System Identification, Lecture 8

Kristiaan Pelckmans (IT/UU, 2338)

Course code: 1RT885, Report code: 61800 - Spring 2017
F, FRI Uppsala University, Information Technology

16 Mai 2017

Things I need to say

1. Projects (Go!).
2. Computer lab reports (22/05).
3. Guest lectures.
4. Presentetations project (26/05)
5. Exam (31/05 - 8-12am).

Projects

What do I expect from you:

1. I give you data + description - you give me *good* model.
2. Single out a SISO problem, make a model and assess why/whynot satisfactory.
3. Set a baseline - where do you want to improve on?
4. Make model of MIMO system.
5. Make plots of the results, and interpret results. What is good? What is not good?
6. Use for intended purpose.
7. What's next?

State Space System

$$\begin{cases} \mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{B}\mathbf{u}_t \\ \mathbf{y}_t = \mathbf{C}\mathbf{x}_t + \mathbf{D}\mathbf{u}_t, \end{cases} \quad \forall t = -\infty, \ldots, \infty.$$

with

- $\{\mathbf{x}_t\}_t \subset \mathbb{R}^n$ the state process.
- $\{\mathbf{u}_t\}_t \subset \mathbb{R}^p$ the input process.
- $\{\mathbf{y}_t\}_t \subset \mathbb{R}^q$ the output process.
- $\mathbf{A} \in \mathbb{R}^{n \times n}$ the system matrix.
- $\mathbf{B} \in \mathbb{R}^{n \times p}$ the input matrix.
- $\mathbf{C} \in \mathbb{R}^{q \times n}$ the output matrix.
- $\mathbf{D} \in \mathbb{R}^{q \times p}$ the feed-through matrix.

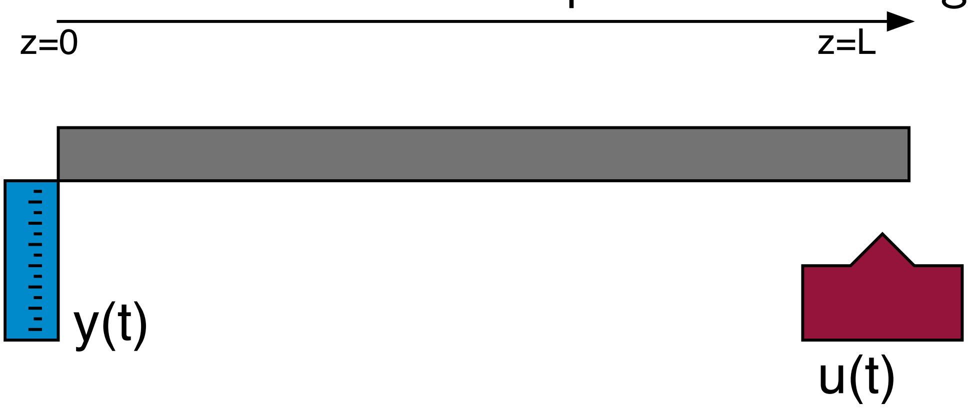
State Space System

Advantages over fractional polynomial models

- Closer to physical modeling.
- Control!
- MIMO systems.
- Noise and Innovations.
- Canonical representation.
- Problems of identiafibility.

State Space System - ex. 1

From PDE to state-space: the heating-rod system:



Let $x(t,z)$ denote temperature at time t, and location z on the rod.

$$\frac{\partial x(t,z)}{\partial t} = \kappa \frac{\partial^2 x(t,z)}{\partial z^2}$$

The heating at the far end means that

$$\left.\frac{\partial x(t,z)}{\partial z}\right|_{z=L} = K u(t),$$

The near-end is insulated such that

$$\left.\frac{\partial x(t,z)}{\partial z}\right|_{z=0} = 0.$$

The measurements are

$$y(t) = x(t,0) + v(t), \forall t = 1, 2, \ldots$$

The unknown parameters are

$$\theta = \begin{bmatrix} \kappa \\ K \end{bmatrix}$$

This can be approximated as a system with n states

$$\mathbf{x}(t) = \Big(x(t,z_1), x(t,z_2), \ldots, x(t,z_n)\Big)^T \in \mathbb{R}^n$$

with $z_k = L(k-1)/(n-1)$.. Then we use the approximation that

$$\frac{\partial^2 x(t,z)}{\partial z^2} \approx \frac{x(t,z_{k+1}) - 2x(t,z_k) + x(t,z_{k-1})}{(L/(n-1))^2}$$

where $z_k = \operatorname{argmin}_{z_1,,\ldots,z_n} \|z - z_k\|$. Hence the continuous state-space approximation becomes

$$\begin{cases} \dot{\mathbf{x}}(t) = \left(\frac{n-1}{L}\right)^2 \kappa \begin{bmatrix} -1 & 1 & & & \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \end{bmatrix} \mathbf{x}(t) + \begin{bmatrix} 0 \\ \vdots \\ K \end{bmatrix} u(t) \\ y(t) = \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix} \mathbf{x}(t) + v(t) \end{cases}$$

and a discrete Euler approximation

$$\begin{cases} \mathbf{x}_{t+1} - \mathbf{x}_t & = \Delta' \left(\frac{n-1}{L}\right)^2 \kappa \\ & \quad \begin{bmatrix} -1 & 1 & & & \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \end{bmatrix} \mathbf{x}_t + \Delta' \begin{bmatrix} 0 \\ \vdots \\ K \end{bmatrix} \int_{\Delta'} u(t) \\ y_t & = \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix} \mathbf{x}_t + \int_{\Delta'} v(t) \end{cases}$$

State Space System - ex. 2

Models for the future size of the population (UN, WWF).

