When is a pole spurious?

I. Markovsky

joint work with
J. Boets, B. Vanluyten, K. De Cock, B. De Moor

K.U.Leuven, ESAT-SISTA

Outline

Modal analysis $\leftrightarrow$ System identification

modal analysis: given data from a mechanical structure, find eigen frequencies and dampings

linear system identification: given observed data, find LTI system that fits that data

the frequencies and dampings are parameters of an LTI system

$$\Downarrow$$

modal analysis is a special linear system identification problem

operational modal analysis $\leftrightarrow$ ARMA or autonomous SYSID

Detect physical poles $\leftrightarrow$ Choose model order

important questions

- in modal analysis: detect physical poles
- in system identification: choose the model order

tools used to answer these questions

- in modal analysis: stabilization diagram
- in system identification:
  - balanced model reduction
  - decay of singular values (in subspace identification)
  - Kumaresan–Tufts's method
  - Akaike's information criterion
  - Cross-validation
  - . . .

Our aim

give a definition of a physical/spurious pole

show that

"detecting physical poles" and "choosing model order"

are equivalent problems

compare the methods

- stabilization diagram
- balanced model reduction
- subspace identification
- Kumaresan–Tufts's

on simulation example

Representations and parameters of LTI systems

$\mathscr{L}$ — discrete-time LTI model class

σ — shift operator $(\sigma w)(t) := w(t+1)$

any $\mathscr{B} \in \mathscr{L}$ can be represented as a solution set of a constant coefficients difference equation (kernel representation)

$$R_0\sigma^0 w + R_1\sigma^1 w + \cdots + R_\mathtt{l}\sigma^\mathtt{l} w = 0 \qquad \text{(DE)}$$

or a first order system of equations (input/state/output repr.)

$$w = \Pi\,\mathrm{col}(u,y), \quad \sigma x = Ax + Bu, \quad y = Cx + Du, \qquad \text{(SS)}$$

where Π is a permutation matrix

the polynomial matrix $R(\xi) := R_0\xi^0 + R_1\xi^1 + \cdots + R_\mathtt{l}\xi^\mathtt{l}$ and the matrices (A,B,C,D,Π) are parameters of $\mathscr{B}$

Representations and parameters of LTI systems

notation:

$$\mathscr{B}(R) := \{\, w \in (\mathbb{R}^\mathtt{w})^\mathbb{N} \mid \text{(DE) holds} \,\}$$

$$\mathscr{B}_{\mathrm{i/s/o}}(A,B,C,D,\Pi) := \{\, w \in (\mathbb{R}^\mathtt{w})^\mathbb{N} \mid \text{(SS) holds} \,\}$$

$\mathtt{l}$ — lag of (DE)

$\mathtt{n}$ — order of (SS)

given $\mathscr{B} \in \mathscr{L}$, the smallest lag $\mathbf{l}(\mathscr{B})$ and order $\mathbf{n}(\mathscr{B})$ are independent of the representations

$\Downarrow$

minimal kernel and state space representations

Autonomous systems

$\mathscr{L}_0$ — model class of autonomous systems

kernel representation

$$P_0\sigma^0 y + P_1\sigma^1 y + \cdots + P_\mathtt{l}\sigma^\mathtt{l} y = 0, \qquad \det\big(P(\xi)\big) \neq 0 \quad \text{(DE AUT)}$$

state space representation

$$\sigma x = Ax, \quad y = Cx \qquad \text{(SS AUT)}$$

$\mathscr{B}(P)$, $\mathscr{B}(A,C)$ — systems defined by (DE AUT) and (SS AUT)

$\lambda(\mathscr{B})$ — the poles of $\mathscr{B}$ (zeros of $\det(P(\xi)) = 0$)

if $\mathscr{B}(A,C)$ is minimal, the set of eigenvalues $\lambda(A) \equiv \lambda(\mathscr{B})$

"Physical" poles $\leftrightarrow$ "True" data generating system

Consider given: y_d — data and $\mathscr{M}$ — model class

the intuition behind "physical" pole is:

$$\boxed{\begin{array}{c}\lambda \text{ is a physical pole for } y_\mathrm{d} \\ \Updownarrow \\ \lambda \text{ is a pole of a true data generating system } \bar{\mathscr{B}} \in \mathscr{M} \text{ for } y_\mathrm{d}\end{array}}$$

The question "when is a pole spurious" is actually the question:

What does it mean "$\bar{\mathscr{B}} \in \mathscr{M}$ is a data generating system for y_d"?

