Computational Control

Identification

Saverio Bolognani

Automatic Control Laboratory (IfA)
ETH Zurich

Model requirements

The methods that we have seen so far, i.e.,

- LQR
- Model predictive control

require a state space model of the system.

$$x_{t+1} = Ax_t + Bu_t$$
$$y_t = Cx_t$$

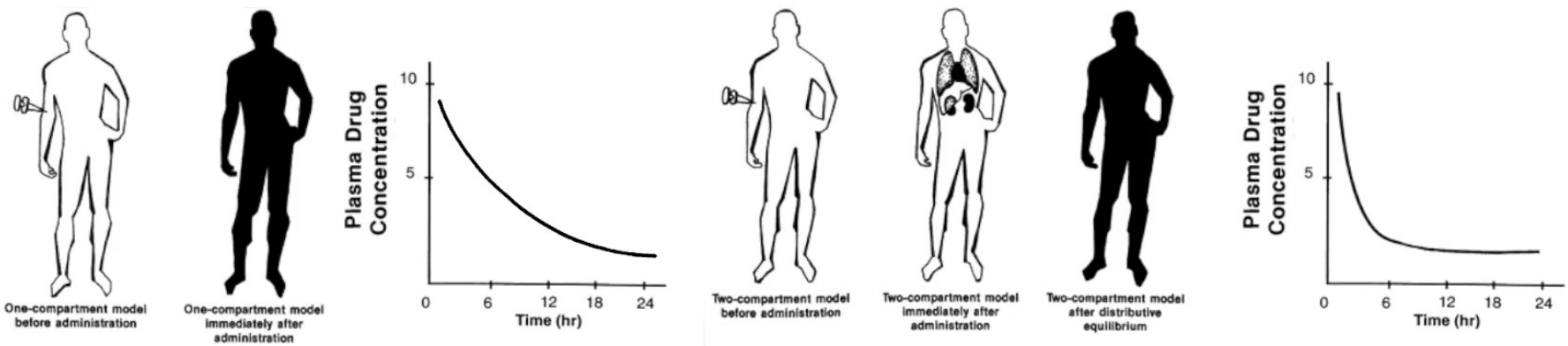
Knowing a state-space representation of a system implies a **very good knowledge** of the plant.

- What are the parameters of the update equation?
- What are the parameters of the input (and output) maps?
- What is the Markovian state?





- No first principles to select the state
- No idea of the dimension of the state
- Accurate modeling sometimes suggests a huge state dimensions
- Even if we can postulate a Markovian state, we cannot measure it



In fact, state-space representations are not even unique.

$$x_{t+1} = Ax_t + Bu_t$$
$$y_t = Cx_t$$

Consider an invertible square matrix T and the system

$$w_{t+1} = T^{-1}ATw_t + T^{-1}Bu_t$$
$$y_t = CTw_t$$

The two systems have exactly the same **input-output behavior**.

Proof: Because of linearity, we only need to check that the response to each input $u_t = e_i\delta_t$ is the same, where

$$\delta_t = \begin{cases} 1 & \text{if } t = 0 \\ 0 & \text{otherwise} \end{cases}$$

In the first case $x_t = 0, b_i, Ab_i, A^2b_i, \ldots$ and $y_k = 0, Cb_i, CAb_i, CA^2b_i, \ldots$.
In the second case $w_t = 0, T^{-1}b_i, T^{-1}Ab_i, T^{-1}A^2b_i, \ldots$ and
$y_k = 0, Cb_i, CAb_i, CA^2b_i, \ldots$.

Impulse response

$$y_t = \underbrace{CA^t x_0}_{\text{initial response}} + \underbrace{\sum_{j=0}^{t-1} CA^{t-1-j} B u_j}_{\text{convolution}}$$

SISO system

Scalar input $u_t \in \mathbb{R}$, scalar output $y_t \in \mathbb{R}$.

Impulse response

$$y_t = CA^{t-1}B \quad t > 0$$

- Are you convinced that the impulse-response is a complete representation of the input-output relation of the system?
- How would you estimate is on a real plant?

