Nonlinear observers

basic concepts; nonlinear systens and DAE

Erik Frisk

Dept. of Electrical Engineering, Linköping University

27 november 2018

LINKÖPING UNIVERSITY

Today's topics, reading material and exercises

- Systems described by differential-algebraic equations and non-linear differential-equations.
- Reading material
  1. Article on DAE:s by E. Yip and R.F. Sincovec
  2. Article on non-linear observers by Besançon (section 4 and 5 can be skipped).
- Exercises Le2.1 - Le2.8 described in this lecture material.

Reading material, cont.

Texts that are available, but possibly not recommended in a first reading of the course. This can serve as depth-material later.

- An (advanced) basic reference in non-linear controllability and observability by R. Hermann and A.J. Krener. The text is formal and rather difficult, but covers fundamental questions.
- A chapter from Torkel Glad:s text on non-linear control theory. Starts out easy, but quickly becomes advanced. Primarily for those who have taken a course in non-linear control already.
- If you want to know (much) more about DAE:s, including observers, check out "*Singular control systems*", 1989, L. Dai, Springer-Verlag.

Outline

Outline

- Introduction
- *Differential-Algebraic Models*
  - *What is a DAE?*
  - *Observability for linear DAE models*
  - *What does an observer for a DAE model look like?*
- Nonlinear models
  - Basic definitions
  - Test observability
- Asymptotic observers
- Appendix

What is a DAE?

A linear DAE can be written in different forms, the most common is

$$\begin{aligned} E\dot{x} &= Ax + Bu, \quad x(0) = x_0 \\ y &= Cx \end{aligned}$$

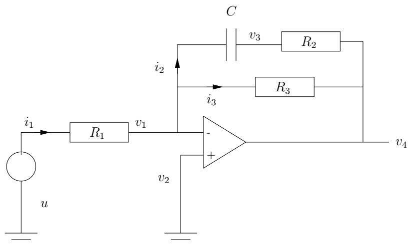
where E can be a singular matrix (and even non-quadratic).

Useful in different contexts. Models written in Modelica are typically in this form.

We will not go in-depth on DAE in this course (see the simulation course or a pure DAE course for more details)

However, we need some basic results on solvability and canonical form.

DAE, an example



$(e_1): \; u - v_1 = R_1 i_1$	[resistance]	$(e_5): \; v_1 = v_2$	[op-amplifier]
$(e_2): \; v_3 - v_4 = R_2 i_2$	[resistance]	$(e_6): \; v_2 = 0$	[ground]
$(e_3): \; v_1 - v_4 = R_3 i_3$	[resistance]	$(e_7): \; i_1 = i_2 + i_3$	[Kirchoff]
$(e_4): \; C\dfrac{d}{dt}(v_1 - v_3) = i_2$	[inductor]		

A component based model gives a DAE model

$$E\dot{x} = Ax + Bu$$

with $x = (i_1, i_2, i_3, v_1, v_2, v_3, v_4)$.

Solvability for DAE models

Solvability for DAE:s has some additional things to consider compared to ODE:s

$$y = \dot{u} \quad \Leftrightarrow \quad \begin{aligned} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \dot{x} &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} x + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u \\ y &= x_2 \end{aligned}$$

- Any $x(0)$ is not feasible ($x_1(t) = u(t)$, $x_2(t) = \dot{u}(t)$)
- If E is quadratic and x_0 is consistent with the model then the model has a unique solution if $\lambda E - A$ has full rank for any $\lambda \in \mathbb{C}$.

You can write down the transfer function, just as for ODE models

$$y(t) = G(p)u(t) = C(pE - A)^{-1}Bu(t)$$

A canonical form for DAE models

A canonical form for a linear DAE that is useful is the following

Theorem

For a regular DAE model

$$E\dot{x} = Ax + Bu$$

there exists invertible matrices P and T such that when $Tw = x$ and multiplication of the model equations with P from the left

$$\begin{aligned} \dot{w}_1 &= A_1 w_1 + B_1 u \\ E_2 \dot{w}_2 &= w_2 + B_2 u \end{aligned}$$

where matrix E_2 is nilpotent with order $m \leq n$

Partitions model into fast and slow dynamics. Here, we "ignore" technicalities that arises due to impulses.

