KKL observer design for non observable systems

Pauline Bernard

Centre Automatique et Systèmes
MINES ParisTech, Université PSL

Collaboration with L. Praly, M. Spirito, L. Marconi, L. Brivadis, V. Andrieu, U. Serres

Nice, November, 2021

Contents

Standard observation problem

We consider a nonlinear dynamical system

$$\dot{x} = f(x, u) \quad , \quad y = h(x, u) \tag{1}$$

with state x in $\mathbb{R}^{d_x}$, input $u : \mathbb{R} \to \mathbb{R}^{d_u}$, output $y : \mathbb{R} \to \mathbb{R}^{d_y}$.

Standard observation problem

We consider a nonlinear dynamical system

$$\dot{x} = f(x, u) \quad , \quad y = h(x, u) \tag{1}$$

with state x in $\mathbb{R}^{d_x}$, input $u : \mathbb{R} \to \mathbb{R}^{d_u}$, output $y : \mathbb{R} \to \mathbb{R}^{d_y}$.

Observation problem

Whatever u in $\mathcal{U}$,
whatever the initial condition $x_0 \in \mathcal{X}_0$,
find an estimate $\hat{x}(t)$ of $x(t)$ at each time $t \geq 0$, based on $u_{[0,t]}$ and $y_{[0,t]}$ and such that

$$\lim_{t \to +\infty} \hat{x}(t) - x(t) = 0 \ .$$

Standard observation problem

We consider a nonlinear dynamical system

$$\dot{x} = f(x,u) \quad , \quad y = h(x,u) \tag{1}$$

with state x in $\mathbb{R}^{d_x}$, input $u : \mathbb{R} \to \mathbb{R}^{d_u}$, output $y : \mathbb{R} \to \mathbb{R}^{d_y}$.

Observation problem

Whatever u in $\mathcal{U}$,
whatever the initial condition $x_0 \in \mathcal{X}_0$,
find an estimate $\hat{x}(t)$ of $x(t)$ at each time $t \geq 0$, based on $u_{[0,t]}$ and $y_{[0,t]}$ and such that

$$\lim_{t \to +\infty} \hat{x}(t) - x(t) = 0 \ .$$

=> Observer design

Luenberger design : history

In *Observing the state of a linear system*, D. Luenberger, 1964

x-coordinates in $\mathbb{R}^{d_x}$ $\qquad$ ξ-coordinates in $\mathbb{R}^{d_\xi}$

$$\begin{array}{rcl} \dot{x} & = & F\,x \\ y & = & H\,x \end{array} \quad \xrightarrow{\quad T \quad} \quad \dot{\xi} = A\,\xi + B\,y$$

A Hurwitz

$$\hat{x} = \ ? \quad \xleftarrow{\quad T^{-1} \quad} \quad \dot{\hat{\xi}} = A\,\hat{\xi} + B\,y$$

We have

$$\lim \hat{\xi} - \xi = 0$$

and if T invertible,

$$\lim \hat{x} - x = 0$$

Goal : find T invertible such that $\xi = Tx$ verifies $\dot{\xi} = A\xi + B\,y$ with A Hurwitz

Goal : find T invertible such that $\xi = Tx$ verifies $\dot{\xi} = A\xi + B\,y$ with A Hurwitz

=> T must verify the "Luenberger equation"

$$TF = AT + BH$$

Goal : find T invertible such that $\xi = Tx$ verifies $\dot{\xi} = A\xi + B\,y$ with A Hurwitz

=> T must verify the "Luenberger equation"

$$TF = AT + BH$$

- unique solution if F and A have no common eigenvalue
- invertible if (F, H) observable and (A, B) controllable

Goal : find T invertible such that $\xi = Tx$ verifies $\dot{\xi} = A\xi + B\,y$ with A Hurwitz

=> T must verify the "Luenberger equation"

$$TF = AT + BH$$

- unique solution if F and A have no common eigenvalue
- invertible if (F, H) observable and (A, B) controllable

=> **Luenberger observer** :

$$\dot{\hat{\xi}} = A\hat{\xi} + B\,y \qquad , \qquad \hat{x} = T^{-1}\hat{\xi}$$

$$\iff$$

$$\dot{\hat{x}} = F\,\hat{x} + T^{-1}B(y - H\hat{x})$$

Extension to nonlinear systems ?

Local analysis :

A. Shoshitaishvili, *On control branching systems with degenerate linearization*, NOLCOS, 1992

N. Kazantzis, C. Kravaris, *Nonlinear observer design using Lyapunov's auxiliary theorem*, SCL, 1998

A. Krener and M. Xiao, *Nonlinear observer design in the Siegel domain*, SIAM, 2003

Global analysis (autonomous) :

G. Kreisselmeier, R. Engel, *Nonlinear observers for autonomous Lipschitz continuous systems*, TAC, 2003

V. Andrieu, L. Praly, *On the existence of a Kazantzis–Kravaris/Luenberger observer*, SIAM, 2006

V. Andrieu, *Convergence speed of nonlinear Luenberger observers*, SIAM, 2014

Global analysis (time-varying) :

R. Engel, *Nonlinear observers for Lipschitz continuous systems with inputs* , IJC, 2007

P. Bernard, V. Andrieu, *Luenberger observers for non autonomous nonlinear systems*, TAC, 2019

x-coordinates in $\mathbb{R}^{d_x}$ | ξ-coordinates in $\mathbb{R}^{d_\xi}$

$$
\begin{tikzcd}
\begin{array}{rcl} \dot{x} & = & f(x,u) \\ y & = & h(x,u) \end{array} \arrow[r, "T_u(\cdot,t)"] \arrow[d, dashed] & \dot{\xi} = A\,\xi + B\,y \arrow[d] \\
\hat{x} = \ ? & \dot{\hat{\xi}} = A\,\hat{\xi} + B\,y \arrow[l, "\mathcal{T}_u(\cdot,t)"']
\end{tikzcd}
$$