Leslie model: key ideas: discretize population in n aging groups and

- Let $\mathbf{x}_{t,i} \in \mathbb{R}^+$ denote the size of the ith aging group at time t.

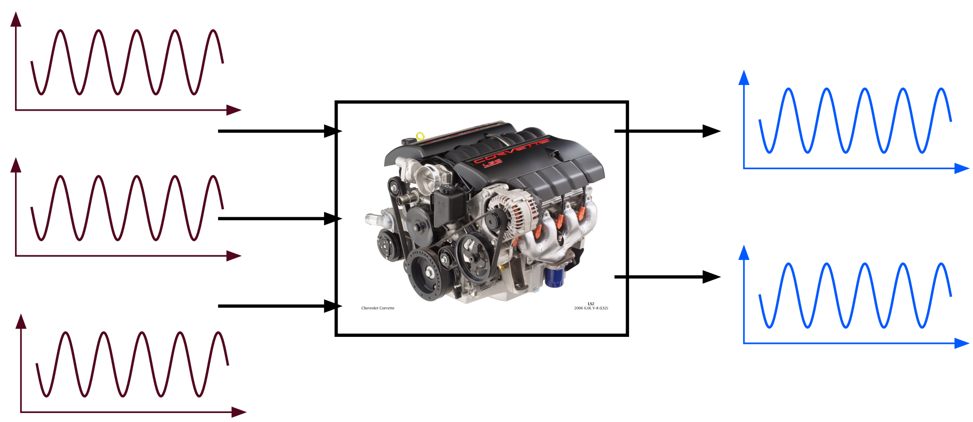
- Let $\mathbf{x}_{t+1,i+1} = s_i \mathbf{x}_{t,i}$ with $s_i \geq 0$ the 'survival' coefficient.
- Let $\mathbf{x}_{t+1,1} = s_0 \sum_{i=1}^n f_i \mathbf{x}_{t,i}$ with $f_i \geq 0$ the 'fertility' rate.

Hence, the dynamics of the population may be captured by the following discrete time model

$$\begin{cases} \mathbf{x}_{t+1} = \begin{bmatrix} s_0 f_1 & s_0 f_2 & \dots & s_0 f_n & \\ s_1 & 0 & & & \\ 0 & s_2 & 0 & & \\ & & \ddots & & \\ & & & s_{n-1} & 0 \end{bmatrix} \mathbf{x}_t + \mathbf{u}_t \\ y_t = \sum_{i=1}^n \mathbf{x}_{t,i} \end{cases}$$

Impulse Response to State Space System

What is now the relation of state-space machines, and the system theoretic tools seen in the previous Part?



Recall impulse response (SISO)

$$y_t = \sum_{\tau=0}^{\infty} h_\tau u_{t-\tau},$$

and MIMO

$$\mathbf{y}_t = \sum_{\tau=0}^{\infty} \mathbf{H}_\tau \mathbf{u}_{t-\tau},$$

where $\{\mathbf{H}_\tau\}_\tau \subset \mathbb{R}^{p\times q}$.

Recall: System identification studies method to build a model from observed input-output behaviors, i.e. $\{\mathbf{u}_t\}_t$ and $\{\mathbf{y}_t\}_t$.

Now it is a simple exercise to see which impulse response matrices $\{\mathbf{H}_\tau\}_\tau$ are implemented by a state-space model with matrices $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$:

$$\mathbf{H}_\tau = \begin{cases} \mathbf{D} & \tau = 0 \\ \mathbf{C}\mathbf{A}^{\tau-1}\mathbf{B} & \tau = 1, 2, \ldots \end{cases}, \ \forall \tau = 0, 1, 2, \ldots$$

Contrast with rational polynomials where typically

$$h_\tau \Leftrightarrow h(q^{-1}) = \frac{b_1 q^{-1} + b_2 q^{-2} + \ldots}{1 + a_1 q^{-1} + a_2 q^{-2} + \ldots}$$

Overlapping: consider FIR model

$$y_t = b_0 u_t + b_1 u_{t-1} + b_2 u_{t-2} + e_t$$

then equivalent state-space with states $\mathbf{x}_t = (u_t, u_{t-1}, u_{t-2})^T \in \mathbb{R}^3$ becomes

$$\begin{cases} \mathbf{x}_t &= \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \mathbf{x}_{t-1} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u_t \\ y_t &= \begin{bmatrix} b_0 & b_1 & b_2 \end{bmatrix} \mathbf{x}_t + e_t \end{cases}$$

and $\mathbf{x}_0 = (u_0, u_{-1}, u_{-2})^T$.