Markov parameters

Generalization to MIMO systems

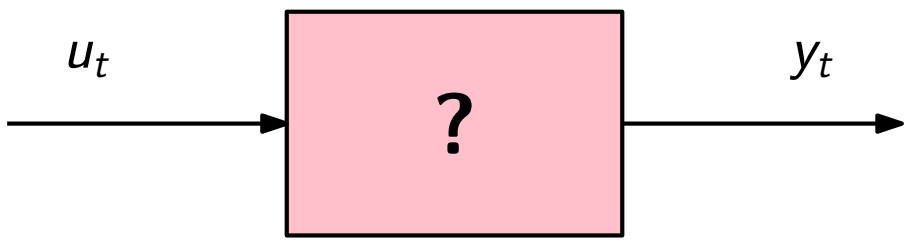
m inputs ($u \in \mathbb{R}^m$) and p outputs ($y \in \mathbb{R}^p$)

$$CA^{t-1}B \in \mathbb{R}^{p \times m}, \quad t > 0$$

The elements in position i, j represent the output i to an unit impulse on the input j.

This sequence of matrices are the **Markov parameters** of the system.

How can you estimate these parameters? With how many experiments?



Controllability and Observability

Controllability

Consider the problem of driving a system from $x_0 = 0$ to a desired $\bar{x} \in \mathbb{R}^n$.

- In **one step**, we can drive the system to

$$x_1 = Bu_0$$

  therefore to any state in the **image of** B.

- In **two steps**, we can drive the system to

$$x_2 = ABu_0 + Bu_1$$

  therefore to any state in the **image of** $\begin{bmatrix} B & AB \end{bmatrix}$.

- ...

- In n **steps**, we can drive the system to any state in the image of

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

  which we call the **controllability matrix** of the system.

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

If the rank of $\mathcal{C}$ is n (**full row rank**) then any state $\bar{x} \in \mathbb{R}^n$ can be reached in n steps.

Remarks

- (obvious) You may be able to control the system to any state in less than n steps
  - Also with a scalar input?
- (non obvious) If you don't have controllability in n steps, adding additional steps will not help.

Given a target state $\bar{x}$, how can you compute the input sequence that drives a controllable system to $\bar{x}$?

If $\mathcal{C}$ is an invertible matrix (which means a scalar input u) then $x_n = \mathcal{C} \begin{bmatrix} u_{n-1} \\ \vdots \\ u_0 \end{bmatrix}$ and this relation can be inverted to compute the input sequence.
In case of non-invertible $\mathcal{C}$, the problem is underdetermined (there are more valid input sequences). We can find one (the one with smallest norm) by using the pseudoinverse $\mathcal{C}^\top(\mathcal{C}\mathcal{C}^\top)^{-1}$ of the matrix $\mathcal{C}$.

Observability

Consider the problem of detecting the effect of an initial condition x_0 in the output of the system.

- At $t = 0$, the output of the system is

$$y_0 = Cx_0$$

  therefore any x_0 in the **kernel of C** is unobservable.

- In **one step**, the output is

$$y_1 = CAx_0$$

  therefore any x_0 in the **kernel of** $\begin{bmatrix} C \\ CA \end{bmatrix}$ is unobservable.

- ...After n steps, any initial state in the kernel of the **observability matrix**

$$\mathcal{O} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

  is unobservable.

$$\mathcal{O} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

If the rank of $\mathcal{O}$ is n (**full column rank**) then any initial state $x_0 \in \mathbb{R}^n$ is observable in the output.

Remarks

- (obvious) You may be able to see the effect of x_0 on the output in less than n steps
  - Also with a scalar output?
- (non obvious) If you don't have observability in n steps, adding additional steps will not help.

Given a desired output sequence $y_0, y_1, \ldots, y_{n-1}$, what is the initial state x_0 that produces that sequence?

If $\mathcal{O}$ is an invertible matrix (which means a scalar output y) then $\begin{bmatrix} y_0 \\ \vdots \\ y_{n-1} \end{bmatrix} = \mathcal{O}x_0$

and this relation can be inverted to compute the initial state.
In case of non-invertible $\mathcal{O}$ we can employ the pseudoinverse as before.

Controllable and observable systems

Remember the goal

We want derive a state-space representation (**realization**) when the only thing we know are input and output data.

$\rightarrow$ We restrict our attention on systems that are **controllable** and **observable**.