A Hurwitz

x-coordinates in $\mathbb{R}^{d_x}$ | ξ-coordinates in $\mathbb{R}^{d_\xi}$

$$
\begin{tikzcd}[column sep=huge, row sep=large]
\begin{array}{rcl} \dot{x} &=& f(x,u) \\ y &=& h(x,u) \end{array} \arrow[r, "T_u(\cdot,t)"] \arrow[d, dashed] & \dot{\xi} = A\,\xi + B\,y \arrow[d] \\
\hat{x} = \ ? & \dot{\hat{\xi}} = A\,\hat{\xi} + B\,y \arrow[l, "\mathcal{T}_u(\cdot,t)"']
\end{tikzcd}
$$

A Hurwitz

Theorem 1 (KKL observer, general statement).

Let u such that the trajectories of interest remain in $\mathcal{X}$ and the system is backward-distinguishable on $\mathcal{X}$.

Then, there exists a set S of zero-Lebesgue measure in $\mathbb{C}^{d_x+1}$ such that for any diagonalizable Hurwitz matrix $A_0 \in \mathbb{C}^{(d_x+1)\times(d_x+1)}$ and any $B_0 \in \mathbb{C}^{d_x+1}$ such that $\mathrm{eig}(A_0) \notin S$ and (A_0, B_0) controllable,

there exists T_u such that

- *$\xi(t) = T_u(x(t),t)$ is solution to $\dot{\xi} = A\xi + By$ with $A = A_0 \otimes I_{d_y}$, $B = B_0 \otimes I_{d_y}$*
- *$x \mapsto T_u(x,t)$ becomes **injective** on $\mathcal{X}$, at least after some time.*

1. Introduction
2. Non observable system
   - Sensorless PMSM with unknown resistance
   - Indistinguishable trajectories
   - KKL observer design
3. Functional KKL observer
   - Existence of T
   - Injectivity of T with respect to q
   - Left-inverse and convergence
4. KKL and indistinguishability

Contents

Sensorless PMSM with unknown resistance

PMSM model in a fixed $\alpha\beta$-frame :

$$\dot{\Psi} = u - R\, i$$

where

- $\Psi \in \mathbb{R}^2$: total flux generated by the windings and the permanent magnet
- $u \in \mathbb{R}^2$, $i \in \mathbb{R}^2$: voltage and intensity of the current in the fixed frame
- R : stator winding resistance

For a non-salient PMSM :

$$\Psi = Li + \Phi \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$$

where L is the inductance, Φ the magnet's flux, and θ the electrical phase.

"Sensorless" PMSM model :

$$\dot{\Psi} = u - R\,i$$

$$\Psi = Li + \Phi \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \iff \begin{cases} |\Psi - Li|^2 - \Phi^2 = 0 \\ \theta = \arg(\Psi - Li) \end{cases}$$

Measured input signals
Parameters assumed known

Goal : estimate θ with R unknown $\iff$ estimate Ψ with R unknown

Model :

$$\begin{cases} \dot{\Psi} &= u - R\,i \\ \dot{R} &= 0 \end{cases} \quad , \quad y = |\Psi - Li|^2 - \Phi^2 = 0 \qquad \text{"virtual measurement"}$$

Observability analysis

For inputs (u, i), the model is

$$\begin{cases} \dot{x} = u - x_3\, i \\ \dot{x}_3 = 0 \\ \\ y = 0 = x^\top x - 2Li^\top x + L^2|i|^2 - \Phi^2 \end{cases}$$

with state $x = (x_1, x_2) \in \mathbb{R}^2$ and $x_3 \in \mathbb{R}$.

Observability analysis

For inputs (u, i), the model is

$$\begin{cases} \dot{x} & = & u - x_3\, i \\ \dot{x}_3 & = & 0 \\ y = 0 & = & x^\top x - 2Li^\top x + L^2|i|^2 - \Phi^2 \end{cases}$$

with state $x = (x_1, x_2) \in \mathbb{R}^2$ and $x_3 \in \mathbb{R}$.

Let (Ψ, R) a particular solution, with the corresponding θ such that

$$\Psi = Li + \Phi \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$$

and $\omega := \dot{\theta}$.

Observability analysis

For inputs (u, i), the model is

$$\begin{cases} \dot{x} = u - x_3\, i \\ \dot{x}_3 = 0 \\ y = 0 = x^\top x - 2Li^\top x + L^2|i|^2 - \Phi^2 \end{cases}$$

with state $x = (x_1, x_2) \in \mathbb{R}^2$ and $x_3 \in \mathbb{R}$.

Let (Ψ, R) a particular solution, with the corresponding θ such that

$$\Psi = Li + \Phi \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$$

and $\omega := \dot{\theta}$.

Denote $i_{dq} = (i_d, i_q) = \mathcal{R}(-\theta) i_{dq}$ and $u_{dq} = (u_d, u_q) = \mathcal{R}(-\theta) u_{dq}$.

Does there exist another solution $t \mapsto (x(t), x_3(t))$ giving the same output ?

Theorem 2 (Finite number of indistinguishable trajectories).

There exist at most 6 indistinguishable solutions, unless

a) $\omega(t) = 0$ *for all* t,

or

b) *there exists* t *such that* $\omega(t) \neq 0$, *but for all* t, $i_d(t) = 0$, *and* $\frac{\omega}{i_q}$ *is constant,*

where there exists an infinite number of indistinguishable solutions.

Theorem 2 (Finite number of indistinguishable trajectories).