Controllability and Observability

A state-space model is said to be Controllable iff for any terminal state $\mathbf{x} \in \mathbb{R}^n$ one has that for all initial state $\mathbf{x}_0 \in \mathbb{R}^n$, there exists an input process $\{\mathbf{u}_t\}_t$ which steers the model from state $\mathbf{x}_0$ to $\mathbf{x}$.

A state-space model is said to be Reachable iff for any initial state $\mathbf{x}_0 \in \mathbb{R}^n$ one has that for all terminal states $\mathbf{x} \in \mathbb{R}^n$ there exists an input process $\{\mathbf{u}_t\}_t$ which steers the model from state $\mathbf{x}_0$ to $\mathbf{x}$.

The mathematical definition goes as follows: Define the reachability matrix $\mathcal{C} \in \mathbb{R}^{n \times np}$ as

$$\mathcal{C} = \begin{bmatrix} \mathbf{B} & \mathbf{AB} & \mathbf{A}^2\mathbf{B} & \dots & \mathbf{A}^{n-1}\mathbf{B} \end{bmatrix}$$

The State space $(\mathbf{A}, \mathbf{B})$ is reachable (controllable) if

$$\text{rank}(\mathcal{C}) = n.$$

Intuition: if the matrix $\mathcal{C}$ is full rank, the image of $\mathcal{C}$ equals $\mathbb{R}^n$, and the superposition principle states that any linear combination of states can be reached by a linear combination of inputs.

A state-space model is Observable iff any two different initial states $\mathbf{x}_0 \neq \mathbf{x}_0' \in \mathbb{R}^n$ lead to a different output $\{\mathbf{y}_s\}_{s\geq 0}$ of the state-space model in the future when the inputs are switched off henceforth (autonomous mode).

Define the Observability matrix $\mathcal{O} \in \mathbb{R}^{qn\times n}$ as

$$\mathcal{O} = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \vdots \\ \mathbf{CA}^{n-1} \end{bmatrix}$$

Hence, a state-space model $(\mathbf{A}, \mathbf{C})$ is observable iff

$$\mathrm{rank}(\mathcal{O}) = n$$

Intuition: if the (right) null space of $\mathcal{O}$ is empty, no two different $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^n$ lead to the $\mathcal{O}\mathbf{x} = \mathcal{O}\mathbf{x}'$.

Let

$$\mathbf{u}_- = (\mathbf{u}_0, \mathbf{u}_{-1}, \mathbf{u}_{-2}, \dots)^T$$

And

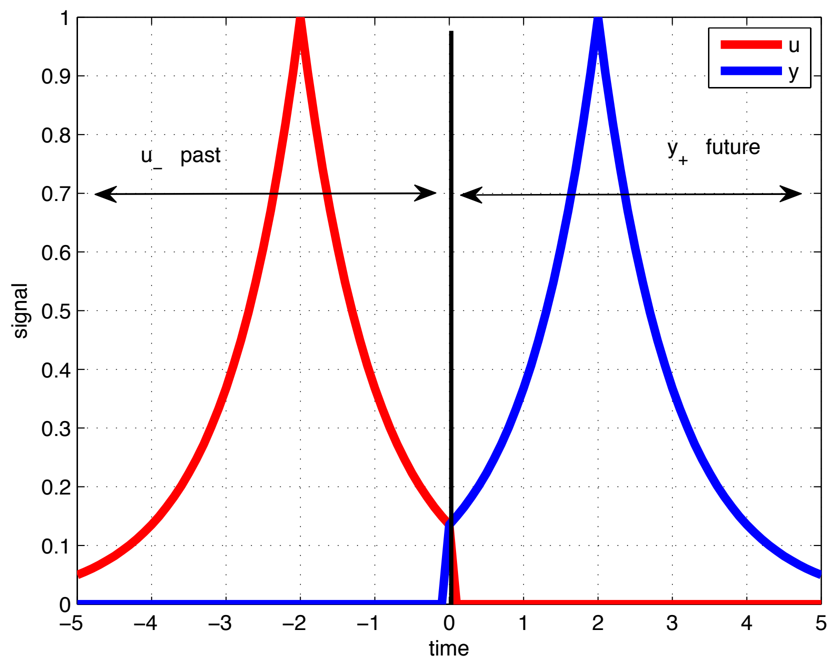
$$\mathbf{y}_+ = (\mathbf{y}_1, \mathbf{y}_2, \dots)^T$$

Then

$$\mathbf{x}_1 \propto \mathcal{C}\mathbf{u}_-$$

and

$$\mathbf{y}_+ \propto \mathcal{O}\mathbf{x}_1$$



Realization Theory

Problem: Given an impulse response sequence $\{\mathbf{H}_\tau\}_\tau$, can we recover $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$?

Def. Minimal Realization. A state-space model $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ is a minimal realization of order n iff the corresponding $\mathcal{C}$ and $\mathcal{O}$ are full rank, that is iff the model is reachable (observable) and controllable.

