}$ matrix of the identified system are 0.9, 0.5 and 0.2
- Diagonalization of $\hat{A}$ by similarity transformation gives the original A matrix