Identifiability of time-lag parameters for nonlinear delay systems with applications in systems biology

Milena Anguelova, Bernt Wennberg

Chalmers University of Technology and Gothenburg University

Motivation

A model of the JAK-STAT signalling pathway:

$$
\begin{aligned}
\dot{x}_1 &= -k_1 x_1 u/k_7 + 2k_4 x_3(t-\tau) \\
\dot{x}_2 &= k_1 x_1 u/k_7 - k_2 x_2^2 \\
\dot{x}_3 &= -k_3 x_3 + 0.5 k_2 x_2^2 \\
\dot{x}_4 &= k_3 x_3 - k_4 x_3(t-\tau) \\
y_1 &= k_5(x_2 + 2x_3) \\
y_2 &= k_6(x_1 + x_2 + 2x_3)
\end{aligned}
$$

Timmer, J., Müller, T. G., Swameye, I., Sandra, O., & Klingmüller, U. (2004). Modelling the nonlinear dynamics of cellular signal transduction. *International Journal of Bifurcation and Chaos, 14*, 2069-2079.

Organisation

- The property of identifiability
- Simple examples of delay systems with single constant time-delays
- Linear-algebraic criteria for identifiability of the delay parameter
- Examples of biological systems
- Several time-delays

The property of identifiability

- Guarantees that the model parameters can be determined uniquely from the available data
- A prerequisite for parameter estimation
- Well-characterised for ODE systems, algorithms exist
- Recently extended to systems of delay differential equations with known time-delays
- Systems with unknown τ?

Identifiability for ODE

Example:

$$\begin{aligned}\dot{x}_1 &= \frac{x_2}{x_1}\\ \dot{x}_2 &= \frac{x_3}{x_2}\\ \dot{x}_3 &= x_1\theta - u\\ y &= x_1\end{aligned}$$

$$\begin{aligned}\dot{y} &= \frac{x_2}{x_1}\\ \ddot{y} &= \frac{x_3}{x_1x_2} - \frac{x_2^2}{x_1^3}\\ y^{(3)} &= \frac{\theta}{x_2} - \frac{u}{x_1x_2} - \frac{x_3^2}{x_1x_2^3} - \frac{3x_3}{x_1^3} + \frac{3x_2^3}{x_1^5}\end{aligned}$$

Sedoglavic, A. (2002). A probabilistic algorithm to test local algebraic observability in polynomial time.

Journal of Symbolic Computation, 33, 735-755.

Explicit relations for the variables/parameters

$$\begin{cases} x_1 & = & y \\ x_2 & = & y\dot{y} \\ x_3 & = & y\dot{y}(\dot{y}^2 + y\ddot{y}) \\ \theta & = & \frac{1}{y}((\dot{y}^2 + y\ddot{y})^2 + y\ddot{y}(3\dot{y}\ddot{y} + yy^{(3)}) - u) \end{cases}$$

The state variables are observable and the parameters identifiable.

Simple example of a delay system, *Example 1:*

$$\begin{cases} \dot{x}_1 &= k_1 x_2(t-\tau) + u(t) \\ \dot{x}_2 &= k_2 x_2(t-\tau) \\ y &= x_1 \\ x(t) &= \varphi(t), \quad t \in [-\tau, 0] \\ u(t) &= u_0(t), \quad t \in [-T, 0] \end{cases}$$

- two state variables, x_1 and x_2 with unknown initial conditions $\varphi(t)$
- two regular parameters k_1 and k_2
- an unknown time-lag parameter τ, $\tau \in [0, T)$, T known
- a controlled input variable u with initial conditions $u_0(t)$
- measured data y

Example 1 continued

Time derivatives of the output at a given point in time produce equations for the state variables and parameters:

$$\begin{aligned} y(t) &= x_1(t) \\ \dot{y}(t) &= k_1 x_2(t-\tau) + u(t) \\ \ddot{y}(t) &= k_1 k_2 x_2(t-2\tau) + \dot{u}(t) \\ y^{(3)}(t) &= k_1 k_2^2 x_2(t-3\tau) + \ddot{u}(t) \end{aligned}$$

Can decide the identifiability of the regular model parameters (Zhang et al., Xia et al.), if τ is known.

Zhang, J., Xia, X., & Moog, C. H. (2006). Parameter identifiability of nonlinear systems with time-delay. *IEEE T. Aut. Contr., 47*, 371-375 and the references therein.