Thm. (Kalman) If $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ and $(\mathbf{A}', \mathbf{B}', \mathbf{C}', \mathbf{D}')$ are two minimal realizations of the same impulse response $\{\mathbf{H}_\tau\}$, then they are linearly related by a nonsingular matrix $\mathbf{T} \in \mathbb{R}^{n \times n}$ such that

$$\begin{cases} \mathbf{A}' = \mathbf{T}^{-1}\mathbf{A}\mathbf{T} \\ \mathbf{B}' = \mathbf{T}^{-1}\mathbf{B} \\ \mathbf{C}' = \mathbf{C}\mathbf{T} \\ \mathbf{D}' = \mathbf{D} \end{cases}$$

Intuition: a linear transformation of the states does not alter input-output behavior; that is, the corresponding $\{\mathbf{H}_\tau\}_\tau$ is the same. The thm states that those are the only transformations for which this is valid.

Hence, it is only possible to reconstruct a minimal realization of a state-space model $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ from $\{\mathbf{H}_\tau\}_\tau$ *up to a linear transformation of the states*.

In case we only observe sequences $\{\mathbf{u}_t\}_{t\geq 1}$ and $\{\mathbf{y}_t\}_{t\geq 1}$, we have to account for the transient effects and need to estimate $\mathbf{x}_0 \in \mathbb{R}^n$ as well. This is in many situations crucial. The above thm. is extended to include $\mathbf{x}_0$ as well.

Now the celebrated Kalman-Ho realization algorithm goes as follows:

- Hankel-matrix

$$\mathbf{H}^n = \begin{bmatrix} \mathbf{H}_1 & \mathbf{H}_2 & \mathbf{H}_3 & \dots & \mathbf{H}_n \\ \mathbf{H}_2 & \mathbf{H}_3 & \mathbf{H}_4 & & \\ & & \ddots & & \\ \mathbf{H}_n & & & & \mathbf{H}_{2n+1} \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{CB} & \mathbf{CAB} & \mathbf{CA}^2\mathbf{B} & \dots & \mathbf{CA}^{n-1}\mathbf{B} \\ \mathbf{CAB} & \mathbf{CA}^2\mathbf{B} & & & \\ & & \ddots & & \\ \mathbf{CA}^{n-1}\mathbf{B} & & & & \mathbf{CA}^{2n-1}\mathbf{B} \end{bmatrix} = \mathcal{OC}$$

- The state space is identifiable up to a non-singular matrix $\mathbf{T} \in \mathbb{R}^{n\times n}$ such that

$$\mathbf{H}^n = \mathcal{O}\mathcal{C} = \mathcal{O}\mathbf{T}\mathbf{T}^{-1}\mathcal{C}$$

.

- Then take the SVD of $\mathbf{H}^n$, such that

$$\mathbf{H}^n = \mathbf{U}\Sigma\mathbf{V}^T$$

with $\mathbf{U} \in \mathbb{R}^{pn\times n}, \mathbf{V} \in \mathbb{R}^{n\times nq}$ and $\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_n) \in \mathbb{R}^{n\times n}$.

- Hence a minimal realization is given as

$$\begin{cases} \mathcal{O}' = \mathbf{U}\sqrt{\Sigma} \\ \mathcal{C}' = \sqrt{\Sigma}\mathbf{V} \end{cases}$$

- From $\mathcal{O}', \mathcal{C}'$ it is not too difficult to extract $(\mathbf{A}, \mathbf{B}, \mathbf{C})$

Stochastic Systems

Stochastic disturbances (no inputs)

$$\begin{cases} X_{t+1} & = \mathbf{A} X_t + W_t \\ Y_t & = \mathbf{C} X_t + V_t \end{cases}$$

with

- $\{X_t\}_t$ the stochastic state process taking values in $\mathbb{R}^n$.
- $\{Y_t\}_t$ the stochastic output process, taking values in $\mathbb{R}^p$.
- $\mathbf{A} \in \mathbb{R}^{n \times n}$ the (deterministic) system matrix.
- $\mathbf{C} \in \mathbb{R}^{p \times n}$ the (deterministic) output matrix.
- $\{W_t\}_t$ the stochastic process disturbances taking values in $\mathbb{R}^n$.

- $\{V_t\}_t$ the stochastic measurement disturbances taking values in $\mathbb{R}^p$.

The stochastic vectors follow a probability law assumed to follow

- $\mathbb{E}[W_t] = 0_n$, and $\mathbb{E}[W_t W_s^T] = \delta_{s,t}\mathbf{Q} \in \mathbb{R}^{n\times n}$.
- $\mathbb{E}[V_t] = 0_p$, and $\mathbb{E}[V_t V_s^T] = \delta_{s,t}\mathbf{R} \in \mathbb{R}^{p\times p}$.
- $\mathbb{E}[W_t V_t^T] = \delta_{s,t}\mathbf{S} \in \mathbb{R}^{n\times p}$.
- W_t, V_t assumed independent of $\ldots, X_t$.

