Proceedings of the 36th
Conference on Decision & Control
San Diego, California USA • December 1997

Observer design in the behavioral approach

Maria Elena Valcher
Dip. di Elettronica ed Informatica, Univ. di Padova
via Gradenigo 6a - 35131 Padova, ITALY
e-mail:meme@dei.unipd.it

Jan C.Willems
Mathematics Institute, University of Groningen
P.O. Box 800, 9700 AV Groningen, The Netherlands
e-mail:J.C.Willems@math.rug.nl

Abstract

Given a dynamical system whose describing variables either can be measured or are unaccessible, the problem of estimating the trajectories of these latter from the knowledge of the former is investigated. Conditions for the existence of asymptotic observers, in particular (strictly) proper ones, as well as a complete parametrization of all possible observers, are provided. These results are then specialized to the case of state-space models.

Finally, the above issues are also analysed in the context of dynamical systems with latent variables.

Keywords: Observers, behaviors, estimation.

1 Introduction

In the last decade the behavioral point of view [4] has received an increasingly broader acceptance as an approach for modelling dynamical systems, and now is generally viewed as a foundational framework for analysing the trajectories a system produces according to its evolution laws. One of the reasons of its success has to be looked for in the fact that it does not assume the input/output point of view for describing how a system interacts with its environment, but focuses its interest on the set of system trajectories, the *behavior*, and hence on the mathematical model describing the relations among the system variables. By assuming this point of view, important aspects of the classical system theory have been translated and solved, thus leading to interesting results, which are powerful generalizations of well-known theorems obtained within the input/output or state-space contexts.

Although several issues have already been analysed in some detail, the important question of estimating some system variables, not available from measurements, from others, which are measured, is still unsolved and the aim of this contribution is to provide an analysis of the problem.

2 Observability and detectability

Consider a dynamical system $\Sigma = (\mathbb{R}, \mathbb{R}^{\mathtt{w}_1+\mathtt{w}_2}, \mathfrak{B})$, whose behavior trajectories $(\mathbf{w}_1, \mathbf{w}_2)$ satisfy a system of differential equations

$$R_2(\frac{d}{dt})\mathbf{w}_2 = R_1(\frac{d}{dt})\mathbf{w}_1, \qquad (1)$$

with $R_i \in \mathbb{R}[\xi]^{\mathtt{r}\times\mathtt{w}_i}$ and $\mathtt{w}_i := \dim \mathbf{w}_i$, $i = 1, 2$. Throughout the paper we make the assumption that $[\, R_2 \quad -R_1 \,]$ has full row rank $\mathtt{r}$. Also, we assume that the trajectories $(\mathbf{w}_1, \mathbf{w}_2)$ belong to $\mathfrak{L}^{loc}(\mathbb{R}, \mathbb{R}^{\mathtt{w}_1+\mathtt{w}_2})$, the space of locally integrable functions from $\mathbb{R}$ to $\mathbb{R}^{\mathtt{w}_1+\mathtt{w}_2}$, equipped with the topology which ensures that $(\mathbf{w}_{1n}, \mathbf{w}_{2n}) \xrightarrow[n\to\infty]{} 0$ iff $\int_{t_0}^{t_1} \|(\mathbf{w}_{1n}, \mathbf{w}_{2n})\|\, dt \xrightarrow[n\to\infty]{} 0$ for all $t_0, t_1 \in \mathbb{R}$.

If $\mathbf{w}_1$ is measured and $\mathbf{w}_2$ is unknown, the goal is that of determining necessary and sufficient conditions for the existence of an estimator of $\mathbf{w}_2$ from the knowledge of $\mathbf{w}_1$, such that the estimation error goes to zero asymptotically. As a first step we introduce observability and detectability notions.