$k_2 = (\ddot{y}(t) - \dot{u}(t))/(\dot{y}(t-\tau) - u(t-\tau))$, k_1 and x_2 unidentifiable/unobservable

Input-output equation

From the above equations, we can extract an external input-output representation of the system, given by the input-output equation

$$(\ddot{y}(t) - \dot{u}(t))(\ddot{y}(t-\tau) - \dot{u}(t-\tau)) - \\ -(y^{(3)} - \ddot{u})(\dot{y}(t-\tau) - u(t-\tau)) = 0$$

- We can calculate the time-lag parameter τ by (numerically) finding the zeros of the function $\xi_{t_0}(\tau)$

$$\xi_{t_0}(\tau) = (\ddot{y}(t_0) - \dot{u}(t_0))(\ddot{y}(t_0-\tau) - \dot{u}(t_0-\tau)) - \\ -(y^{(3)}(t_0) - \ddot{u}(t_0))(\dot{y}(t_0-\tau) - u(t_0-\tau))$$

Estimating τ

- We have used the `dde23.m` differential equation solver in `Matlab` to simulate an output for the above system by choosing
  $k_1 = -2, k_2 = -3$
  $\varphi_1(t) = t + 1, \varphi_2(t) = t^2 + 1$
  $u(t) = t$ and $\tau = 1$
- Plotted the function

$$\begin{aligned} \xi_6(\tau) = \quad & (\ddot{y}(6) - \dot{u}(6))(\ddot{y}(6-\tau) - \dot{u}(6-\tau)) - \\ & -(y^{(3)}(6) - \ddot{u}(6))(\dot{y}(6-\tau) - u(6-\tau)) \end{aligned}$$

for τ in the interval $[0, 2]$.

A plot of the output and the function $\xi(\tau)$

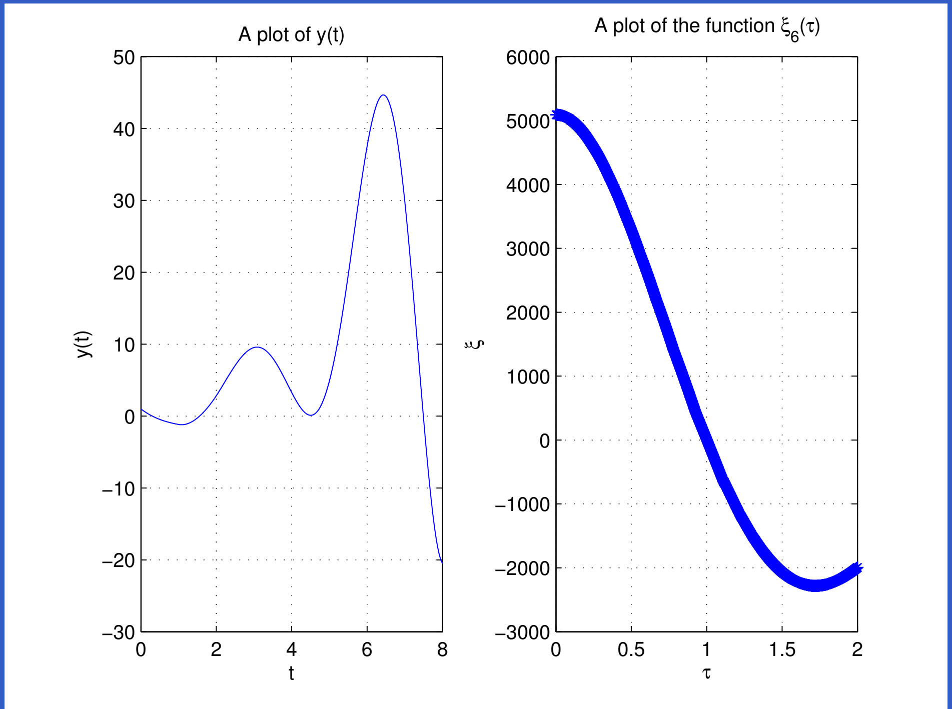

Figure 1: *The output* $y(t)$ *and the function* $\xi(\tau)$*.*

A delay system with unidentifiable time-lag parameter

Example 2:

$$\begin{cases} \dot{x}_1 & = & x_2^2(t-\tau) \\ \dot{x}_2 & = & x_2 \\ y & = & x_1 \\ x(t) & = & \varphi(t), \quad t \in [-\tau, 0] \end{cases}$$