Main questions:

- Covariance matrix states $\mathbb{E}[X_t X_t^T] = \Pi$:

$$\Pi = \mathbf{A}\Pi\mathbf{A}^T + \mathbf{Q}$$

- Lyapunov, stable.

- Covariance matrix outputs $\mathbb{E}[Y_t Y_t^T]$.

This model can equivalently be described in its innovation form

$$\begin{cases} X'_{t+1} & = \mathbf{A}X'_t + \mathbf{K}D_t \\ Y_t & = \mathbf{C}X'_t + D_t \end{cases}$$

with $\mathbf{K} \in \mathbb{R}^{n \times p}$ the Kalman gain, such that $\mathbf{P}, \mathbf{K}$ solves

$$\begin{cases} \mathbf{P} = \mathbf{APA} + (\mathbf{G} - \mathbf{APC}^T)(\Lambda_0 - \mathbf{CPC}^T)^{-1}(\mathbf{G} - \mathbf{APC}^T)^T \\ \mathbf{K} = (\mathbf{G} - \mathbf{APC}^T)(\Lambda_0 - \mathbf{CPC}^T) \end{cases}$$

and

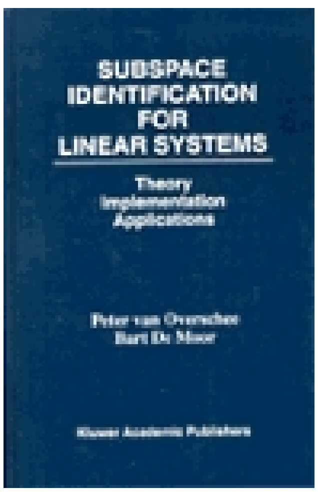
- $\mathbb{E}[D_t {D'_t}^T] = (\Lambda_0 - \mathbf{CPC}^T)$
- $\mathbf{P} = \mathbb{E}[X'_t {X'_t}^T]$

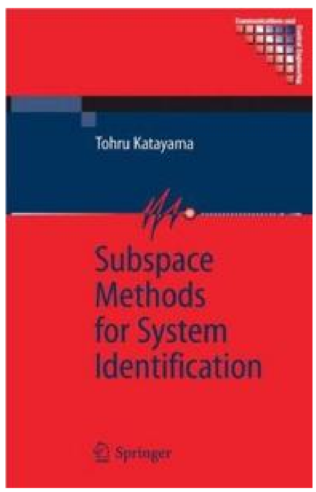
Overview Subspace Identification

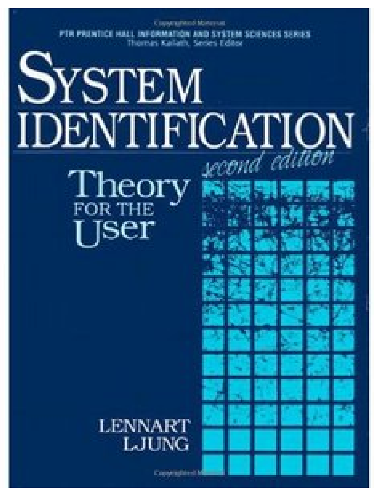
1. Deterministic.

2. Stochastic.

3. Extensions.





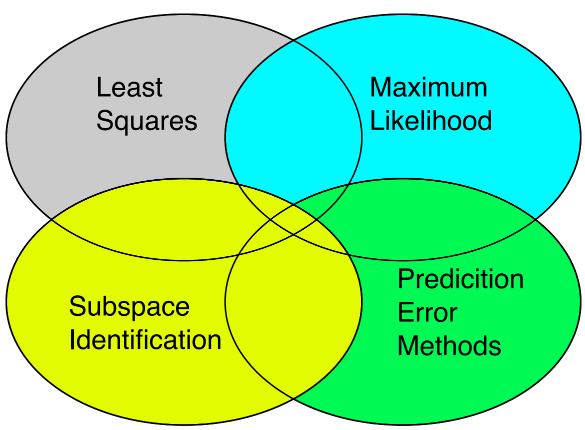


K. De Cock, B. De Moor, "Subspace Identification Methods", report, 2003.

Motivation

Why?

- MIMO.
- State space models.
- Inherent identifiability 'up to $\mathbf{T}$'.
- Numerical matching.
- Numerical Robust techniques (perturbations).
- Connection to systems theory.



State Space System

$$\begin{cases} \mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{B}\mathbf{u}_t \\ \mathbf{y}_t = \mathbf{C}\mathbf{x}_t + \mathbf{D}\mathbf{u}_t, \end{cases} \quad \forall t = -\infty, \ldots, \infty.$$

with

- $\{\mathbf{x}_t\}_t \subset \mathbb{R}^n$ the state process.
- $\{\mathbf{u}_t\}_t \subset \mathbb{R}^p$ the input process.
- $\{\mathbf{y}_t\}_t \subset \mathbb{R}^q$ the output process.
- $\mathbf{A} \in \mathbb{R}^{n \times n}$ the system matrix.
- $\mathbf{B} \in \mathbb{R}^{n \times p}$ the input matrix.
- $\mathbf{C} \in \mathbb{R}^{q \times n}$ the output matrix.
- $\mathbf{D} \in \mathbb{R}^{q \times p}$ the feed-through matrix.