Definition In the dynamical system $\Sigma = (\mathbb{R}, \mathbb{R}^{\mathtt{w}_1+\mathtt{w}_2}, \mathfrak{B})$, with behavior $\mathfrak{B}$ described by (1), $\mathbf{w}_2$ is said to be

• *observable* from $\mathbf{w}_1$, if $(\mathbf{w}_1, \mathbf{w}_2), (\mathbf{w}_1, \bar{\mathbf{w}}_2) \in \mathfrak{B}$ implies $\mathbf{w}_2 = \bar{\mathbf{w}}_2$;

• *detectable* from $\mathbf{w}_1$, if $(\mathbf{w}_1, \mathbf{w}_2), (\mathbf{w}_1, \bar{\mathbf{w}}_2) \in \mathfrak{B}$ implies $\mathbf{w}_2(t) - \bar{\mathbf{w}}_2(t) \xrightarrow[t\to\infty]{} 0$.

Proposition 2.1 *Given a dynamical system* $\Sigma = (\mathbb{R}, \mathbb{R}^{\mathtt{w}_1+\mathtt{w}_2}, \mathfrak{B})$, *with behavior* $\mathfrak{B}$ *described as in (1),* *i)* $\mathbf{w}_2$ *is observable from* $\mathbf{w}_1$ *if and only if* R_2 *is a right prime polynomial matrix, or, equivalently, if and only if* $R_2(\lambda)$ *has full column rank for all* $\lambda \in \mathbb{C}$*;*

ii) $\mathbf{w}_2$ *is detectable from* $\mathbf{w}_1$ *if and only if* R_2 *is full column rank and the g.c.d. of its maximal order minors is Hurwitz, or, equivalently, if and only if* $R_2(\lambda)$ *has full column rank for all* $\lambda \in \mathbb{C}^+ := \{\xi \in \mathbb{C} : \Re e(\xi) \geq 0\}$.

PROOF (i) Suppose that R_2 is right prime and let U be a unimodular matrix such that $U(\xi)R_2(\xi) = \begin{bmatrix} I_{\mathtt{w}_2} \\ 0 \end{bmatrix}$. Once we set $\begin{bmatrix} N_1(\xi) \\ D_1(\xi) \end{bmatrix} := U(\xi)R_1(\xi)$, then $\mathfrak{B} = \{(\mathbf{w}_1, \mathbf{w}_2) \in \left(\mathbb{R}^{\mathtt{w}_1+\mathtt{w}_2}\right)^{\mathbb{R}} : D_1(\frac{d}{dt})\mathbf{w}_1 = 0$ and $\mathbf{w}_2 = N_1(\frac{d}{dt})\mathbf{w}_1\}$. Consequently, $\mathbf{w}_2$ is observable from $\mathbf{w}_1$. On the other hand, if R_2 were not right prime, its kernel would include some nonzero trajectory $\mathbf{v}$. So, $(\mathbf{w}_1, \mathbf{w}_2) \in \mathfrak{B}$ would imply $(\mathbf{w}_1, \mathbf{w}_2 + \mathbf{v}) \in \mathfrak{B}$, thus ruling out observability.

(ii) Follows the same lines of the previous one, upon realizing that $\mathbf{w}_2$ is detectable from $\mathbf{w}_1$ iff ker $(R_2(\frac{d}{dt}))$ is a stable autonomous behavior. ■

The previous definitions are consistent with the well-known definitions of observability and detectability for state space models. In fact, given an (n-dimensional) state space model, with m inputs and p outputs

$$\frac{d\mathbf{x}}{dt} = F\mathbf{x} + G\mathbf{u}, \tag{2}$$

$$\mathbf{y} = H\mathbf{x} + J\mathbf{u}. \tag{3}$$

the set of its trajectories is equivalently described, in behavioral terms, as the set $\mathfrak{B}$ of all $(\mathbf{x}, \mathbf{u}, \mathbf{y})$ satisfying

$$\begin{bmatrix} \frac{d}{dt}I_{\mathtt{n}} - F \\ H \end{bmatrix} \mathbf{x} = \begin{bmatrix} G & 0 \\ -J & I_{\mathtt{p}} \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{y} \end{bmatrix}. \tag{4}$$

By the previous proposition, $\mathbf{x}$ is observable (detectable) from $(\mathbf{u}, \mathbf{y})$ if and only if rank $\begin{bmatrix} \xi I_{\mathtt{n}} - F \\ H \end{bmatrix} = n, \forall\, \xi \in \mathbb{C} (\forall\, \xi \in \mathbb{C}^+)$. The previous conditions are the well-know observability and detectability PBH tests for state space models.