Calculating time-derivatives of the output function as above, we obtain

$$\begin{aligned} \dot{y} &= x_2^2(t-\tau) \\ \ddot{y} &= 2(x_2(t-\tau))^2 \end{aligned}$$

Example 2, continued

An output equation of lowest degree (of derivation) for the above system is

$$\ddot{y}(t) - 2\dot{y}(t) = 0$$

- τ does not appear in the external representation of the system
- τ is not identifiable, as there is a symmetry involving the functions of initial conditions φ and τ

Example 2, continued

Setting

$$\begin{cases} \varphi_1(t) &= c \\ \varphi_2(t) &= e^{t+\tau} \end{cases}, \quad t \in [-\tau, 0] \quad ,$$

where c is a constant, leads to the solution

$$\begin{cases} x_1(t) &= \frac{e^{2t}}{2} + c - \frac{1}{2} \\ x_2(t) &= e^{t+\tau} \end{cases}, \quad \forall t \geq 0 \quad .$$

Since $y(t) = x_1(t)$, τ cannot be identified from the output.

Identifiability of the delay parameter

- Determined by the form of the external input-output representation of the system
- For simpler systems with few states and parameters, the time-lag can be identified directly from the input-output equations
- For more complex systems, the explicit i-o representation is difficult to obtain
- Linear-algebraic criteria eliminate the need for an explicit calculation

  Anguelova, M & Wennberg, B. State elimination and identifiability of the delay parameter for nonlinear time-delay systems. *Submitted*.

General form for the nonlinear delay systems

$$
\begin{cases}
\dot{x}(t) &= f(x(t), x(t-\tau), u, u(t-\tau)) \\
y(t) &= h(x(t), x(t-\tau), u, u(t-\tau)) \\
x(t) &= \varphi(t), \quad t \in [-\tau, 0] \\
u(t) &= u_0(t), \quad t \in [-T, 0]
\end{cases},
$$

Several outputs, inputs allowed.

The above system form also allows for model parameters, which can simply be considered as state variables with time-derivative zero.

Mathematical framework for delay systems

We have used the framework developed by

- Moog, C. H., Castro-Linares, R., Velasco-Villa, M.,& Márquez-Martínez, L. A. (2000). The disturbance decoupling problem for time-delay nonlinear systems. *IEEE Transactions on Automatic Control, 45*, 305-309.
- Márquez-Martínez, L. A., Moog, C. H., & Velasco-Villa, M. (2000). The structure of nonlinear time delay systems. *Kybernetika, 36*, 53-62.
- Xia, X., Márquez, L. A., Zagalak, P., & Moog, C. H. (2002). Analysis of nonlinear time-delay systems using modules over noncommutative rings. *Automatica, 38*, 1549-1555.

Mathematical framework for delay systems

- the time-shift operator δ is defined by:

$$\delta(\xi(t)) = \xi(t - \tau)$$

$$\delta d\xi(t) = d\xi(t - \tau)$$

- $\mathcal{K}$ - the field of meromorphic functions of a finite number of variables from $\{x(t - k\tau), u(t - k\tau), \ldots, u^{(l)}(t - k\tau), \quad k, l \in \mathbb{Z}^{+}\}$.

Mathematical framework for delay systems

- the time-shift operator δ is defined by:

$$\delta(\xi(t)) = \xi(t-\tau)$$

$$\delta d\xi(t) = d\xi(t-\tau)$$

- $\mathcal{K}$ - the field of meromorphic functions of a finite number of variables from $\{x(t-k\tau), u(t-k\tau), \ldots, u^{(l)}(t-k\tau), \quad k, l \in \mathbb{Z}^{+}\}$.

- $\mathcal{K}(\delta]$ - the set of polynomials of the form

$$a(\delta] = a_0(t) + a_1(t)\delta + \ldots + a_{r_a}(t)\delta^{r_a} \quad , a_j(t) \in \mathcal{K}$$

- $\mathcal{K}(\delta]$ is a non-commutative, Noetherian, left Ore ring which allows for row elimination in a matrix consisting of elements from it

Preliminary definitions

We have the equations

$$y_i^{(j)} = h_i^{(j)}$$

for the variables (and parameters) in the known outputs.