There exist at most 6 indistinguishable solutions, unless

a) $\omega(t) = 0$ *for all* t,

or

b) *there exists* t *such that* $\omega(t) \neq 0$*, but for all* t, $i_d(t) = 0$*, and* $\frac{\omega}{i_q}$ *is constant,*

where there exists an infinite number of indistinguishable solutions.

In particular, at each time t, how many solutions in (x, x_3) may the following equation have

$$H_3(x, x_3, t) = 0 \ = \ \left(h(x,t) \ , \ \dot{\overbrace{h(x,t)}} \ , \ \ddot{\overbrace{h(x,t)}} \right) \ ?$$

Theorem 2 (Finite number of indistinguishable trajectories).

There exist at most 6 indistinguishable solutions, unless

a) $\omega(t) = 0$ *for all* t,

or

b) *there exists* t *such that* $\omega(t) \neq 0$, *but for all* t, $i_d(t) = 0$, *and* $\frac{\omega}{i_q}$ *is constant,*

where there exists an infinite number of indistinguishable solutions.

In particular, at each time t, how many solutions in (x, x_3) may the following equation have

$$H_3(x, x_3, t) = 0 \ = \ \left(h(x,t) \ , \ \dot{\overbrace{h(x,t)}} \ , \ \ddot{\overbrace{h(x,t)}} \right) \ ?$$

If $\omega(t) \neq 0$ and $i_d(t) \neq 0$, there are as many solutions (x, x_3) as the number of distinct real roots of a polynomial of degree six :

$$P(x_3, t) = \omega(t)^6 \Phi^6 \left[\left(\mathbf{1} - \frac{(R - x_3)}{\omega(t)\Phi} \left(\dot{\overbrace{\left(\frac{i_d}{\omega}\right)}}(t) - 2 i_q(t) \right) + \frac{(R - x_3)^2}{\omega(t)^2 \Phi^2} \mu(t) |i(t)|^2 \right)^2 \right.$$
$$\left. - \left(\mathbf{1} + \frac{(R - x_3)}{\omega(t)\Phi} 2 i_q + \frac{(R - x_3)^2}{\omega(t)^2 \Phi^2} |i(t)|^2 \right)^3 \right]$$

Theorem 2 (Finite number of indistinguishable trajectories).

There exist at most 6 indistinguishable solutions, unless

a) $\omega(t) = 0$ *for all* t,

or

b) *there exists* t *such that* $\omega(t) \neq 0$, *but for all* t, $i_d(t) = 0$, *and* $\frac{\omega}{i_q}$ *is constant,*

where there exists an infinite number of indistinguishable solutions.

In particular, at each time t, how many solutions in (x, x_3) may the following equation have

$$H_3(x, x_3, t) = 0 \;=\; \left(h(x,t)\ ,\ \dot{\overbrace{h(x,t)}}\ ,\ \ddot{\overbrace{h(x,t)}} \right) \ ?$$

If $\omega(t) \neq 0$ and $i_d(t) \neq 0$, there are as many solutions (x, x_3) as the number of distinct real roots of a polynomial of degree six :

$$P(x_3, t) = \omega(t)^6 \Phi^6 \left[\left(1 - \frac{(R - x_3)}{\omega(t)\Phi} \left(\dot{\overbrace{\left(\frac{i_d}{\omega}\right)}}(t) - 2 i_q(t) \right) + \frac{(R - x_3)^2}{\omega(t)^2 \Phi^2} \mu(t) |i(t)|^2 \right)^2 \right.$$
$$\left. - \left(1 + \frac{(R - x_3)}{\omega(t)\Phi} 2 i_q + \frac{(R - x_3)^2}{\omega(t)^2 \Phi^2} |i(t)|^2 \right)^3 \right]$$

If besides $i_q(t) \neq -\dot{\overbrace{\left(\frac{i_d}{\omega}\right)}}(t)$, then $P(x_3, t)$ admits at least two distinct real roots.

Particular case : ω, i_d and i_q constant

Theorem 3 (Case of constant ω, i_d and i_q).

Assume $\omega \neq 0$ and $i_d \neq 0$.

There exist exactly two indistinguishable solutions (x, x_3), of the form (Ψ, R) and (Ψ_δ, R_δ) with

$$R_\delta = R + \frac{2\Phi\omega\, i_q}{|i|^2} \quad , \quad i_{d,\delta} = i_d \quad , \quad i_{q,\delta} = -i_q \ ,$$
$$\omega_\delta = \dot\theta_\delta = \dot\theta = \omega \ .$$

Particular case : ω, i_d and i_q constant

Theorem 3 (Case of constant ω, i_d and i_q).

Assume $\omega \neq 0$ and $i_d \neq 0$.

There exist exactly two indistinguishable solutions (x, x_3), of the form (Ψ, R) and (Ψ_δ, R_δ) with

$$R_\delta = R + \frac{2\Phi\omega\, i_q}{|i|^2} \quad , \quad i_{d,\delta} = i_d \quad , \quad i_{q,\delta} = -i_q \ ,$$
$$\omega_\delta = \dot{\theta}_\delta = \dot{\theta} = \omega \ .$$

Besides, if $(\hat{R}, \hat{\theta})$ is one of the solutions (R, θ) or $(R_\delta, \theta_\delta)$, then the other solution is

$$\left(\hat{R} + \frac{2\Phi\hat{\omega}\,\widehat{i_q}}{|i|^2} \ , \ \hat{\theta} + \arctan_2\left(2\frac{\widehat{i_q}\,\widehat{i_d}}{|i|^2}, 1 - 2\frac{\widehat{i_q}^{\,2}}{|i|^2} \right) \right) \ .$$

