System Estimation Methods III: Subspace Identification

Many estimation procedures come from:

- Statistical Principles: (Fisher, Bayes, Åström, Bohlin, Box, Jenkins, ...) — ML, PEM, Bayesian Methods
- Stochastic Realization Theory: (Ho, Kalman, Akaike, Faurre, Kailath, Lindquist, Picci, ...) — Subspace Identification Methods

SUBSPACE METHODS

Advantages:

- Non-iterative
- Numerically robust
- Natural extension to MIMO

Disadvantages:

- Difficult to analyze
- Non-efficient in general (*exception*: Larimore CVA method)

MODEL

System (MIMO):

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t + w_t \\ y_t &= Cx_t + Du_t + v_t \end{aligned}$$

Input:	$u_t \in \mathbb{R}^m$	Process noise:	$w_t \in \mathbb{R}^n$
State:	$x_t \in \mathbb{R}^n$	Measurement noise:	$v_t \in \mathbb{R}^p$
Output:	$y_t \in \mathbb{R}^p$		

Assumptions: $\{w_t\}$ and $\{v_t\}$ are white noise sequences

There are no constraints on A, B, C, D

BASIC IDEA

If not only u_t and y_t are measured, but also x_t, then estimating A, B, C, D is just a LS problem!

$$Y_t = \Theta \Phi_t + E_t$$

where

$$Y_t := \begin{bmatrix} x_{t+1} \\ y_t \end{bmatrix}, \qquad \Theta := \begin{bmatrix} A & B \\ C & D \end{bmatrix}, \qquad \Phi_t := \begin{bmatrix} x_t \\ u_t \end{bmatrix}, \qquad E_t := \begin{bmatrix} w_t \\ v_t \end{bmatrix}$$

Remaining problem: How can we estimate the state x_t?

Idea: For some realizations, x_t can be interpreted as *predictors*

PREDICTORS

From the system equations:

$$
\begin{aligned}
y_t &= Cx_t + Du_t + v_t \\
y_{t+1} &= Cx_{t+1} + Du_{t+1} + v_{t+1} \\
&= CAx_t + CBu_t + Du_{t+1} + Cw_t + v_{t+1} \\
&\vdots \\
y_{t+k} &= CA^k x_t + \\
&\quad CA^{k-1}Bu_t + CA^{k-2}Bu_{t+1} + \cdots + CBu_{t+k-1} + Du_{t+k} + \\
&\quad CA^{k-1}w_t + CA^{k-2}w_{t+1} + \cdots + Cw_{t+k-1} + v_{t+k}
\end{aligned}
$$

PREDICTORS (CONT.)

or

$$\boxed{Y_t^r = O^r x_t + S^r U_t^r + V_t} \qquad \textbf{Fundamental Equation}$$

where

$$Y^r := \begin{bmatrix} y_t \\ y_{t+1} \\ \vdots \\ y_{t+r-1} \end{bmatrix}, \quad U^r := \begin{bmatrix} u_t \\ u_{t+1} \\ \vdots \\ u_{t+r-1} \end{bmatrix}, \quad V_t := \begin{bmatrix} v_t \\ Cw_t + v_{t+1} \\ \vdots \\ CA^{r-2}w_t + CA^{r-3}w_{t+1} + \cdots + Cw_{t+r-2} + v_{t+r-1} \end{bmatrix}$$

$$O^r := \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{r-1} \end{bmatrix}, \qquad S^r := \begin{bmatrix} D & 0 & \cdots & 0 & 0 \\ CB & D & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ CA^{r-2}B & CA^{r-3}B & \cdots & CB & D \end{bmatrix}$$

PREDICTORS (CONT.)

If $u_t = \cdots = u_{t+r-1} = 0$, then $O^r x_t$ is the predictor $\hat{Y}^r_{t|t-1} = E\{Y^r \mid y_{t-1}, u_{t-1}, \ldots\}$.
This predictor thus correspond to a state vector in a particular realization.
Furthermore, if $\hat{\mathbf{Y}} := [\hat{Y}^r_{1|0} \quad \cdots \quad \hat{Y}^r_{N|N-1}]$, then as $N \to \infty$:

1. The model has minimal order n iff $\operatorname{rank}\{\hat{\mathbf{Y}}\} = n$ for all $r \geq n$
2. In innovations form: $x_t = L\hat{Y}^r_{t|t-1}$ for some $L \in \mathbb{R}^{n \times pr}$.

$\Rightarrow$ $\boxed{\text{We could estimate } n \text{ from the "rank" of } \hat{\mathbf{Y}}\text{, and choose } x_t \text{ as } L\hat{Y}^r_{t|t-1} \text{ for some } L}$

Estimating the Predictors: Basic 4SID Approach

Idea: Estimate the predictors $\hat{Y}^r_{t|t-1}$ via LS!

$$\boxed{Y_t^r = \Theta \varphi_t^s + \Gamma U_t^l + \mathrm{E}_t}$$

where

$$\varphi_t^s := [y_{t-1}^T \quad \cdots \quad y_{t-s_1}^T \quad u_{t-1}^T \quad \cdots \quad u_{t-s_2}^T]^T$$

$$Y_t^r := [y_t^T \quad \cdots \quad y_{t+r-1}^T]^T, \quad U_t^l := [u_t^T \quad \cdots \quad u_{t+l-1}^T]^T$$

$$\mathrm{E}_t := [\varepsilon_t^T \quad \cdots \quad \varepsilon_{t+r-1}^T]^T$$

$$\Theta := [\theta_1 \quad \cdots \quad \theta_r]^T, \quad \Gamma := [\gamma_1 \quad \cdots \quad \gamma_r]^T$$

and then:

$$\boxed{\hat{Y}^r_{t|t-1} = \hat{\Theta} \varphi_t^s}$$

User choices: Integers s_1, s_2, r and l

SUMMARY OF BASIC SUBSPACE ALGORITHM

1. Choose s_1, s_2, r and l, and apply LS to estimate Θ, Γ from

$$Y_t^r = \Theta \varphi_t^s + \Gamma U_t^l + \mathrm{E}_t$$

Then, construct $\hat{Y}_{t|t-1}^r = \hat{\Theta}\varphi_t^s$ and $\hat{\mathbf{Y}} = [\hat{Y}_{1|0}^r \quad \cdots \quad \hat{Y}_{N|N-1}^r]$

2. Estimate $\operatorname{rank}\{\hat{\mathbf{Y}}\}$ and determine L to construct a well-conditioned $\hat{x}_t = L\hat{Y}_{t|t-1}^r$

3. Estimate A, B, C, D via LS on

$$\begin{bmatrix} \hat{x}_{t+1} \\ y_t \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \hat{x}_t \\ u_t \end{bmatrix} + \begin{bmatrix} w_t \\ v_t \end{bmatrix}$$