Problem Statement

Problem SI: Given multivariate timeseries $\{\mathbf{u}_t\}_{t=0}^N \subset \mathbb{R}^p$ and $\{\mathbf{y}\}_{t=0}^N \subset \mathbb{R}^q$, can you figure out the order n, matrices $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ and $\{\mathbf{x}_t\}_t \subset \mathbb{R}^n$?

Realization: Given impulse response matrices $\{\mathbf{H}_\tau\}_\tau$, recover n and $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$.

A first (naive) approach:

(1) Estimate IR matrices $\{\hat{\mathbf{H}}_\tau\}_\tau$ by solving/approximating

$$\begin{bmatrix} \mathbf{y}_n^T \\ \mathbf{y}_{n+1}^T \\ \mathbf{y}_{n+2}^T \\ \vdots \\ \mathbf{y}_N^T \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1^T & \mathbf{u}_2^T & \dots & \mathbf{u}_n^T \\ \mathbf{u}_2^T & \mathbf{u}_3^T & & \mathbf{u}_{n+1}^T \\ \mathbf{u}_3^T & \mathbf{u}_4^T & & \mathbf{u}_{n+2}^T \\ \vdots & \vdots & & \vdots \\ \mathbf{u}_{N-n+1}^T & \mathbf{u}_{N-n+2}^T & & \mathbf{u}_N^T \end{bmatrix} \begin{bmatrix} \mathbf{H}_{n-1}^T \\ \mathbf{H}_{n-2}^T \\ \vdots \\ \mathbf{H}_0^T \end{bmatrix}$$

(2) Realization: transform $\{\hat{\mathbf{H}}_\tau\}_\tau$ into $\hat{n}$ and $(\hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}, \hat{\mathbf{D}})$

But:

- Computational burdensome.
- Not robust.
- PE...
- Numerically ill-conditioned.
- Process Noise.
- State-Space structure.

That's why subspace ID:

- N4SID (*'enforce it'*) (Numerical algorithm for Subspace State-space System ID)
- MOESP (Multivariate Output Error State sPace)

The Deterministic Case

(From T. Katayama, 2005) Matrix matching

$$\begin{bmatrix} \mathbf{y}_t \\ \vdots \\ \mathbf{y}_{t+k-1} \end{bmatrix} = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \mathbf{CA}^2 \\ \vdots \\ \mathbf{CA}^{k-1} \end{bmatrix} \mathbf{x}_t + \begin{bmatrix} \mathbf{D} & & & \\ \mathbf{CB} & \mathbf{D} & & \\ \vdots & & \ddots & \\ \mathbf{CA}^{k-2}\mathbf{B} & & \dots & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{u}_t \\ \vdots \\ \mathbf{u}_{t+k-1} \end{bmatrix}$$

In shorthand:

$$\mathbf{y}_k(t) = \mathcal{O}_k \mathbf{x}_t + \Psi_k \mathbf{u}_k(t)$$

This holds for any $t = 1, 2, \dots, N$, or

$$\begin{bmatrix} \mathbf{y}_k(0) & \mathbf{y}_k(1) & \dots & \mathbf{y}_k(i-1) \end{bmatrix} = \mathcal{O}_k \begin{bmatrix} \mathbf{x}_0 & \mathbf{x}_1 & \dots & \mathbf{x}_{i-1} \end{bmatrix} + \Psi_k \begin{bmatrix} \mathbf{u}_k(0) & \mathbf{u}_k(1) & \dots & \mathbf{u}_k(i-1) \end{bmatrix}$$

Or in even shorter hand

$$\mathbf{Y}_{k,0} = \mathcal{O}_k \mathbf{X}_0 + \Psi_k \mathbf{U}_{k,0}$$

Now the same trick for for data $k, \dots, k+i-1$

$$\begin{cases} \mathbf{Y}_{k,s} = \begin{bmatrix} \mathbf{y}_k(s) & \mathbf{y}_k(1) & \dots & \mathbf{y}_k(s+i-1) \end{bmatrix} \\ \mathbf{U}_{k,s} = \begin{bmatrix} \mathbf{u}_k(s) & \mathbf{u}_k(1) & \dots & \mathbf{u}_k(s+i-1) \end{bmatrix} \\ \mathbf{X}_s = (\mathbf{x}_s, \dots, \mathbf{x}_{s+i-1}) \end{cases}$$

Hence one has for all $s = 0, 1, \dots, N-i$.

$$\mathbf{Y}_{k,s} = \mathcal{O}_k \mathbf{X}_s + \Psi_k \mathbf{U}_{k,s}.$$

We will use in our exposition

$$\begin{cases} \mathbf{Y}_{k,0} = \mathcal{O}_k \mathbf{X}_0 + \Psi_k \mathbf{U}_{k,0} \\ \mathbf{Y}_{k,k} = \mathcal{O}_k \mathbf{X}_k + \Psi_k \mathbf{U}_{k,k}. \end{cases}$$