Observer design

Find a transformation $(x, x_3, t) \mapsto T_\lambda(x, x_3, t)$ mapping

$$\left\{ \begin{array}{rcl} \dot{x} & = & u - x_3\, i \\ \dot{x}_3 & = & 0 \end{array} \right.$$

into a Hurwitz form

$$\dot{\xi}_\lambda = -\lambda \xi_\lambda + y \qquad , \quad y = x^\top x - 2Li^\top x + L^2|i|^2 - \Phi^2$$

Observer design

Find a transformation $(x, x_3, t) \mapsto T_\lambda(x, x_3, t)$ mapping

$$\left\{ \begin{array}{rcl} \dot{x} & = & u - x_3\, i \\ \dot{x}_3 & = & 0 \end{array} \right.$$

into a Hurwitz form

$$\dot{\xi}_\lambda = -\lambda \xi_\lambda + y \qquad , \quad y = x^\top x - 2Li^\top x + L^2|i|^2 - \Phi^2$$

We find

$$T_\lambda(x, x_3, t) = \lambda^2\, x^\top x + \lambda\, c_\lambda(t)^\top x + \lambda\, x_3\, b_\lambda(t)^\top x + a_\lambda(t) x_3 + d_\lambda(t) x_3^2 - e_\lambda$$

with

$$\begin{array}{rcl} \dot{a_\lambda} & = & -\lambda\,(a_\lambda - c_\lambda^\top i + b_\lambda^\top u) \\ \dot{b_\lambda} & = & -\lambda\,(b_\lambda - 2i) \\ \dot{c_\lambda} & = & -\lambda\,(c_\lambda + 2u + 2\lambda L i) \\ \dot{d_\lambda} & = & -\lambda\,(d_\lambda - b_\lambda^\top i) \\ \dot{e_\lambda} & = & -\lambda\,(e_\lambda - c_\lambda^\top u + \lambda^2 L^2 |i|^2 - \lambda^2 \Phi^2)\ . \end{array}$$

Take

$$\begin{aligned} T(x, x_3, t) &= \begin{pmatrix} T_{\lambda_1}(x, x_3, t) \\ T_{\lambda_2}(x, x_3, t) \\ T_{\lambda_3}(x, x_3, t) \end{pmatrix} \\ &= m(\Lambda)\, x^\top x + M(\Lambda)\,(c(t) + x_3\, b(t))\, x + a(t)\, x_3 + d(t)\, x_3^2 - e(t) \end{aligned} \tag{2}$$

for $\lambda_k \in \mathbb{C}$ with $\Re(\lambda_k) > 0$.

Take

$$T(x, x_3, t) = \begin{pmatrix} T_{\lambda_1}(x, x_3, t) \\ T_{\lambda_2}(x, x_3, t) \\ T_{\lambda_3}(x, x_3, t) \end{pmatrix} \tag{2}$$

$$= m(\Lambda)\, x^\top x + M(\Lambda)\,(c(t) + x_3\, b(t))\, x + a(t)\, x_3 + d(t)\, x_3^2 - e(t)$$

for $\lambda_k \in \mathbb{C}$ with $\Re(\lambda_k) > 0$.

Since $y(t) = 0$ for all t, a possible solution is $\xi_\lambda(t) = 0$ for all t and for all λ, so we solve online at each time

$$T(\hat{x}, \hat{x}_3, t) = 0$$

Take

$$T(x, x_3, t) = \begin{pmatrix} T_{\lambda_1}(x, x_3, t) \\ T_{\lambda_2}(x, x_3, t) \\ T_{\lambda_3}(x, x_3, t) \end{pmatrix} \tag{2}$$
$$= m(\Lambda)\, x^\top x + M(\Lambda)\,(c(t) + x_3\, b(t))\, x + a(t)\, x_3 + d(t)\, x_3^2 - e(t)$$

for $\lambda_k \in \mathbb{C}$ with $\Re(\lambda_k) > 0$.

Since $y(t) = 0$ for all t, a possible solution is $\xi_\lambda(t) = 0$ for all t and for all λ, so we solve online at each time

$$T(\hat{x}, \hat{x}_3, t) = 0$$

Number of solutions proved to be equal to the number of solutions to $H_3(x, x_3, t) = 0$ in the differential observability analysis for sufficiently fast eigenvalues !

Take

$$T(x, x_3, t) = \begin{pmatrix} T_{\lambda_1}(x, x_3, t) \\ T_{\lambda_2}(x, x_3, t) \\ T_{\lambda_3}(x, x_3, t) \end{pmatrix} \tag{2}$$

$$= m(\Lambda)\, x^\top x + M(\Lambda)\,(c(t) + x_3\, b(t))\, x + a(t)\, x_3 + d(t)\, x_3^2 - e(t)$$

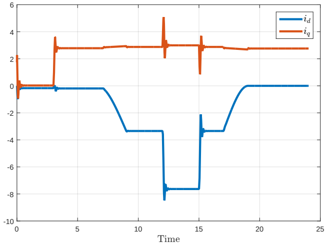
for $\lambda_k \in \mathbb{C}$ with $\Re(\lambda_k) > 0$.

Since $y(t) = 0$ for all t, a possible solution is $\xi_\lambda(t) = 0$ for all t and for all λ, so we solve online at each time

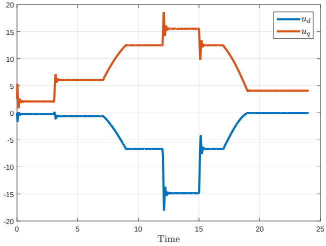
$$T(\hat{x}, \hat{x}_3, t) = 0$$

Number of solutions proved to be equal to the number of solutions to $H_3(x, x_3, t) = 0$ in the differential observability analysis for sufficiently fast eigenvalues !

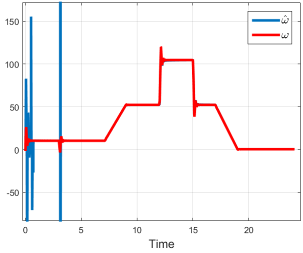
$\Longrightarrow$ Elimination of $x^\top x$, then x and finally grid in x_3 !