Outline

Observability for DAE

The definition of observability is no different from previous definitions.

Definition (Observability)

A linear, regular, DAE is observable if $x(t)$ is uniquely determined by $y(t)$.

Consider the canonical form

$$\begin{aligned} \dot{x}_1 &= A_1 x_1 + B_1 u \\ E_2 \dot{x}_2 &= x_2 + B_2 u, \qquad E_2^k \neq 0, k < m, E_2^m = 0 \\ y &= \begin{pmatrix} C_1 & C_2 \end{pmatrix} x \end{aligned}$$

Simple calculations give that

$$\begin{aligned} x_2(t) &= E_2\dot{x}_2 - Bu = E_2(E_2\ddot{x}_2 - B\dot{u}) - Bu = E_2^2\ddot{x}_2 - E_2 B\dot{u} - Bu = \\ &= \cdots = E_2^m x_2^{(m)} - \sum_{i=0}^{m-1} E_2^i B_i u^{(i)}(t) = -\sum_{i=0}^{m-1} E_2^i B_i u^{(i)}(t) \end{aligned}$$

Observability for a DAE model

This gives that observability for a DAE

$$\begin{aligned} \dot{x}_1 &= A_1 x_1 + B_1 u \\ E_2 \dot{x}_2 &= x_2 + B_2 u, \qquad E_2^k \neq 0, k < m, E_2^m = 0 \\ y &= \begin{pmatrix} C_1 & C_2 \end{pmatrix} x \end{aligned}$$

is equivalent to (A_1, C_1) is an observable pair, i.e., that

$$\begin{bmatrix} C_1 \\ C_1 A_1 \\ \vdots \\ C_1 A_1^{n_1 - 1} \end{bmatrix} \text{ have full column rank or } \begin{bmatrix} \lambda I - A_1 \\ C_1 \end{bmatrix} \text{ full rank for all } \lambda \in \mathbb{C}$$

Computing the canonical form for a general DAE, i.e., finding matrices P and T, is it really necessary to determine observability?

DAE model - exercise

Exercise Le2.1: Prove that a regular DAE

$$E\dot{x} = Ax + Bu, \quad y = Cx$$

is observable if and only if

$$\text{rank} \begin{pmatrix} \lambda E - A \\ C \end{pmatrix} = n, \quad \forall \lambda \in \mathbb{C}$$

Tips:

- Use the canonical form

Observability for DAE models

- $\mathcal{C}$-observability - complete observability

$$\begin{pmatrix} \alpha E - \beta A \\ C \end{pmatrix} \text{ full rank for all } \alpha, \beta \in \mathbb{C}^2 \setminus \{(0,0)\}$$

- $\mathcal{R}$-observability - reachable observability

$$\begin{pmatrix} \lambda E - A \\ C \end{pmatrix} \text{ full rank for all } \lambda \in \mathbb{C}$$

- $\mathcal{I}$-observability - Impulse observability

$$\begin{pmatrix} E \\ K_{E^T}^T A \\ C \end{pmatrix} \text{ full rank}$$

where the rows in K_{E^T} spans $\ker E^T$

Message: There are several definitions, $\mathcal{R}$-observability is the one normally referred to.

Outline

What does an observer for a DAE model look like?

Look at an ob sever for

$$E\dot{x} = Ax + Bu, \quad y = Cx$$

you'll see formulations like (singular/normal)

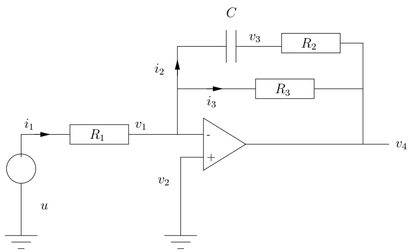
$$\begin{aligned} E_c\dot{w} &= A_c w + B_c u + Gy \\ \hat{x} &= F_c w + Fu + Hy \end{aligned} \qquad \begin{aligned} \dot{w} &= A_c w + B_c u + Gy \\ \hat{x} &= F_c w + Fu + Hy \end{aligned}$$

where the right one often is more common/desired.