3 Asymptotic observers design

Definition Consider the dynamical system $\Sigma = (\mathbb{R}, \mathbb{R}^{\mathtt{w}_1+\mathtt{w}_2}, \mathfrak{B})$, with behavior $\mathfrak{B}$ described as in (1), and let P and Q be two polynomial matrices of suitable dimensions. We call

$$Q(\frac{d}{dt})\hat{\mathbf{w}}_2 = P(\frac{d}{dt})\mathbf{w}_1 \tag{5}$$

an *asymptotic observer of* $\mathbf{w}_2$ *from* $\mathbf{w}_1$ if for every $(\mathbf{w}_1, \mathbf{w}_2) \in \mathfrak{B}$ and every $\hat{\mathbf{w}}_2$ s.t. $(\mathbf{w}_1, \hat{\mathbf{w}}_2)$ satisfies (5), we have $\Delta(\frac{d}{dt})(\mathbf{w}_2 - \hat{\mathbf{w}}_2) = 0$ for some Hurwitz matrix $\Delta \in \mathbb{R}[\xi]^{\mathtt{w}_2 \times \mathtt{w}_2}$ (i.e. a nonsingular matrix whose determinant is a Hurwitz polynomial).

The difference variable $\mathbf{e} := \mathbf{w}_2 - \hat{\mathbf{w}}_2$ represents the *estimation error* of the asymptotic observer. Let $\mathfrak{B}_e$ denote its behavior. It is clear that if (5) is an asymptotic observer, then the set of admissible estimation errors is a subset of ker $(\Delta(\frac{d}{dt}))$. In the sequel we will choose Δ so that $\mathfrak{B}_e = \ker\Delta(\frac{d}{dt})$ and refer to the determinant of Δ as to the *error-dynamics characteristic polynomial* (see [5]).

Proposition 3.1 Given a dynamical system $\Sigma = (\mathbb{R}, \mathbb{R}^{\mathtt{w}_1+\mathtt{w}_2}, \mathfrak{B})$, whose behavior $\mathfrak{B}$ is described as in (1), a necessary and sufficient condition for the existence of an asymptotic observer of $\mathbf{w}_2$ from $\mathbf{w}_1$ is that $\mathbf{w}_2$ is detectable from $\mathbf{w}_1$.

PROOF Assume, first, that there exists an asymptotic observer of $\mathbf{w}_2$ from $\mathbf{w}_1$. If $\mathbf{w}_2$ were not detectable from $\mathbf{w}_1$, there would be $(\mathbf{w}_1, \mathbf{w}_2)$ and $(\mathbf{w}_1, \bar{\mathbf{w}}_2)$ in $\mathfrak{B}$ such that $\mathbf{w}_2(t) - \bar{\mathbf{w}}_2(t)$ does not estinguish as t goes to $+\infty$. If $\hat{\mathbf{w}}_2$ is an estimate provided by the asymptotic observer corresponding to $\mathbf{w}_1$, then it should be $\mathbf{w}_2(t) - \hat{\mathbf{w}}_2(t) \to 0$ as well as $\bar{\mathbf{w}}_2(t) - \hat{\mathbf{w}}_2(t) \to 0$, for $t \to +\infty$, and, hence, $[\mathbf{w}_2(t) - \bar{\mathbf{w}}_2(t)] = [\mathbf{w}_2(t) - \hat{\mathbf{w}}_2(t)] - [\bar{\mathbf{w}}_2(t) - \hat{\mathbf{w}}_2(t)]$ should asymptotically estinguish, a contradiction.

