# System Identification, Lecture 9

Kristiaan Pelckmans (IT/UU, 2338)

Course code: 1RT880, Report code: 61800 - Spring 2013
F, FRI Uppsala University, Information Technology

2 Mai 2013

# Project Works

What:

1. Identification of a Multimedia stream.
2. Identification of an industrial Petrochemical plant.
3. Identification of an Acoustic Impulse Response.
4. Identification of Financial Stock Markets.

Before Final Presentation:

1. Report Comp.lab. 5
2. Report Process Lab.
3. Report Project Work.

Expected:

1. *Visualize* the data, point out characterizing properties and state the problem you're after.

2. Do some simple (*SISO*) simulations: e.g. what is the best constant prediction (mean). This can often be done using the ident tool.

3. What is a proper method for identification of the system (*MIMO*) and perform the simulations. Most importantly, verify the result: why is this result satisfactory? How does it compare to the naive estimates of (2)?

4. *Describe* a full identification experiment: why is this (not) possible in practice? What would be the benefit if it were possible? What are further important todo's?

5. *Summarize* your contribution in an 'abstract' and 'conclusions' of your report. Those different steps (sections) should show up in the report to be handed in.

6. *Report*: A well-manicured report describing the achieved results, motivating the design decisions and verifying the estimated models. Make sure sufficient care is given to (a) Avoid Typos. (b) Use of the English language: think about what you write and how you write it up. (c) Be concise: reread your own text and throw out what is not needed for supporting the conclusions. (d) Figures: name axes, and give units. Add a legend explaining the curves we see, and add a caption explaining what we see and should conclude from the present figure. (e) A guideline would be a report of 3-4 single column, 11pt, letter pages.

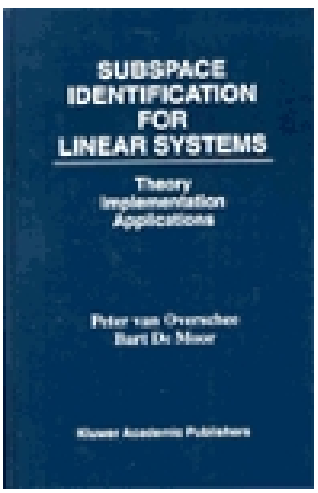
7. *Presentation*: each group is assigned a slot of 10 minutes to defend their results. Specifically, try to convince the audience of the following bullets: (a) What are the conclusions of the effort, and how do you get there?(b) How do you improve over earlier/simpler solutions? (c) What is the contribution of each of the groupmembers? (d) What are possible applications for your work? (e) Suppose I were your manager at a company: why should I invest 1000$ to implement your model? (f) Suppose I were your teacher: why would I award you a grade 5 for your work?

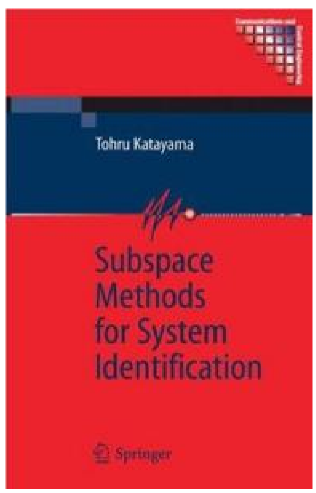
# Overview Subspace Identification

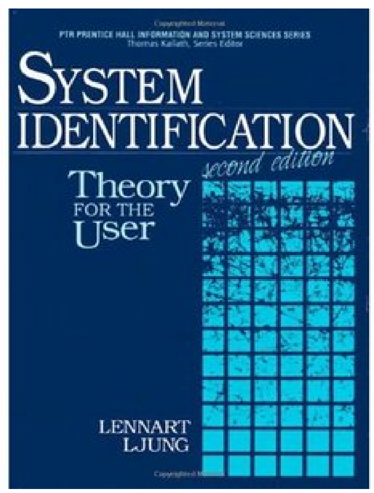
1. Deterministic.