Theorem

For any observable DAE there exists a normal observer of order rank E.

DAE, exercise



$(e_1): \; u - v_1 = R_1 i_1$	[resistance]	$(e_5): \; v_1 = v_2$	[op-amplifier]
$(e_2): \; v_3 - v_4 = R_2 i_2$	[resistance]	$(e_6): \; v_2 = 0$	[ground]
$(e_3): \; v_1 - v_4 = R_3 i_3$	[resistance]	$(e_7): \; i_1 = i_2 + i_3$	[Kirchoff]
$(e_4): \; C\dfrac{d}{dt}(v_1 - v_3) = i_2$	[inductor]		

DAE exercise, cont.

Exercise Le2.2:

a) Write the model equations in the form

$$E\dot{x} = Ax + Bu$$

with $x = (i_1, i_2, i_3, v_1, v_2, v_3, v_4)$.

b) Assume possible measurements are voltages $v_1, \dots, v_4$ and currents $i_1, \dots, i_3$. Determine which measurement points that makes the model observable and which that doesn't make it observable.

c) Choose one such signal and design an observer.

If it makes it less messy, invent simple values for resistor and inductor parameters.

Outline

Nonlinear models

- observability is not necessarily a global property
- Excitation conditions could be needed, e.g., requirement on $u(t)$

$$\dot{x}_1 = x_2 u$$
$$\dot{x}_2 = -x_2$$
$$y = x_1$$

Requires finer and more specialized concepts to define what observability exactly means.

Indistinguishability

$$\dot{x} = f(x, u), \quad x(0) = x_0, \quad y = h(x), x \in \mathcal{M}$$

Definition (Indistinguishability)

Initial states x_1 and x_2 is said to be indistinguishable if for all $u(t)$ they give the same output.

Let $I(x)$ be the set of states indistinguishable from x.

Definition (Observable at x^0)

A system is said to be observable at x^0 if $I(x^0) = \{x^0\}$. The system is observable on $\mathcal{M}$ if it is observable for all $x^0 \in \mathcal{M}$.

Thus, there only need to exist a control signal $u(t)$ such that states are distinguishable, does not need to be for all $u(t)$.

Observability

For the model

$$\dot{x} = u$$
$$y = \sin(x)$$

it holds that

$$I(x^0) = \{x | x = x^0 + k2\pi\}$$

This system is this not observable on $\mathbb{R}$.

A more local property

With $\mathcal{M} = \mathbb{R}^+$, then the model below is observable

$$\dot{x} = -\epsilon x, \quad y = h(x) = \begin{cases} x & x < 1 \\ 0 & \text{otherwise} \end{cases}$$

but it can take a very long time before the difference between initial states is visible.

Definition

$\mathcal{U}$-indistinguishable Let $\mathcal{U} \subseteq \mathcal{M}$, then $x_1 \, I_{\mathcal{U}} \, x_2$ ($\mathcal{U}$-indistinguishable) if for all $u(t)$ such that $x_i(t) \in \mathcal{U}$ it holds that the outputs are equal.

Definition (locally observable)

A model is*locally observable* at x_1 if for **all** $\mathcal{U} \subseteq \mathcal{M}$ it holds that $I_{\mathcal{U}}(x_0) = \{x_0\}$.

Local observability is a stronger property than observability. There exists a $u(t)$ such that you immediately see a difference.

Local observability too strong condition

- for linear systems it holds that if $x^1 I x^2$, $x^2 I x^3$ then $x^1 I x^3$ (transitivity is intuitive). This is not true for local observability.
- Too restrictive, wee need something less strict.
- It is often enough to differentiate between "neighbors", i.e., it is only when states that are close are not distinguishable that it is problematic.