- "well-posed" relation between state and input/output signals
- exclude the pathological case of extra states that have no effect on the I/O
- aim at **minimal** realizations

Example: Consider a plant whose output is simply a moving average of the input

$$y_t = a_1 u_{t-1} + a_2 u_{t-2} + a_3 u_{t-3}$$

What is a minimal state-space representation? Is it observable and controllable?

We use the state to store a delayed copy of the input:

$$x_t = \begin{bmatrix} u_{t-1} \\ u_{t-2} \\ u_{t-3} \end{bmatrix}$$

This is easily accomplished via the state-space model

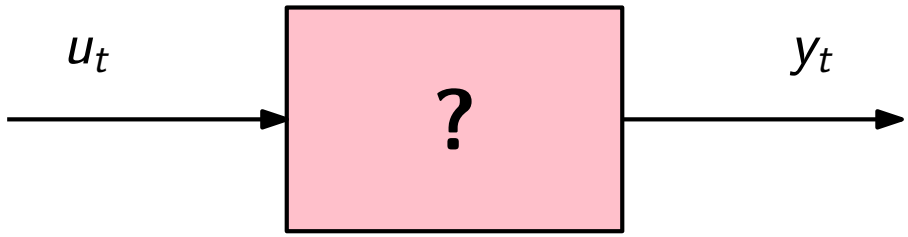
$$x_{t+1} = \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}}_{A} x_t + \underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}}_{B} u_t$$

which we then combine with the output equation

$$y_t = \underbrace{\begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix}}_{C} x_t$$

The controllability matrix is $\mathcal{C} = I$ and the observability matrix is $\mathcal{O} = \begin{bmatrix} a_1 & a_2 & a_3 \\ a_2 & a_3 & 0 \\ a_3 & 0 & 0 \end{bmatrix}$.

Identification from data



Suppose that you only have the Markov parameters of the system.

- For example: sequence of impulse-response experiments on the real plant.

How to compute A, B, C (realization)?

More general system ID

Given generic input-output data, construct a realization A, B, C.

Hankel matrix

$$\text{Hankel matrix } \mathcal{H} := \mathcal{OC} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix} \begin{bmatrix} B & AB & A^2B & \dots & A^{n-1}B \end{bmatrix}$$

$$= \begin{bmatrix} CB & CAB & CA^2B & \dots & CA^{n-1}B \\ CAB & CA^2B & CA^3B & \dots & CA^nB \\ CA^2B & CA^3B & \dots & & \dots \\ \dots & & & & \dots \\ CA^{n-1}B & \dots & & \dots & CA^{2n-2}B \end{bmatrix}$$

Remember the **Markov parameters**

$$\{ CB, \quad CAB, \quad CA^2B, \quad \dots, \quad CA^{t-1}B, \quad \dots \}$$

Ho-Kalman algorithm

Not the Kalman filter!

1. Construct the Hankel matrix $\mathcal{H}$ from the Markov parameters (experiments)
2. Factorize $\mathcal{H}$ as $\mathcal{O}\mathcal{C}$
3. Derive realization A, B, C from $\mathcal{O}$ and $\mathcal{C}$

Step 1: Construct $\mathcal{H}$

Problem: how many Markov parameters do we need? We don't know n!

We are looking for a realization that is minimal (controllable and observable).
$\rightarrow$ Both $\mathcal{O}$ and $\mathcal{C}$ will have rank n, so the Hankel matrix has rank n.

When do you stop adding Markov parameters? **Try!**

Low-pass filter

$$x_{t+1} = ax_t + u_t$$
$$y_t = x_t$$

Frictionless cart

$$v_{t+1} = v_t + u_t$$
$$z_{t+1} = z_t + v_t$$
$$y_t = z_t$$

Step 2: How do you factorize $\mathcal{H}$ in $\mathcal{OC}$?

Interestingly, it does not matter!

How is that possible? Different $\mathcal{O}$ and $\mathcal{C}$ must return different A, B, C!