(a) $i_{dq} = \mathcal{R}(-\theta)i$

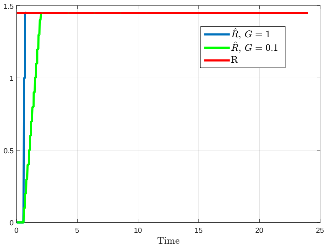


(b) $u_{dq} = \mathcal{R}(-\theta)u$

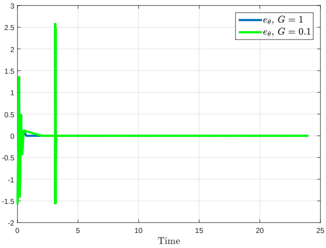


(c) Rotation speed $\omega = \dot{\theta}$ and estimated rotation speed $\hat{\omega} = \dot{\hat{\theta}}$.

Additional knowledge : $i_q > 0$ because motor mode $\Rightarrow$ removes indistinguishability

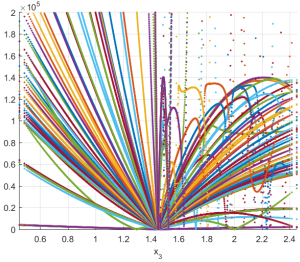


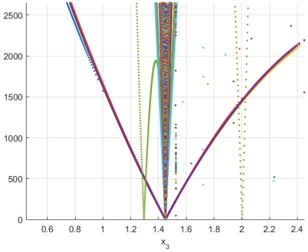
(d) $\hat{R}$



(e) $e_\theta = \hat{\theta} - \theta$ [rad]

Figure – Results of the observer algorithm with $\lambda_1 = 20$, $\lambda_2 = 30$, $\lambda_3 = 40$, $dt_R = 0.1$, and two grids with amplitude $G = 1$ and $G = 0.1$ respectively. The estimation starts at $\bar{t} = 0.5$.





(b) Zoom around $R = 1.45$

Figure – Plot of the criterion $|J(\cdot, t)|$ on the grid with $G = 1$ at each iteration where $\hat{R}$ is updated, i.e. every $dt_R = 0.1$.

Experimental data with constant i_d, i_q and ω

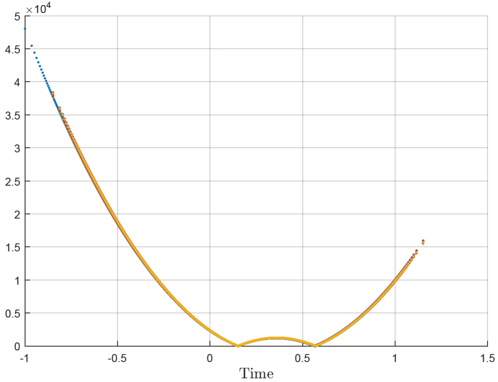


Figure – Real data : Plot of the criterion $|J(\cdot, t)|$ on the grid with $G = 1$ at each iteration where $\hat{R}$ is updated, i.e. every $dt_R = 0.1$.

Conclusion

- KKL observer design possible for a non observable system
- the inversion of the transformation enables to recover the different possibilities predicted by differential observability analysis
- comparison with high gain :
    - transformation independent from input derivatives
    - non observability prevents having a triangular normal form (set-valued last line)

Contents

KKL design

x-coordinates in $\mathbb{R}^{d_x}$ $\qquad$ ξ-coordinates in $\mathbb{R}^{d_\xi}$

$$
\begin{tikzcd}
\begin{array}{rcl} \dot{x} & = & f(x) \\ y & = & h(x) \end{array} \arrow[r, "T"] \arrow[d, dashed] & \dot{\xi} = A\,\xi + B\,y \arrow[d] \\
\hat{x} = \ ? & \dot{\hat{\xi}} = A\,\hat{\xi} + B\,y \arrow[l, "\mathcal{T}"']
\end{tikzcd}
$$

A Hurwitz

Existence of a left-inverse $\mathcal{T}$ requires observability !

$\Rightarrow$ What if the system is not observable ?

$\Rightarrow$ Estimate only the observable part !

Functional KKL design

Consider a continuous map $q : \mathbb{R}^{d_x} \to \mathbb{R}^{d_q}$.

Goal : estimate $z := q(x)$

x-coordinates in $\mathbb{R}^{d_x}$ | ξ-coordinates in $\mathbb{R}^{d_\xi}$

$$
\begin{tikzcd}
\begin{array}{rcl} \dot{x} &=& f(x) \\ y &=& h(x) \end{array} \arrow[r, "T"] \arrow[d, dashed] & \dot{\xi} = A\,\xi + B\,y \arrow[d] \\
\widehat{z} = \text{?} & \dot{\hat{\xi}} = A\,\hat{\xi} + B\,y \arrow[l, "\mathcal{T}"']
\end{tikzcd}
$$

A Hurwitz

Existence of T doesn't change!

Can we show existence of $\mathcal{T}$ such that $\lim_{t\to+\infty} |q(x) - \hat{z}| = 0$ under observability of $q(x)$?

Functional KKL design

Consider a continuous map $q : \mathbb{R}^{d_x} \to \mathbb{R}^{d_q}$.

Goal : estimate $z := q(x)$

x-coordinates in $\mathbb{R}^{d_x}$ | ξ-coordinates in $\mathbb{R}^{d_\xi}$

$$\begin{tikzcd}
\begin{array}{rcl} \dot{x} & = & f(x) \\ y & = & h(x) \end{array} \arrow[r, "T"] \arrow[d, dashed] & \dot{\xi} = A\xi + B\,y \arrow[d] \\
\hat{z} = \ ? & \dot{\hat{\xi}} = A\hat{\xi} + B\,y \arrow[l, "\mathcal{T}"']
\end{tikzcd}$$

A Hurwitz

Existence of T doesn't change!