Which we will denote as the matrix input-output relations of 'past' and 'future'.

or

$$\left\{ \mathbf{U}_{k,0} = \begin{bmatrix} \mathbf{u}_0 & \mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_{i-1} \\ \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 & \dots & \mathbf{u}_i \\ \vdots & & & & \vdots \\ \mathbf{u}_{k-1} & \mathbf{u}_k & & \dots & \mathbf{u}_{k+i-2} \end{bmatrix} \in \mathbb{R}^{kp\times i} \right.$$

$$\left\{ \mathbf{Y}_{k,0} = \begin{bmatrix} \mathbf{y}_0 & \mathbf{y}_1 & \mathbf{y}_2 & \dots & \mathbf{y}_{i-1} \\ \mathbf{y}_1 & \mathbf{y}_2 & \mathbf{y}_3 & \dots & \mathbf{y}_i \\ \vdots & & & & \vdots \\ \mathbf{y}_{k-1} & \mathbf{y}_k & & \dots & \mathbf{y}_{k+i-2} \end{bmatrix} \in \mathbb{R}^{kq\times i} \right.$$

$$\left\{ \mathbf{U}_{k,k} = \begin{bmatrix} \mathbf{u}_k & \mathbf{u}_{k+1} & \mathbf{y}_{k+2} & \dots & \mathbf{u}_{k+i-1} \\ \mathbf{u}_{k+1} & \mathbf{u}_{k+2} & \mathbf{y}_{k+3} & \dots & \mathbf{u}_{k+i} \\ \vdots & & & & \vdots \\ \mathbf{u}_{2k-1} & \mathbf{u}_k & & \dots & \mathbf{u}_{k+i-2} \end{bmatrix} \in \mathbb{R}^{kq\times i} \right.$$

$$\left\{ \mathbf{Y}_{k,k} = \begin{bmatrix} \mathbf{y}_k & \mathbf{y}_{k+1} & \mathbf{y}_{k+2} & \dots & \mathbf{y}_{k+i-1} \\ \mathbf{y}_{k+1} & \mathbf{y}_2 & \mathbf{y}_3 & \dots & \mathbf{y}_{k+i} \\ \vdots & & & & \vdots \\ \mathbf{y}_{2k-1} & \mathbf{y}_{2k} & & \dots & \mathbf{y}_{2k+i-1} \end{bmatrix} \in \mathbb{R}^{kq\times i} \right.$$

Let

$$\mathbf{W}_- = \begin{bmatrix} \mathbf{U}_{k,0} \\ \mathbf{Y}_{k,0} \end{bmatrix} \qquad \mathbf{W}_+ = \begin{bmatrix} \mathbf{U}_{k,k} \\ \mathbf{Y}_{k,k} \end{bmatrix}$$

Now we study the relation of $\mathbf{W}_-, \mathbf{W}_+$ and $\mathbf{H}$. From above, we have that

$$\mathbf{W}_- = \begin{bmatrix} \mathbf{U}_{k,0} \\ \mathbf{Y}_{k,0} \end{bmatrix} = \begin{bmatrix} I_{kp} & 0 \\ \psi_k & \mathcal{O}_k \end{bmatrix} \begin{bmatrix} \mathbf{U}_{k,0} \\ \mathbf{X}_0 \end{bmatrix}$$

Or

$$\mathbf{W}_- = \begin{bmatrix} \mathbf{U}_{k,0} \\ \mathbf{Y}_{k,0} \end{bmatrix} = \begin{bmatrix} I_{kp} & 0 \\ \psi_k & \mathcal{O}_k\mathcal{C}_k \end{bmatrix} \begin{bmatrix} \mathbf{U}_{k,0} \\ \mathbf{U}_{k,0} \end{bmatrix}$$

Relation - MOESP

Using a LQ (QR)-decomposition one can bring any $\mathbf{W}_-$ into this structure, from which we have the matrix $\mathbf{H}_k$, and can apply realization. This approach is taken in MOESP

1. Using LQ to recover matrix $\mathcal{O}_k\mathcal{C}_k$
2. Use realization to recover $\mathbf{A}, \mathbf{B}$, and then $\mathbf{B}, \mathbf{D}$.
3. Then use Kalman filter to obtain corresponding state sequence.

Relation - N4SID

A different road:

- Recover the order and the state *subspace* by relating $\mathbf{W}_-$ to $\mathbf{W}_+$,
- Then recover $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ by LS.

How does that work?

Thm. $\mathrm{span}(\mathbf{W}_-) \cap \mathrm{span}(\mathbf{W}_+) = \mathrm{span}(\mathbf{X}_k)$, or

$$\mathbf{Y}_{k,0} = \mathcal{O}_k \mathbf{X}_0 + \Psi \mathbf{U}_{k,0}$$

So find the subspace by oblique projection (SVD).

$$\Pi_{\mathbf{U}}^{+} = I - \mathbf{U}^T (\mathbf{U}\mathbf{U}^T)^{-1} \mathbf{U}$$

Then $\mathbf{Y}_{k,0} \Pi_{\mathbf{U}}^{+} = \mathcal{O}_k \mathbf{X}_0 \Pi_{\mathbf{U}}^{+}$.