Consider the change of basis: $\tilde{A} = T^{-1}AT$, $\tilde{B} = T^{-1}B$, $\tilde{C} = CT$.

$$\tilde{\mathcal{O}} = \begin{bmatrix} \tilde{C} \\ \tilde{C}\tilde{A} \\ \vdots \end{bmatrix} = \begin{bmatrix} CT \\ CTT^{-1}AT \\ \vdots \end{bmatrix} = \mathcal{O}T$$

Similarly one can show that $\tilde{\mathcal{C}} = T^{-1}\mathcal{C}$, so that $\tilde{\mathcal{O}}\tilde{\mathcal{C}} = \mathcal{O}TT^{-1}\mathcal{C} = \mathcal{OC}$.

$\rightarrow$ different factorizations return you the same system in different coordinates!

- You can take $\mathcal{O} = I$ and $\mathcal{C} = \mathcal{H}$
- You can take $\mathcal{O} = \mathcal{H}$ and $\mathcal{C} = I$ (do you remember?)
- There are preferrable ways to do it (for numerical reasons)

Singular value decomposition

Let H be a $m \times n$ matrix of rank r.

$$H = U\Sigma V^\top$$

- U is a $m \times m$ matrix
- V is a $n \times n$ matrix
- Σ is a "diagonal" $m \times n$ matrix

$$H = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0_{r \times n-r} \\ 0_{m-r \times r} & 0_{m-r \times n-r} \end{bmatrix} \begin{bmatrix} V_1 & V_2 \end{bmatrix}^\top = U_1 \Sigma_1 V_1^\top$$

Your linear algebra workhorse!

- computationally very fast and numerically robust
- U_1 columns span the image of H
- the dimension of Σ_1 is $r \times r$, where r is the rank of H
- factorizes a matrix is a numerically "balanced" way (orthonormal U, V)

Do you see how we could use this?

Extended Hankel matrix

Let us build a Hankel matrix of size $k \times \ell$, where k and ℓ are larger than n.

$$\mathcal{H}_{k,\ell} = \begin{bmatrix} G_1 & G_2 & \dots & G_\ell \\ G_2 & G_3 & \dots & G_{\ell+1} \\ G_3 & \dots & & \\ \dots & & & \dots \\ G_k & G_{k+1} & \dots & G_{k+\ell-1} \end{bmatrix}$$

How do you pick k, ℓ larger than n?

If these are the Markov parameters of an A, B, C system, then $\mathcal{H}_{k,\ell}$ is the product between

$$\mathcal{O}_k = \begin{bmatrix} C \\ CA \\ CA^2 \\ \dots \\ CA^{k-1} \end{bmatrix} \quad \text{and} \quad \mathcal{C}_\ell = \begin{bmatrix} B & AB & A^2B & \dots & A^{\ell-1}B \end{bmatrix}$$

called **extended** observability/controllability matrices (what is their rank?).

SVD of the Hankel matrix

Let's compute the SVD of $\mathcal{H}_{k,\ell}$.

$$\mathcal{H}_{k,\ell} = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0_{n\times \ell-n} \\ 0_{k-n\times n} & 0_{k-n\times \ell-n} \end{bmatrix} \begin{bmatrix} V_1 & V_2 \end{bmatrix}^\top = U_1 \Sigma_1 V_1^\top$$

$\rightarrow$ dimension of the system n from Σ_1.

Decompose $\mathcal{H}_{k,\ell}$ as

$$\mathcal{H}_{k,\ell} = \underbrace{U_1 \Sigma_1^{1/2}}_{\mathcal{O}_k} \underbrace{\Sigma_1^{1/2} V_1^\top}_{\mathcal{C}_\ell}$$

to obtain C and B

$$\mathcal{O}_k = \begin{bmatrix} C \\ CA \\ CA^2 \\ \dots \\ CA^{k-1} \end{bmatrix} \quad \text{and} \quad \mathcal{C}_\ell = \begin{bmatrix} B & AB & A^2B & \dots & A^{\ell-1}B \end{bmatrix}$$

How to compute A?