To show the converse, assume that R_2 has the required properties. Then, there exists a unimodular matrix U such that $U(\xi)R_2(\xi) = \begin{bmatrix} D_2(\xi) \\ 0 \end{bmatrix}$, with D_2 Hurwitz. Once we set $\begin{bmatrix} N_1(\xi) \\ D_1(\xi) \end{bmatrix} := U(\xi)R_1(\xi)$, we obtain for the behavior the equivalent description

$$D_2(\frac{d}{dt})\mathbf{w}_2 = N_1(\frac{d}{dt})\mathbf{w}_1 \tag{6}$$

$$0 = D_1(\frac{d}{dt})\mathbf{w}_1. \tag{7}$$

Clearly, $D_2(\frac{d}{dt})\hat{\mathbf{w}}_2 = N_1(\frac{d}{dt})\mathbf{w}_1$ is an asymptotic observer of $\mathbf{w}_2$ from $\mathbf{w}_1$, with $\Delta = D_2$. ■

From now on we will assume that $\mathbf{w}_2$ is detectable from $\mathbf{w}_1$ and that the plant Σ is represented as in (6)÷(7), with D_2 Hurwitz and D_1 having full row rank $\mathtt{d}_1$. The following technical lemma, whose simple proof is omitted for sake of brevity, allows to take a first step toward a complete parametrization of all asymptotic observers.

Lemma 3.2 *Given any asymptotic observer (5) there exists an equivalent one with* Q *full row rank.*

Proposition 3.3 *Consider a plant* Σ *whose behavior* $\mathfrak{B}$ *is described as in (6)÷(7), with* D_2 *Hurwitz and* D_1 *full row rank, and let* P *and* Q *be polynomial matrices* *of suitable dimensions, with Q full row rank. Then $Q(\frac{d}{dt})\hat{\mathbf{w}}_2 = P(\frac{d}{dt})\mathbf{w}_1$, is an asymptotic observer of $\mathbf{w}_2$ from $\mathbf{w}_1$ if and only if there exists a Hurwitz matrix $Y \in \mathbb{R}[\xi]^{\mathrm{w}_2 \times \mathrm{w}_2}$ and a polynomial matrix $X \in \mathbb{R}[\xi]^{\mathrm{w}_2 \times \mathrm{d}_1}$ such that*

$$[\,Q(\xi) \quad -P(\xi)\,] = [\,Y(\xi) \quad X(\xi)\,]\begin{bmatrix} D_2(\xi) & -N_1(\xi) \\ 0 & -D_1(\xi) \end{bmatrix}. \quad (8)$$

Moreover, the set $\mathfrak{B}_e$ of all possible error trajectories coincides with ker $\left(Q(\frac{d}{dt})\right)$, *or, equivalently, $\Delta = Q$.*

PROOF Suppose, first, that $Q(\frac{d}{dt})\hat{\mathbf{w}}_2 = P(\frac{d}{dt})\mathbf{w}_1$, with Q full row rank, is an asymptotic observer of $\mathbf{w}_2$ from $\mathbf{w}_1$. If $(\mathbf{w}_1, \mathbf{w}_2)$ is a behavior trajectory and $(\mathbf{w}_1, \hat{\mathbf{w}}_2)$ a trajectory correspondingly provided by the observer, $\mathbf{e} := \mathbf{w}_2 - \hat{\mathbf{w}}_2 \in \mathfrak{B}_e$, and for every $\mathbf{v} \in \ker\left(Q(\frac{d}{dt})\right)$, the trajectory $(\mathbf{w}_1, \hat{\mathbf{w}}_2 + \mathbf{v})$ is also admissible for the observer. Therefore $\mathbf{e} - \mathbf{v}$, and hence $\mathbf{v}$, must be error trajectories, thus showing that $\ker\left(Q(\frac{d}{dt})\right) \subseteq \mathfrak{B}_e$. As the zero error trajectory $\mathbf{e} = 0$ possibly occurs, for every $(\mathbf{w}_1, \mathbf{w}_2) \in \mathfrak{B}$ it must be $Q(\frac{d}{dt})\mathbf{w}_2 = P(\frac{d}{dt})\mathbf{w}_1$, and therefore there exist polynomial matrices X and Y such that (8) holds true. Moreover, as ker $(Q(\frac{d}{dt}))$ must be autonomous and stable, Y must be nonsingular Hurwitz.