Can we show existence of $\mathcal{T}$ such that $\lim_{t\to+\infty} |q(x) - \hat{z}| = 0$ under observability of $q(x)$?

Assumption 1 (Bounded trajectories).

The trajectories of $\dot{x} = f(x)$ to be estimated evolve in a compact set $\mathcal{X}$.

Existence of T

Goal : Find a transformation $x \mapsto T(x)$ such that $\xi = T(x)$ verifies

$$\dot{\xi} = A\xi + B\,y \quad , \quad A \text{ Hurwitz}$$

Existence of T

Goal : Find a transformation $x \mapsto T(x)$ such that $\xi = T(x)$ verifies

$$\dot{\xi} = A\xi + B\,y \quad , \quad A \text{ Hurwitz}$$

$\Rightarrow$ We want $\dot{\overbrace{T(x)}} = A\,T(x) + B\,h(x)$

Existence of T

Goal : Find a transformation $x \mapsto T(x)$ such that $\xi = T(x)$ verifies

$$\dot{\xi} = A\xi + B\,y \quad , \quad A \text{ Hurwitz}$$

$\Rightarrow$ We want $\dot{\overbrace{T(x)}} = A\,T(x) + B\,h(x)$

=> Find a solution to the PDE

$$\frac{\partial T}{\partial x}(x) f(x) = A\,T(x) + B\,h(x) \tag{3}$$

Existence of T

Goal : Find a transformation $x \mapsto T(x)$ such that $\xi = T(x)$ verifies

$$\dot{\xi} = A\xi + B\,y \quad , \quad A \text{ Hurwitz}$$

$\Rightarrow$ We want $\overset{\cdot}{\overbrace{T(x)}} = A\,T(x) + B\,h(x)$

=> Find a solution to the PDE

$$\frac{\partial T}{\partial x}(x) f(x) = A\,T(x) + B\,h(x) \tag{3}$$

Theorem 4 (Existence of T).

There exists $\ell > 0$ such that for any $A \in \mathbb{R}^{d_\xi \times d_\xi}$ with $\mathrm{Re}(\mathrm{eig}(A)) < -\ell$, and any $B \in \mathbb{R}^{d_\xi \times d_y}$, the map

$$T(x) = \int_{-\infty}^{0} e^{-As} B\,h(\check{X}(x,s))ds$$

is C^1 and verifies PDE (3), where $s \mapsto \check{X}(x,s)$ is the solution initialized at x of $\dot{x} = \check{f}(x)$ with $\check{f}$ bounded equal to f on $\mathcal{X}$.

Injectivity of T with respect to q

Let an open set $\mathcal{S}$ such that $\mathcal{X} \subset \mathcal{S}$.

Assumption 2 (Backward $\mathcal{O}$-distinguishability with respect to q).

There exist $\delta_d > \delta_\Upsilon$*, such that for each pair* (x_a, x_b) *in* $\mathcal{O} + \delta_\Upsilon$ *verifying* $q(x_a) \neq q(x_b)$*, there exists* $t \in (\max\{\sigma^-_{\mathcal{O}+\delta_d}(x_a), \sigma^-_{\mathcal{O}+\delta_d}(x_b)\}, 0]$*, such that*

$$h(X(x_a, t)) \neq h(X(x_b, t)).$$

Injectivity of T with respect to q

Let an open set $\mathcal{S}$ such that $\mathcal{X} \subset \mathcal{S}$.

Assumption 2 (Backward $\mathcal{O}$-distinguishability with respect to q).

There exist $\delta_d > \delta_\Upsilon$, such that for each pair (x_a, x_b) in $\mathcal{O} + \delta_\Upsilon$ verifying $q(x_a) \neq q(x_b)$, there exists $t \in (\max\{\sigma^-_{\mathcal{O}+\delta_d}(x_a), \sigma^-_{\mathcal{O}+\delta_d}(x_b)\}, 0]$, such that

$$h(X(x_a, t)) \neq h(X(x_b, t)).$$

Theorem 5 (Injectivity with respect to q).

Assume the system is backward $\mathcal{O}$-distinguishable with respect to q.

Then, there exists $\ell > 0$ and a set S of zero-Lebesgue measure in $\mathbb{C}_\ell^{d_x+1}$ with

$$\mathbb{C}_\ell = \{z \in \mathbb{C} \ : \ \mathrm{Re}(z) < -\ell\}$$

such that for any diagonalizable matrix $A_0 \in \mathbb{C}^{(d_x+1)\times(d_x+1)}$ and any $B_0 \in \mathbb{C}^{d_x+1}$ such that $\mathrm{eig}(A_0) \in \mathbb{C}_\ell^{d_x+1} \setminus S$ and (A_0, B_0) controllable,

the map T defined for $A = A_0 \otimes I_{d_y}$ and $B = B_0 \otimes I_{d_y}$ is injective with respect to q on $\mathcal{X}$, i.e, for all $(x_a, x_b) \subset \mathcal{X} \times \mathcal{X}$,

$$T(x_a) = T(x_b) \implies q(x_a) = q(x_b) \ .$$

Left-inverse and convergence

Taking real and imaginary parts, $T : \mathbb{R}^{d_x} \to \mathbb{R}^{d_\xi}$ with $d_\xi = 2(d_x + 1)d_y$, such that for all $(x_a, x_b) \in \mathcal{X} \times \mathcal{X}$,

$$T(x_a) = T(x_b) \implies q(x_a) = q(x_b)\ .$$

Left-inverse and convergence

Taking real and imaginary parts, $T : \mathbb{R}^{d_x} \to \mathbb{R}^{d_\xi}$ with $d_\xi = 2(d_x + 1)d_y$, such that for all $(x_a, x_b) \in \mathcal{X} \times \mathcal{X}$,

$$T(x_a) = T(x_b) \implies q(x_a) = q(x_b) \ .$$

Therefore, there exists $\mathcal{T} : \mathbb{R}^{d_\xi} \to \mathbb{R}^{d_x}$ and ρ of class $\mathcal{K}$ such that