Preliminary definitions

We have the equations

$$y_i^{(j)} = h_i^{(j)}$$

for the variables (and parameters) in the known outputs.

Gather all "independent" output functions in the set

$$S = (h_1, ..., h_1^{(s_1-1)}, \ldots, h_p, \ldots, h_p^{(s_p-1)}) \quad ,$$

$s_i - 1$ corresponds to the highest derivative of the output function h_i in S.

Preliminary definitions

We have the equations

$$y_i^{(j)} = h_i^{(j)}$$

for the variables (and parameters) in the known outputs.

Gather all "independent" output functions in the set

$$S = (h_1, ..., h_1^{(s_1-1)}, \ldots, h_p, \ldots, h_p^{(s_p-1)}) \quad ,$$

$s_i - 1$ corresponds to the highest derivative of the output function h_i in S.

$$rank_{\mathcal{K}(\delta]} \frac{\partial S}{\partial x} \leq n \quad .$$

Linear-algebraic identifiability criteria

If at least one of the following is true:

1. $\frac{\partial h_i^{(j)}(t)}{\partial u_r^{(k)}(t-s\tau)} \neq 0$ for some $1 \leq i \leq p$, $0 \leq j \leq s_i$, $s \geq 1$, $1 \leq r \leq m$ and $k \geq 0$, i.e. a delayed input-variable $u_r^{(k)}$ occurs in some of the functions in $\{S, h_1^{(s_1)}, \ldots, h_p^{(s_p)}\}$;

2. $rank_{\mathcal{K}(\delta]} \frac{\partial S}{\partial x} \neq rank_{\mathcal{K}} \frac{\partial (S, h_1^{(s_1)}, \ldots, h_p^{(s_p)})}{\partial x}$;

then τ is locally identifiable.

Otherwise, the system can be realized in a generalized sense as a system of ODEs.

Linear-algebraic identifiability criteria, translation

- The local identifiability of τ depends on the presence or absence of τ in the input-output equations for the system

Linear-algebraic identifiability criteria, translation

- The local identifiability of τ depends on the presence or absence of τ in the input-output equations for the system
- Whether τ is present in the i-o equations is decided by
  - The occurrence of a delayed input variable in the time-derivatives of the output functions
  - Analysing whether $\frac{\partial(S, h_1^{(s_1)}, \ldots, h_p^{(s_p)})}{\partial x}$ can be row-reduced without δ

Example 1, revisited

$$\begin{cases} \dot{x}_1 &= k_1 x_2(t-\tau) + u(t) \\ \dot{x}_2 &= k_1 x_2(t-\tau) \\ y &= x_1 \end{cases}$$

$$\begin{pmatrix} dy \\ d\dot{y} - du \\ d\ddot{y} - d\dot{u} \\ dy^{(3)} - d\ddot{u} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & k_1\delta & x_2(t-\tau) & 0 \\ 0 & k_1 k_2 \delta^2 & k_2 x_2(t-2\tau) & k_1 x_2(t-2\tau) \\ 0 & k_1 k_2^2 \delta^3 & k_2^2 x_2(t-3\tau) & 2k_1 k_2 x_2(t-3\tau) \end{pmatrix} \begin{pmatrix} dx_1 \\ dx_2 \\ dk_1 \\ dk_2 \end{pmatrix} .$$

- Rank 4 over $\mathcal{K}$ and rank 3 over $\mathcal{K}(\delta]$
- τ is locally identifiable
- The system is not weakly observable (k_1 and x_2) as the matrix above is not of full-rank over $\mathcal{K}(\delta]$

Example 2, revisited

$$\begin{cases} \dot{x}_1 & = & x_2^2(t-\tau) \\ \dot{x}_2 & = & x_2 \\ y & = & x_1 \end{cases}$$

$$\begin{pmatrix} dy \\ d\dot{y} \\ d\ddot{y} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 2x_2(t-\tau)\delta \\ 0 & 4x_2(t-\tau)\delta \end{pmatrix} \begin{pmatrix} dx_1 \\ dx_2 \end{pmatrix} .$$

- Rank 2 over both $\mathcal{K}$ and $\mathcal{K}(\delta]$
- τ is not identifiable