Classical approaches to define "true" system

let $\mathscr{M}$ be the class of autonomous LTI systems $\mathscr{L}_0$

output error: y_d is a noise corrupted trajectory of $\bar{\mathscr{B}}$

$$y_\mathrm{d} = \bar{y} + \tilde{y}, \quad \text{where} \quad \bar{y} \in \bar{\mathscr{B}} \in \mathscr{L}_0 \quad \text{and} \quad \tilde{y} \sim \mathrm{N}(0, \sigma^2 I) \quad \text{(OE)}$$

ARMA: $\operatorname{col}(y_\mathrm{d}, e_\mathrm{d}) \in \bar{\mathscr{B}}$, where $e_\mathrm{d} \sim \mathrm{N}(0, \sigma^2 I)$

the classical approaches are stochastic, they

- assume that $\bar{\mathscr{B}}$ exists in the model class, but
- even if $\bar{\mathscr{B}} \in \mathscr{M}$ exists, it is not computable from y_d

Alternative deterministic approach

Definition (Most powerful unfalsified model (MPUM))

$\mathscr{B}_\mathrm{mpum}(y_\mathrm{d})$ is the MPUM of $y_\mathrm{d} \in (\mathbb{R}^\mathrm{p})^T$ in the model class $\mathscr{L}^\mathrm{p}$ if

1. $\mathscr{B}_\mathrm{mpum}(y_\mathrm{d})$ is unfalsified by y_d, *i.e.*, $y_\mathrm{d} \in \mathscr{B}_\mathrm{mpum}(y_\mathrm{d})$,
2. $\mathscr{B}_\mathrm{mpum}(y_\mathrm{d})$ is in the model class, *i.e.*, $\mathscr{B}_\mathrm{mpum}(y_\mathrm{d}) \in \mathscr{L}^\mathrm{p}$, and
3. any other unfalsified model in $\mathscr{L}^\mathrm{p}$ is less powerful, *i.e.*,

$$y_\mathrm{d} \in \mathscr{B} \in \mathscr{L}^\mathrm{p} \implies \mathscr{B}_\mathrm{mpum}(y_\mathrm{d}) \subseteq \mathscr{B}.$$

Notes:

- the MPUM exists and is unique (no prior assumptions)
- there are algorithms for computing it from y_d
- $\mathscr{B}_\mathrm{mpum}(y_\mathrm{d})$ is an exact model for y_d

Alternative deterministic approach

Definition (Physical poles)

The poles $\lambda\big(\mathscr{B}_\mathrm{mpum}(y_\mathrm{d})\big)$ of the MPUM of $y_\mathrm{d} \in (\mathbb{R}^\mathrm{p})^T$ are called physical poles w.r.t. the data y_d. Any $z \in \mathbb{C}$, such that $z \notin \lambda\big(\mathscr{B}_\mathrm{mpum}(y_\mathrm{d})\big)$, is called a spurious pole.

$\Longrightarrow$ by computing the MPUM, we compute the physical poles

question:

Assuming there is a "true" data generating system $\bar{\mathscr{B}}$ for y_d, i.e., $y_\mathrm{d} \in \bar{\mathscr{B}}$, under want conditions $\mathscr{B}_\mathrm{mpum}(y_\mathrm{d})$ coincides with $\bar{\mathscr{B}}$?

Recovering the data generating system

Theorem (Identifiability)

If $\mathscr{B} \in \mathscr{L}_{0,\mathtt{l}}^{\mathtt{p},\mathtt{n}}$ and $y_{\mathrm{d}} \in (\mathbb{R}^{\mathtt{p}})^T$ satisfy the following conditions:

1. *y_{d} is an exact trajectory of $\mathscr{B}$,* i.e., *$y_{\mathrm{d}} \in \mathscr{B}|_{[1,T]}$, and*
2. *y_{d} is persistently exciting of order $\mathbf{n}(\mathscr{B})$,*

then $\mathscr{B} = \mathscr{B}_{\mathrm{mpum}}(y_{\mathrm{d}})$.

condition 2 is restrictive for practical applications
$\rightsquigarrow$ need of approximation

Simulation setup

exact data generating system $\bar{\mathscr{B}} \in \mathscr{L}_{0,4}^{1}$ with physical poles

$$\lambda(\bar{\mathscr{B}}) = \{0.8556 \pm 0.4674j, 0.8980 \pm 0.3797j\}$$

exact trajectory $y_{\mathrm{d}} \in \mathbb{R}^{250}$, persistently exciting of order 4