- $\mathcal{T}(T(x)) = q(x)$ for all $x \in \mathcal{X}$
- $|\mathcal{T}(\xi_a) - \mathcal{T}(\xi_b)| \leq \rho(|\xi_a - \xi_b|)$ for all $(\xi_a, \xi_b) \in \mathbb{R}^{d_\xi} \times \mathbb{R}^{d_\xi}$

Left-inverse and convergence

Taking real and imaginary parts, $T : \mathbb{R}^{d_x} \to \mathbb{R}^{d_\xi}$ with $d_\xi = 2(d_x + 1)d_y$, such that for all $(x_a, x_b) \in \mathcal{X} \times \mathcal{X}$,

$$T(x_a) = T(x_b) \implies q(x_a) = q(x_b) \ .$$

Therefore, there exists $\mathcal{T} : \mathbb{R}^{d_\xi} \to \mathbb{R}^{d_x}$ and ρ of class $\mathcal{K}$ such that

- $\mathcal{T}(T(x)) = q(x)$ for all $x \in \mathcal{X}$
- $|\mathcal{T}(\xi_a) - \mathcal{T}(\xi_b)| \leq \rho(|\xi_a - \xi_b|)$ for all $(\xi_a, \xi_b) \in \mathbb{R}^{d_\xi} \times \mathbb{R}^{d_\xi}$

$\Rightarrow$ Uniformly continuous left-inverse of T with respect to q !

Left-inverse and convergence

Taking real and imaginary parts, $T : \mathbb{R}^{d_x} \to \mathbb{R}^{d_\xi}$ with $d_\xi = 2(d_x+1)d_y$, such that for all $(x_a, x_b) \in \mathcal{X} \times \mathcal{X}$,

$$T(x_a) = T(x_b) \implies q(x_a) = q(x_b) \ .$$

Therefore, there exists $\mathcal{T} : \mathbb{R}^{d_\xi} \to \mathbb{R}^{d_x}$ and ρ of class $\mathcal{K}$ such that

- $\mathcal{T}(T(x)) = q(x)$ for all $x \in \mathcal{X}$
- $|\mathcal{T}(\xi_a) - \mathcal{T}(\xi_b)| \leq \rho(|\xi_a - \xi_b|)$ for all $(\xi_a, \xi_b) \in \mathbb{R}^{d_\xi} \times \mathbb{R}^{d_\xi}$

$\Rightarrow$ Uniformly continuous left-inverse of T with respect to q !

Since $x(t) \in \mathcal{X}$ for all $t \geq 0$, taking $\xi_a = \hat{\xi}$ and $\xi_b = \xi = T(x)$,

$$|\mathcal{T}(\hat{\xi}(t)) - x(t)| \leq \rho(|\hat{\xi}(t) - \xi(t)|)$$

so that

$$\lim_{t \to +\infty} |\hat{z}(t) - q(x(t))| = 0$$

Functional KKL

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

Theorem 6 (Functional KKL).

Assume

- *the trajectories of interest remain in a compact set* $\mathcal{X}$
- *the system is backward distinguishable with respect to* q.

Then, there exists a map $\mathcal{T} : \mathbb{R}^{d_\xi} \to \mathbb{R}^{d_x}$, *a matrix* $A \in \mathbb{R}^{d_\xi \times d_\xi}$ *and* $B \in \mathbb{R}^{d_\xi \times d_y}$ *with* $d_\xi = 2(d_x + 1)d_y$ *such that*

$$\dot{\hat{\xi}} = A\hat{\xi} + B\,y$$
$$\hat{z} = \mathcal{T}(\hat{\xi})$$

is a functional KKL observer with respect to q, *i.e., for any* $\hat{\xi}(0)$,

$$\lim_{t\to+\infty} |\hat{z}(t) - q(x(t))| = 0$$

Remarks :

- It is actually sufficient to take $d_\xi = (2d_x + 1)d_y$ and the result holds for "almost any" choice of Hurwitz $A_0 \in \mathbb{R}^{(2d_x+1)\times(2d_x+1)}$ and any B making (A, B) controllable.

Remarks :

- It is actually sufficient to take $d_\xi = (2d_x + 1)d_y$ and the result holds for "almost any" choice of Hurwitz $A_0 \in \mathbb{R}^{(2d_x+1)\times(2d_x+1)}$ and any B making (A, B) controllable.

- The result extends to **time-varying systems** with $q(x, t)$ continuous but for $d_\xi = (2d_x + 2)d_y$ and with time-varying $T(x, t)$ and $\mathcal{T}(x, t)$. But uniform injectivity in time not guaranteed for convergence.

Remarks :

- It is actually sufficient to take $d_\xi = (2d_x + 1)d_y$ and the result holds for "almost any" choice of Hurwitz $A_0 \in \mathbb{R}^{(2d_x+1)\times(2d_x+1)}$ and any B making (A, B) controllable.

- The result extends to **time-varying systems** with $q(x, t)$ continuous but for $d_\xi = (2d_x + 2)d_y$ and with time-varying $T(x, t)$ and $\mathcal{T}(x, t)$. But uniform injectivity in time not guaranteed for convergence.