Weak observability

A state x^0 is weakly observable means that it can be distinguished from its neighbors which leads to the following definition

Definition (weakly observable)

A state x^0 is weakly observable if there exists a neighborhood $\mathcal{U}$ to x^0 such that

$$I(x^0) \cap \mathcal{U} = \{x^0\}$$

We still have the problem that it could take a very long time to see the difference.

Weak and local observability

Combining the definitions for weak and local observability; existence of a neighborhood such that within that neighborhood states can immediately be distinguished.

Definition

A state x^0 is weakly locally observable if there exists a neighborhood $\mathcal{U}$ such that for all $\mathcal{V} \subseteq \mathcal{U}$ it holds that $I_{\mathcal{V}}(x^0) = \{x^0\}$.

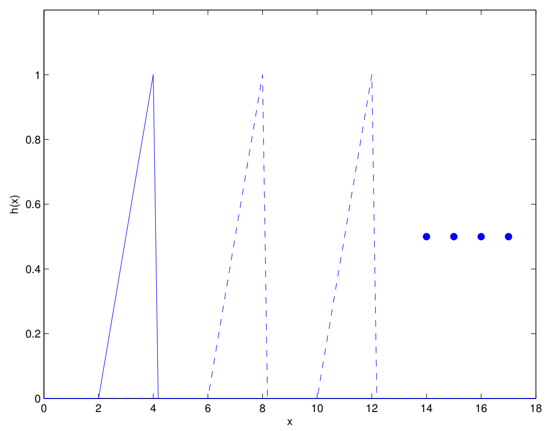
We then have the desired(?) properties that we do not need to wait long (local) to observer differences between neighboring (weak) states.

Relations between definitions

$$\begin{tikzcd}
\text{locally observable} \arrow[r] \arrow[d] & \text{observable} \arrow[d] \\
\text{locally weakly observable} \arrow[r] & \text{weakly observable}
\end{tikzcd}$$

Illustrative example

$$\dot{x} = 1$$
$$y = h(x)$$



Exercises on the example

First assume $h(x)$ without the dashed repetition. The system is not observable on $\mathcal{M} = \mathbb{R}^+$.

Exercise Le2.3: What happens when

a) $\mathcal{M} = (0, 4)$

b) $\mathcal{M} = (2, 4)$

c) With the dashed repetition in $h(x)$ and $\mathcal{M} = \mathbb{R}^+$

Outline

Test observability for nonlinear systems

There are at least 4 different definitions

$$\begin{array}{ccc} \text{locally observable} & \longrightarrow & \text{observable} \\ \downarrow & & \downarrow \\ \text{locally weakly observable} & \longrightarrow & \text{weakly observable} \end{array}$$

To test observability is in general a difficult problem, but for local weak observability there is a simple sufficient condition.

What did we do for linear systems?

For a linear model,

$$\dot{x} = Ax, \quad y = Cx$$

differentiate the measurement signal $n-1$ times

$$Y = \begin{pmatrix} y(t) \\ \dot{y}(t) \\ \vdots \\ y^{(n-1)}(t) \end{pmatrix} = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{pmatrix} x(t) = \mathcal{O}x(t)$$

Observability is then equivalent to if we can solve the system of equations

$$Y = \mathcal{O}x(t)$$

which corresponds to that the matrix $\mathcal{O}$ have full column rank.

Inverse function theorem

Let a function

$$y = f(x)$$

be a differentiable function from $\mathbb{R}^n$ to $\mathbb{R}^m$ such that

$$\text{rang } \frac{\partial}{\partial x} f(x) \bigg|_{x=x^0} = n,$$

then there exist neighborhoods around x^0 and $y^0 = f(x^0)$ such that an inverse function exists.