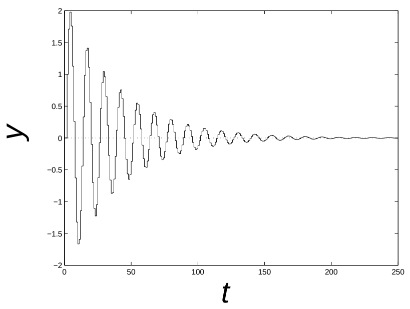
the MPUM of y_{d} is $\mathscr{B}(P)$, where

$$P(\xi) = 0.1278\xi^0 - 0.4716\xi^1 + 0.7037\xi^2 - 0.4961\xi^3 + 0.1414\xi^4$$

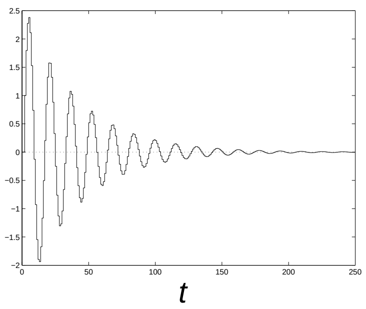
$$\mathscr{B}(P) = \bar{\mathscr{B}} \quad \text{and} \quad \lambda\big(\mathscr{B}(P)\big) = \lambda(\bar{\mathscr{B}})$$

Physical modes in the example

mode $0.8556 \pm 0.4674j$



mode $0.8980 \pm 0.3797j$



Noisy data

perturb the exact data y_{d} by a zero mean white Gaussian noise

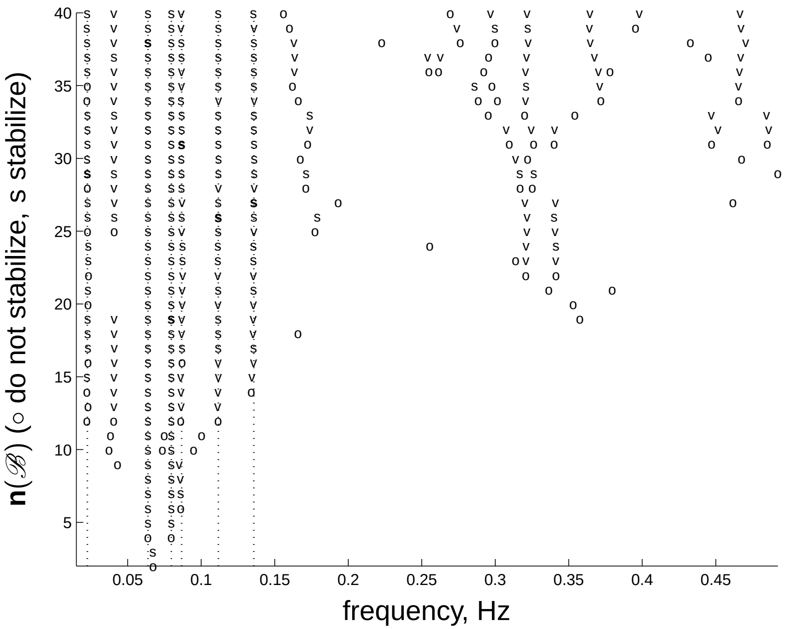
consider a bounded complexity model class $\mathscr{L}_{0,\mathtt{l}}$

for given $\mathtt{l}$, we can find an approximation $\widehat{\mathscr{B}} \in \mathscr{L}_{0,\mathtt{l}}$ of $\mathscr{B}_{\mathrm{mpum}}(y_{\mathrm{d}})$

Main question: How to find the correct $\mathtt{l}$? Compared methods:

1. stabilization diagram
2. method of Kumaresan and Tufts
3. singular value analysis in Kung's algorithm
4. principal angle analysis in subspace identification
5. trade-off curve, used with a maximum likelihood method

Stabilization diagram



Method of Kumaresan and Tufts

assuming exact data, specifying a higher order $L > \mathtt{l}$ implies

$$\lambda\big(\mathscr{B}(\widehat{P})\big) = \lambda\big(\mathscr{B}(P)\big) \cup \lambda\big(\mathscr{B}(S)\big)$$

where P, degree$(P) = L - \mathtt{l} - 1$ is arbitrary

$\Longrightarrow$ $\widehat{\mathscr{B}}$ contains all physical poles and $L - \mathtt{l} - 1$ spurious poles