2. Stochastic.

3. Extensions.





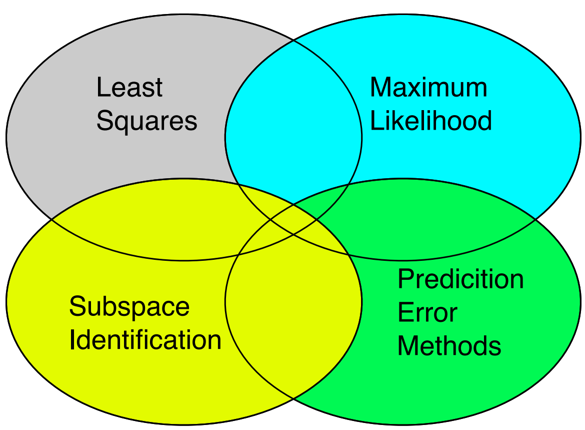



K. De Cock, B. De Moor, "Subspace Identification Methods", report, 2003.

# 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

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

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 $k, \ldots, k+i-1$

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

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

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

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 - ex.1.

Let

$$u = (0, 0, 0, 1, 0, 0, 0, \dots)^T$$

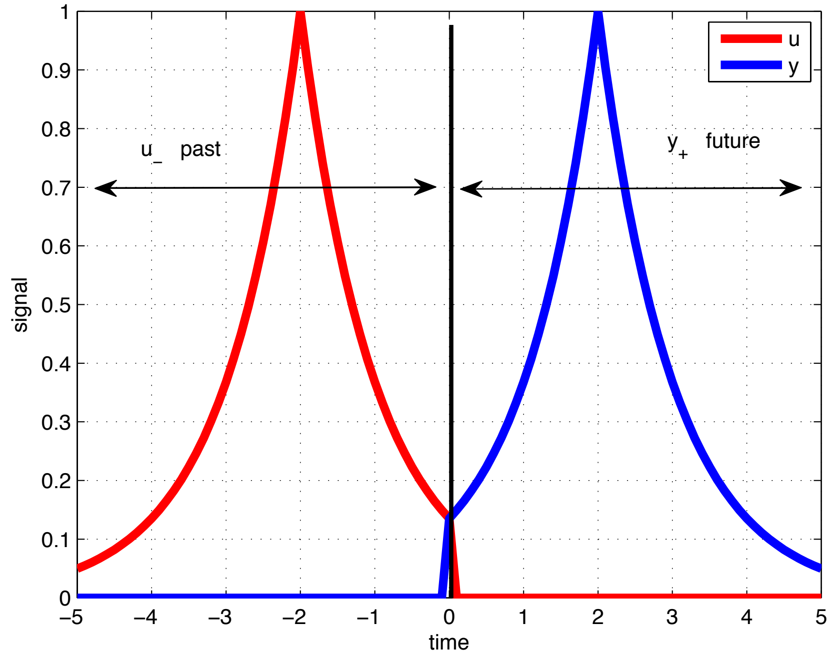
Apply this input signal to a noiseless LTI, and suppose the outcome is

$$y = (0, 0, 0, g_1, g_1, g_3, g_4, g_5, \dots)^T$$

Let $k = 4, i = 8$, then we get

$$\mathbf{W}_- = \left[\begin{array}{cccc|cccc} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & g_0 & g_1 & g_2 & g_3 & g_4 \\ 0 & 0 & g_0 & g_1 & g_2 & g_3 & g_4 & g_5 \\ 0 & g_0 & g_1 & g_2 & g_3 & g_4 & g_5 & g_6 \\ g_0 & g_1 & g_2 & g_3 & g_4 & g_5 & g_6 & g_7 \end{array}\right]$$

This datamatrix ressembles $\begin{bmatrix} I_4 & 0 \\ \psi_4 & \mathcal{O}_4\mathcal{C}_4 \end{bmatrix}$ (up to permutation).




This is a general structure, 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 $(\mathbf{A}, \mathbf{C})$.

# Extensions

- Innovation representation.
- Reduced Realization.
- Weighting matrices.
- Positive Real and Stable.
- Relation to Kalman filter.
- SVD and LQ are robust and efficient techniques.

# Combined Stochastic - Deterministic

System

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

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.

Extensions:

- Feedback - rank conditions.
- Bilinear (States $\times$ Inputs) and nonlinear systems (Hammerstein).
- Recursive.
- Selection of the order.
- Statistical analysis.
- Finite data.

# Conclusions

To remember

- Problem.
- Subspace as extended realization.
- SVD and LQ.
- Stochastic.
- Combined Deterministic - Stochastic.
- Optimality?