- In presence of **inputs**, the time-varying paradigm applies, but with $T_u(x, t)$ and $\mathcal{T}_u(x, t)$ depending implicitly on the the input :
  - -> either we can compute this dependence explicitly (for instance via filters as in the context of electrical machines)
  - -> or we design T and $\mathcal{T}$ for a particular class of inputs generated by

$$\dot{w} = s(w) \quad , \quad u = l(w)$$

via a functional observer on the extended system with state (x, w) and extended output $(h(x, l(w)), l(w))$, leading to

$$\dot{\hat{\xi}} = A\hat{\xi} + B\begin{pmatrix} y \\ u \end{pmatrix}$$

Remarks :

- It is actually sufficient to take $d_\xi = (2d_x + 1)d_y$ and the result holds for "almost any" choice of Hurwitz $A_0 \in \mathbb{R}^{(2d_x+1)\times(2d_x+1)}$ and any B making (A, B) controllable.

- The result extends to **time-varying systems** with $q(x, t)$ continuous but for $d_\xi = (2d_x + 2)d_y$ and with time-varying $T(x, t)$ and $\mathcal{T}(x, t)$. But uniform injectivity in time not guaranteed for convergence.

- In presence of **inputs**, the time-varying paradigm applies, but with $T_u(x, t)$ and $\mathcal{T}_u(x, t)$ depending implicitly on the the input :
  - -> either we can compute this dependence explicitly (for instance via filters as in the context of electrical machines)
  - -> or we design T and $\mathcal{T}$ for a particular class of inputs generated by

$$\dot{w} = s(w) \quad , \quad u = l(w)$$

  via a functional observer on the extended system with state (x, w) and extended output $(h(x, l(w)), l(w))$, leading to

$$\dot{\hat{\xi}} = A\hat{\xi} + B\begin{pmatrix} y \\ u \end{pmatrix}$$

- When the maps T and $\mathcal{T}$ are not known explicitly, we can compute them numerically via neural networks ⇒ "numerical KKL"

L. Da Costa Ramos, F. Di Meglio, L. F. Figueira Da Silva, V. Morgenthaler, P. Bernard, *Numerical design of Luenberger observers for nonlinear systems*, CDC, 2020

Contents

Differential non-observability

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

The system is always instantaneously backward-distinguishable with respect to the map

$$q(x,t) = \begin{pmatrix} h(x,t) \\ L_f h(x,t) \\ \vdots \\ L_f^{m-1} h(x,t) \end{pmatrix}$$

for any $m \in \mathbb{N}$.

$\Rightarrow$ $q(x,t)$ can always be estimated via functional KKL observer !

$$\dot{\hat{\xi}} = A\hat{\xi} + B\,y \qquad , \qquad \hat{z} = \mathcal{T}(\hat{\xi},t)$$

$\Rightarrow$ Recover all possible $\hat{x}$ such that

$$\hat{x} \in \mathrm{Argmin}_{x\in\mathcal{X}} \, |q(x,t) - \hat{z}|^2$$

$\Rightarrow$ No need for normal triangular form unlike in high-gain design !

Indistinguishability

Given $\breve{f}$ bounded such that $\breve{f} = f$ on an open set $\mathcal{O}$ containing $\mathcal{X}$:

Consider the set $\breve{\mathcal{I}} \subset \mathbb{R}^{d_x} \times \mathbb{R}^{d_x}$ of pairs (x_a, x_b) backward-indistinguishable for $\dot{x} = \breve{f}(x)$, i.e., such that

$$h(\breve{X}(x_a), t)) = h(\breve{X}(x_b), t)) \quad \forall t \leq 0$$

Then, by definition, for all $(x_a, x_b) \in \breve{\mathcal{I}}$,

$$T(x_a) = \int_{-\infty}^{0} e^{-As} B \, h(\breve{X}(x_a, s)) ds = \int_{-\infty}^{0} e^{-As} B \, h(\breve{X}(x_b, s)) ds = T(x_b)$$

$\Rightarrow$ Indistinguishable by T !

Indistinguishability

Given $\breve{f}$ bounded such that $\breve{f} = f$ on an open set $\mathcal{O}$ containing $\mathcal{X}$:

Consider the set $\breve{\mathcal{I}} \subset \mathbb{R}^{d_x} \times \mathbb{R}^{d_x}$ of pairs (x_a, x_b) backward-indistinguishable for $\dot{x} = \breve{f}(x)$, i.e., such that

$$h(\breve{X}(x_a), t)) = h(\breve{X}(x_b), t)) \quad \forall t \leq 0$$

Then, by definition, for all $(x_a, x_b) \in \breve{\mathcal{I}}$,

$$T(x_a) = \int_{-\infty}^{0} e^{-As} B\, h(\breve{X}(x_a, s))ds = \int_{-\infty}^{0} e^{-As} B\, h(\breve{X}(x_b, s))ds = T(x_b)$$

$\Rightarrow$ Indistinguishable by T !

Conversely, applying the proof of KKL on the open set $(\mathbb{R}^{d_x} \times \mathbb{R}^{d_x}) \setminus \breve{\mathcal{I}}$ instead of

$$\{(x_a, x_b) \in \mathbb{R}^{d_x} \times \mathbb{R}^{d_x} \;:\; x_a \neq x_b\}$$

should show that for "almost all" choice of (A, B) of dimension $(2d_x + 1)d_y$,

$$T(x_a) = T(x_b) \implies (x_a, x_b) \in \breve{\mathcal{I}}$$

i.e., T is injective with respect to the distinguishable states !

Conclusion

- KKL design well suited for non observable systems since the dynamics are well-defined and everything is in the inversion of the transformation
- very general theoretical answer to the observation problem under weak assumptions
- **problem** : numerical computation of T and $\mathcal{T}$ challenging in practice !
- how generic is the case of "finite-indistinguishability" ?