Useful notation: Lie derivatives

With

$$\dot{x} = f(x)$$
$$y = h(x)$$

then

$$\dot{y}(t) = h_x f = L_f h(x)$$

Define Lie derivatives as

$$L_f^k h(x) = \frac{\partial}{\partial x}\left(L_f^{k-1} h(x)\right) f(x), \quad L_f^0 h(x) = h(x)$$

i.e., the derivative of $h(x)$ along $f(x)$.

With this notation, a simple notation for

$$y^{(k)} = L_f^k h(x)$$

Observability rank condition

Do the same as for linear systems

$$Y = \begin{pmatrix} y(t) \\ \dot{y}(t) \\ \vdots \\ y^{(q)}(t) \end{pmatrix} = \begin{pmatrix} h(x) \\ L_f h(x) \\ \vdots \\ L_f^q h(x) \end{pmatrix} = F(x)$$

According to the inverse function theorem, a **sufficient** condition for a unique x, in a neighborhood of $x = x^0$ and $y^0 = F(x^0)$, is that

$$\exists q. \text{ rang } \frac{\partial}{\partial x} F(x) \bigg|_{x=x^0} = n$$

A simple algebraic condition that can be tested (for a given q) given that the system is smooth enough.

Theorem

A model is locally weakly observable at $x = x_0$ if the rank condition above is fulfilled in $x = x^0$.

Observability rank condition

Often it is enough to test, as for linear systems, with $q = n$ but it is not generally true.

Exercise Le2.4: Consider the two models below, determine for each of them if they are observable and if the fulfill the rank condition, and if so for which value of q.

$$\dot{x} = 1 \qquad\qquad \dot{x} = 1$$
$$y = x^3 \qquad\qquad y = x\, e^{-1/x^2}$$

What about the linearized model?

A possibility to determine local observability for a model

$$\dot{x} = f(x, u), \quad y = h(x)$$

at a point $x = x^0$ is to look at the observability for the pair(A, C)

$$A = f_x(x)|_{x=x^0}, \quad C = h_x(x)|_{x=x^0}$$

Exercise Le2.5: Analyze the model $f(x, u) = u$ and $h(x) = x^2$ in a point $x^0 = 0$. Is the non-linear system observable, and if so how? What conclusions can be drawn from the linearization?

Outline

Asymptotic observers

Just observability is not enough in general; you also want to design an observer such that $\hat{x}(t) \to x(t)$, exponential convergence is good, robustness against model uncertainty etc.

For a model and observer

$$\begin{aligned} \dot{x} &= f(x, u) \qquad\qquad \dot{\hat{x}} = f(\hat{x}, u) + K(y - h(\hat{x})) \\ y &= h(x) \end{aligned}$$

it is then properties of the error dynamics

$$\dot{e} = f(x, u) - f(\hat{x}, u) - K(y - h(\hat{x}))$$

that are of interest.

Asymptotic observers

Analysis of the error dynamics, and thereby the design of the observer, falls back on a basic stability problem.

many design methods build upon Lyapunov techniques, i.e., define a measure of the size of the error, a positive definite function $V(e)$, and design the observer such that

$$\dot{V}(e) < 0$$

because then it holds that $V(e) \to 0$ which implies that $e \to 0$.

This often gets technical, but I'll include one typical result.

Example from Besançon

For a model

$$\dot{x} = A(u)x + B(u)$$
$$y = Cx$$

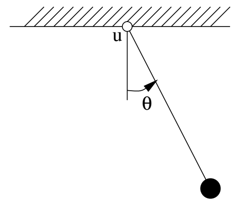
a candidate observer is

$$\dot{\hat{x}} = A(u)\hat{x} + B(u) + S^{-1}C^T(y - C\hat{x})$$
$$\dot{S} = -\theta S - A^T(u)S - SA(u) + C^T C$$
$$S(0) = S^T(0) > 0$$

Why this? Why not $2C^T C$ in the $\dot{S}$-equations?

Design example - a pendulum

Exercise Le2.6: Consider a simple pendulum model with viscous damping:



Assume you can control the pendulum with a torque and that the angle θ can be measured.