The JAK-STAT signalling pathway model, revisited

$$\begin{cases} \dot{x}_1 &= -k_1 x_1 u/k_7 + 2k_4 x_3(t-\tau) \\ \dot{x}_2 &= k_1 x_1 u/k_7 - k_2 x_2^2 \\ \dot{x}_3 &= -k_3 x_3 + 0.5 k_2 x_2^2 \\ \dot{x}_4 &= k_3 x_3 - k_4 x_3(t-\tau) \\ y_1 &= k_5(x_2 + 2x_3) \\ y_2 &= k_6(x_1 + x_2 + 2x_3) \end{cases},$$

$$\begin{aligned} \dot{y}_2 &= 2k_6(k_4 x_3(t-\tau) - k_3 x_3) \\ \ddot{y}_2 &= 2k_6(-k_3 k_4 x_3(t-\tau) + 0.5 k_2 k_4 x_2^2(t-\tau) + \\ &\quad + k_3^2 x_3 - 0.5 k_2 k_3 x_2^2) \\ y_2^{(3)} &= k6(-2k_3 k_2 x_2 k_1 x_1 u + 2k_3 k_2^2 x_2^3 k_7 - 2k_3^3 k_7 x_3 + \\ &\quad + k_3^2 k_7 k_2 x_2^2 + 2k_4 k_2 x_2(t-\tau) k_1 x_1(t-\tau) u(t-\tau) - \ldots \end{aligned}$$

$u(t-\tau)$ appears in $y_2^{(3)}$ and τ is identifiable

A gene expression model for Hes1

- Two state variables P and M and six parameters $p = (\alpha_m, P_0, n, \mu_m, \alpha_p, \mu_p)$

$$
\begin{cases}
\dot{M} &= \frac{\alpha_m}{1+(P(t-\tau)/P_0)^n} - \mu_m M \\
\dot{P} &= \alpha_p M - \mu_p P \\
y_1 &= M \\
y_2 &= P \quad ,
\end{cases}
$$

Monk, N. A. M.. (2003). Oscillatory expression of Hes1, p53, and NF-κB driven by transcriptional time delays. *Curr. Biol., 13*, 1409-1413.

The gene expression model for Hes1, continued

$$\dot{y}_1 = \frac{\alpha_m}{1 + (\delta P/P_0)^n} - \mu_m M$$

$$\ddot{y}_1 = h_1^{(2)}(P(t-\tau), M(t-\tau), P, M, \ldots)$$

$$y_1^{(3)} = h_1^{(3)}(P(t-2\tau), P(t-\tau), M(t-\tau), P, M, \ldots)$$

$$y_1^{(4)} = h_1^{(3)}(M(t-2\tau), P(t-2\tau), P(t-\tau), M(t-\tau), P, M, \ldots$$

- $\frac{\partial h_1^{(j)}}{\partial M}$ or $\frac{\partial h_1^{(j)}}{\partial P}$ polynomial in δ of degree higher than for $j-1$
- $rank_{\mathcal{K}} \frac{\partial(S, h_1, \ldots, h_1^{(8)})}{\partial x, p} = 9$, greater than $rank_{\mathcal{K}(\delta]} \frac{\partial S}{\partial x, p}$ (limited by the number of variables and parameters)
- τ is identifiable

Checking the criteria in practice

- If for some output function h_i each derivative $h_i^{(j)}, \quad j = 0, \ldots, n$ contains a state-variable that is delayed compared to the previous derivative, then the delay parameter is identifiable.
- If a delayed input variable (or its derivative) appears in the first $s_i + 1$ equations for some output function h_i, then the delay parameter is identifiable.

  In such cases the identifiability of the delay can be decided without any rank calculations.

Several delays

$$\begin{cases} \dot{x}_1 &= x_2(t-\tau_1) \\ \dot{x}_2 &= x_1(t-\tau_2) \\ y_1 &= x_1 \\ y_2 &= x_2(t-\tau_2) \\ x &= \varphi(t), \quad t \in [-T, 0] \end{cases}$$

Input-output equations:

$$\begin{aligned} \ddot{y}_1(t) &= y_1(t-\tau_1-\tau_2) \\ y_2(t-\tau_1) &= \dot{y}_1(t-\tau_2) \end{aligned} .$$

τ_1 and τ_2 identifiable

Acknowledgements

Financial support:

- The National Research School in Genomics and Bioinformatics
- The Swedish Research Council
- The Swedish Foundation for Strategic Research via the Gothenburg Mathematical Modelling Center