Stochastic Realization

Problem: Given $\mathbb{E}[Y_t Y_{t-\tau}^T] = \Lambda(\tau)$ for $\tau = 0, 1, 2, \ldots$, find a realization $(\mathbf{A}, \mathbf{B})$ such that the outcome $\{Y_t\}$ of the system

$$\begin{cases} X'_{t-1} = \mathbf{A} X'_t + \mathbf{K} D_t \\ Y_t = \mathbf{C} X'_t + D_t \end{cases}$$

when driven by white noise $\{D_t\}$ taking values in $\mathbb{R}^n$ has properties $\{\Lambda(\tau)\}_\tau$. Richer in history: Parzen, Akaike,Kalman, Faurre, De moor/Van Overschee, but Messier in results

Build up the data matrices $\mathbf{Y}_{k,0}$ and $\mathbf{Y}_{k,k}$, and use those to reconstruct the internal states. One common way to do that is using Canonical Correlation Analysis, solving

$$\max_{\mathbf{a},\mathbf{b}} \frac{\mathbf{a}^T \mathbf{Y}_{k,0} \mathbf{Y}_{k,k}^T \mathbf{b}}{\sqrt{\mathbf{a}^T \mathbf{Y}_{k,0} \mathbf{Y}_{k,0}^T \mathbf{a}} \sqrt{\mathbf{b}^T \mathbf{Y}_{k,k} \mathbf{Y}_{k,k}^T \mathbf{b}}}$$

- Solutions given by generalized eigenvalue problem.

- Detection of n by number of significant eigenvalues of $\Sigma_{--}^{-1/2}\Sigma_{-+}\Sigma_{++}^{-1/2}$ where

$$\begin{cases}\Sigma_{--} = \frac{1}{N}\mathbf{Y}_{k,0}\mathbf{Y}_{k,0}^T \\ \Sigma_{-+} = \frac{1}{N}\mathbf{Y}_{k,0}\mathbf{Y}_{k,k}^T \\ \Sigma_{--} = \frac{1}{N}\mathbf{Y}_{k,k}\mathbf{Y}_{k,k}^T\end{cases}$$

- Basis given by corresponding eigenvectors.

- Again, compute matrices $\mathcal{O}_k$ and $\mathcal{C}_k$, and realize a $(\mathbf{A}, \mathbf{C})$.

Combined Stochastic - Deterministic

System

$$\begin{cases} X_{t+1} = \mathbf{A} X_t + \mathbf{B}\mathbf{u}_t + V_t \\ Y_t = \mathbf{C} X_t + \mathbf{D}\mathbf{u}_t + W_t, \end{cases} \quad \forall t = -\infty, \ldots, \infty.$$

with

- $\{\mathbf{x}_t\}_t \subset \mathbb{R}^n$ the state process.
- $\{\mathbf{u}_t\}_t \subset \mathbb{R}^p$ the input process.
- $\{V_t\}_t \subset \mathbb{R}^n$ the process noise with covariance $\mathbf{R}$.
- $\{\mathbf{y}_t\}_t \subset \mathbb{R}^q$ the output process.
- $\{V_t\}_t \subset \mathbb{R}^n$ the measurement noise with covariance $\mathbf{Q}$.
- $\mathbf{A} \in \mathbb{R}^{n \times n}$ the system matrix.
- $\mathbf{B} \in \mathbb{R}^{n \times p}$ the input matrix.
- $\mathbf{C} \in \mathbb{R}^{q \times n}$ the output matrix.
- $\mathbf{D} \in \mathbb{R}^{q \times p}$ the feed-through matrix.

Problem: Given $\{\mathbf{u}_t\}_t \subset \mathbb{R}^p$ and $\{\mathbf{y}_t\}_t \subset \mathbb{R}^q$, find $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}, \mathbf{P}, \mathbf{Q})$ and $\{\mathbf{x}_t\}_t$.

Basic equation

$$\mathbf{Y}_{k,0} = \mathcal{O}_k \mathbf{X}_0 + \Psi \mathbf{U}_{k,0} + \mathbf{V}$$

- Razor away $\mathbf{U}$ by oblique projection.
- Razor away $\mathbf{V}$ using appropriate instruments.

Algorithm:

- Build data matrices.
- Estimate $\mathcal{O}_k$, or $\{\mathbf{x}_t\}_t$.
- Recover $(\hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}, \hat{\mathbf{D}})$.
- Estimate $\mathbf{P}, \mathbf{Q}$ from sample covariance of residuals.

Conclusions

- State-space systems for MIMO - distributed parameter systems.
- Relation impulse response - state-space models.
- Controllability - Observability
- Kalman - Ho
- Stochastic Systems
- Subspace as extended realization.
- SVD and LQ.
- Stochastic.
- Combined Deterministic - Stochastic.