Consider the **shifted Hankel matrix**

$$\mathcal{H}_{k,\ell} = \begin{bmatrix} G_1 & G_2 & \dots & G_\ell \\ G_2 & G_3 & \dots & G_{\ell+1} \\ G_3 & \dots & & \\ \dots & & & \dots \\ G_k & G_{k+1} & \dots & G_{k+\ell-1} \end{bmatrix} \qquad \mathcal{H}^{\uparrow}_{k,\ell} = \begin{bmatrix} G_2 & G_3 & \dots & G_{\ell+1} \\ G_3 & G_4 & \dots & G_{\ell+2} \\ G_4 & \dots & & \\ \dots & & & \dots \\ G_{k+1} & G_{k+2} & \dots & G_{k+\ell} \end{bmatrix}$$

And notice that

$$\mathcal{H}^{\uparrow}_{k,\ell} = \begin{bmatrix} CAB & CA^2B & CA^3B & \dots \\ CA^2B & CA^3B & CA^4B & \dots \\ CA^3B & CA^4B & \dots & \\ \dots & \dots & & \end{bmatrix} = \underbrace{\begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \end{bmatrix}}_{\mathcal{O}_k} A \underbrace{\begin{bmatrix} B & AB & A^2B & \dots \end{bmatrix}}_{\mathcal{C}_\ell}$$

$$\rightarrow \quad A = \mathcal{O}_k^{\dagger} \mathcal{H}^{\uparrow}_{k,\ell} \mathcal{C}_\ell^{\dagger}$$

Kalman-Ho algorithm

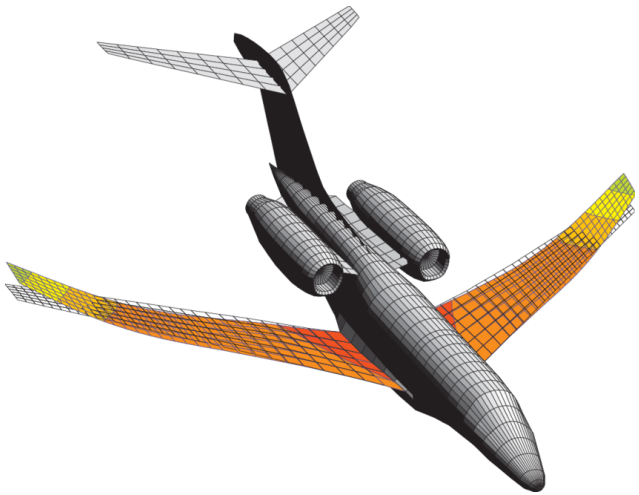
1. Collect impulse reponses of the system as Markov parameters
2. Estimate $k, \ell > n$
3. Build Hankel matrix $\mathcal{H}_{k,\ell}$ and shifted Hankel matrix $\mathcal{H}^{\uparrow}_{k,\ell}$
4. Compute the SVD of $\mathcal{H}_{k,\ell}$ to
   - determine n
   - compute extended observability and controllability matrices
5. Derive A, B, C.

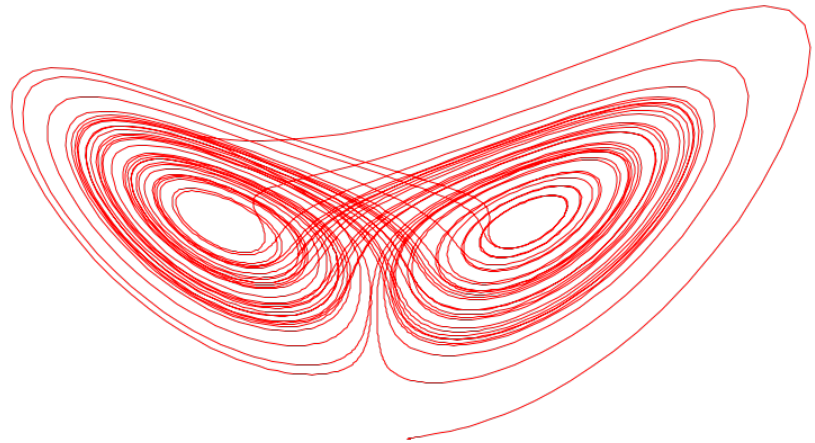
Take home message

- When a state-space system model is not available, it can be identified.
- Data contains all the model information (data = model?)
- The control design methods that we have seen can be applied to the identified model.

Underlying assumptions today

- A real LTI finite-dimensional system exists





- There is no noise



- This was just a sample of **system identification** (nonlinear systems, model reduction, stochastic data, missing data, priors, parametric/non-parametric, and much more!)



This work is licensed under a
Creative Commons Attribution-NonCommercial 4.0 International License

https://bsaver.io/COCO