Design example - a pendulum, cont.

With $x_1 = \theta$ and $x_2 = \dot{\theta}$ a normalized model is given by

$$\dot{x} = f(x, u) = \begin{pmatrix} x_2 \\ -x_2 - \sin x_1 + u \end{pmatrix}, \quad y = x_1$$

The objective of the exercise is to design a proven stable observer in the form

$$\dot{\hat{x}} = f(\hat{x}, u) + K(y - \hat{x}_1)$$

and verify that the system is observable (and how).

Design example - a pendulum, cont.

Let $e = (e_1, e_2)$ be the estimation error and define a quadratic Lyapunov function

$$V(e) = e_1^2 + \beta e_2^2, \quad \beta > 0$$

Choose K such that $\dot{V} < 0$ which then guarantees that $e(t) \to 0$.

Hint: The inequality

$$0 \leq \frac{\sin x - \sin y}{x - y} \leq 1, \quad -\pi/2 \leq x, y \leq \pi/2$$

is useful.

Relative degree and observability

Exercise Le2.7: Show why a non-linear SISO system,

$$\dot{x} = f(x) + g(x)u$$
$$y = h(x)$$

with relative degree equal to the number of states always is observable (which kind?).

Hint: Start with a linear example and look at a suitable choice of state variables.

Exercise

Exercise Le2.8: Assume a model described by the equations

$$\dot{x} = \begin{pmatrix} x_1^2 x_2^2 + u \\ 1 - x_1 x_2^3 - x_2 - \frac{x_2}{x_1} u \end{pmatrix}$$
$$y = x_1 x_2$$

on the set $\mathcal{M} = \{(x_1, x_2) | x_1 \neq 0\}$.

Prove that the system is observable and design an observer with proven global error stability.

Outline

Relative degree

A SISO system

$$\dot{x} = f(x) + g(x)u$$
$$y = h(x)$$

is said to have relative degree r in a point x^0 if we have to differentiate $y(t)$ r times for the control input signal u to appear explicitly in a neighborhood of x^0. Equivalent to

$$L_g L_f^k h(x) = 0, \text{for all } x \text{ in a neighborhood of } x^0 \text{ and all } k < r-1$$
$$L_g L_f^{r-1} h(x^0) \neq 0$$

For a linear system $b(s)/a(s)$, the relative degree is $r = \deg(a(s)) - \deg(b(s))$.

Change of state variables in nonlinear models

Let $\Phi(x)$ be differentiable, (locally) invertible. A change of variables $z = \Phi(x)$ then gives

$$\begin{aligned} \dot{x} &= f(x,u) \\ y &= h(x) \end{aligned} \quad \Rightarrow \quad \begin{aligned} \dot{z} &= \Phi(x) f(x,u)|_{x=\Phi^{-1}(z)} \\ y &= h(\Phi^{-1}(z)) \end{aligned}$$

If you want to practice, test performing the change of variables

$$z = \begin{pmatrix} x_1 \\ \frac{x_2}{x_1} \end{pmatrix} \quad \text{in the model} \quad \begin{aligned} \dot{x} &= \begin{pmatrix} \frac{x_2^2}{x_1^2} \\ \frac{x_2^3}{x_1^3} - x_2 \end{pmatrix} + \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} u \\ y &= x_1 x_2 \end{aligned}$$

Rank test

A rank test I have used when solving the exercises (you might find other ways, so don't worry if you don't need the result)

Definition (Minor)

If A is an $n \times m$ matrix, then the determinant of a $p \times p$ submatrix obtained from A is called a minor of order p.

Theorem (Rank)

The rank of an $n \times m$ matrix is equal to the order of its largest non-zero minor.

Corollary (Full column rank test)

A tall $n \times m$ matrix is full column rank if and only if there exists a non-zero minor of order m.

Nonlinear observers

basic concepts; nonlinear systens and DAE

Erik Frisk

Dept. of Electrical Engineering, Linköping University

27 november 2018

LINKÖPING UNIVERSITY