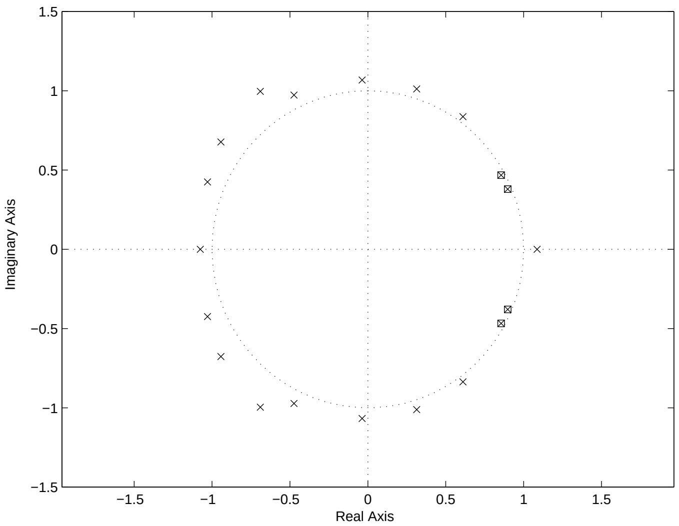
S depends on the identification method

Kumaresan–Tufts method — force the spurious poles to be outside the unit circle

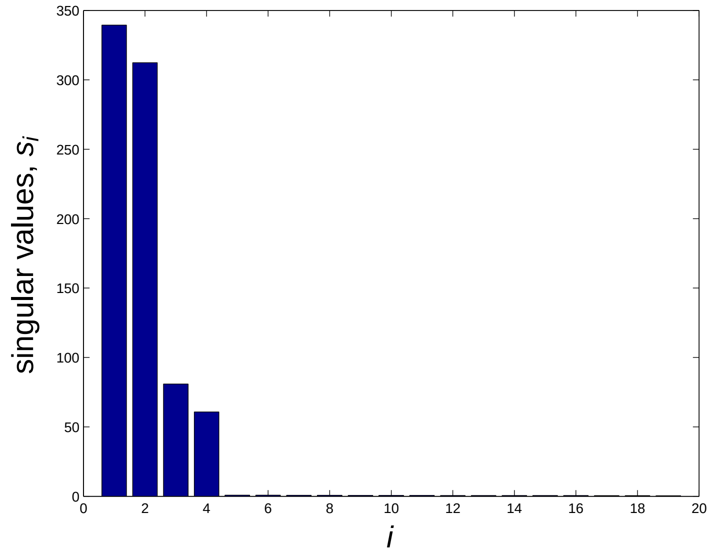
$\Longrightarrow$ the stable physical poles can be extracted

Method of Kumaresan and Tufts $L = 20$

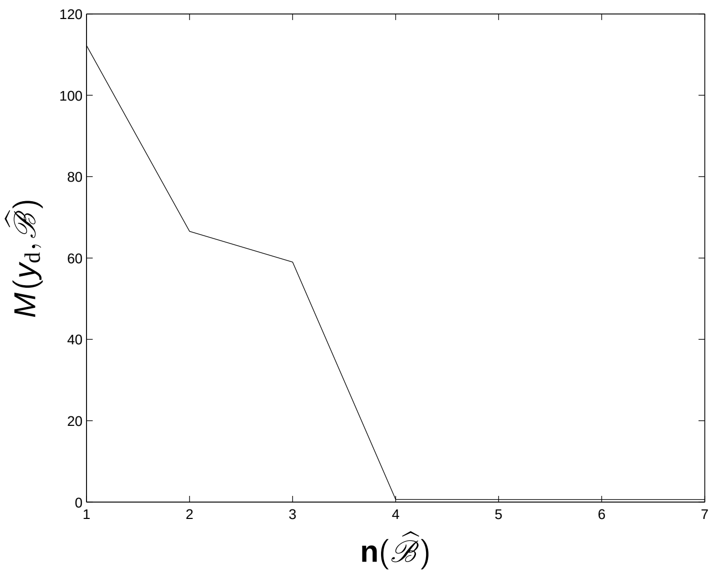
$\times$ — poles of $\widehat{\mathscr{B}}$, $\square$ — physical poles



Singular value analysis in Kung's algorithm



Misfit–complexity trade-off



Conclusions

- detecting spurious poles $\leftrightarrow$ order estimation
- can be done in many ways
  (one of which is the stabilization diagram)
- Kumaresan–Tufts', Kung's, and subspace methods
  1. compute a model of complexity higher than intended
  2. perform model reduction

  they have tests that indicate an appropriate low order

Thank you