Assume, now, that P and Q satisfy (8) for suitable X and Y, with Y Hurwitz. Then, for every $(\mathbf{w}_1, \mathbf{w}_2) \in \mathfrak{B}$ and every estimate $\hat{\mathbf{w}}_2$, correspondingly determined from the observer equations, one gets $Q(\frac{d}{dt})(\mathbf{w}_2 - \hat{\mathbf{w}}_2) = -(XN_1)(\frac{d}{dt})\mathbf{w}_1 = 0$. This immediately proves that $\mathfrak{B}_e$, as a subset of $\ker\left(Q(\frac{d}{dt})\right)$, is a stable autonomous behavior, and hence the observer is an asymptotic one. Moreover, as we have already proved that $\ker\left(Q(\frac{d}{dt})\right) \subseteq \mathfrak{B}_e$, then $\mathfrak{B}_e$ coincides with $\ker\left(Q(\frac{d}{dt})\right)$. ■

REMARK Notice that if $\mathfrak{B}$ is described as in (1), with $\mathbf{w}_2$ detectable from $\mathbf{w}_1$, the asymptotic observers of $\mathbf{w}_2$ are those and only those that can be expressed as in (5) for some polynomial pair $(Q(\xi), P(\xi))$, with $Q(\xi)$ Hurwitz, satisfying $[\,Q(\xi) \quad -P(\xi)\,] = T(\xi)\,[\,R_2(\xi) \quad R_1(\xi)\,]$ for some polynomial $T(\xi)$.

Corollary 3.4 *Consider a dynamical system whose behavior $\mathfrak{B}$ is described as in (6)÷(7), with D_2 Hurwitz and D_1 full row rank. If $d_2(\xi) := \det D_2(\xi)$, then*

i) for every Hurwitz polynomial $\delta \in \mathbb{R}[\xi]$ with $d_2 \mid \delta$, there exists an asymptotic observer with error-dynamics characteristic polynomial $\det \Delta = \delta$;

ii) for every asymptotic observer of $\mathbf{w}_2$ from $\mathbf{w}_1$, the error-dynamics characteristic polynomial $\det \Delta$ is a Hurwitz polynomial satisfying $d_2 \mid \det \Delta$.

PROOF Easily follows from Proposition 3.3. ■

As in the previous section, it is interesting to see how the above results apply to the case of state-space models. Indeed, if we consider the behavior description given in (4), and assume that (F, H) is a detectable pair, we can apply the previous reasonings and obtain an asymptotic state estimator described by

$$Q(\frac{d}{dt})\hat{\mathbf{x}} = P_u(\frac{d}{dt})\mathbf{u} + P_y(\frac{d}{dt})\mathbf{y}, \quad (9)$$

where matrices (Q, P_u, P_y) satisfy the following constraints: i) $Q(\xi)$ is nonsingular Hurwitz; ii) $[\,Q(\xi) \quad -P_u(\xi) \quad -P_y(\xi)\,] = [\,Y(\xi) \quad X(\xi)\,]\begin{bmatrix} \xi I_{\mathrm{n}} - F & -G & 0 \\ H & J & -I_{\mathrm{p}} \end{bmatrix}$, for some polynomial pair (X, Y).

It is easily seen that among all asymptotic observers, there are, in particular, Luenberger (full-order feedback) observers. In order to obtain them it is sufficient to assume in (ii) $Y(\xi) = I_{\mathrm{n}}$ and $X(\xi) = -L$, thus getting

$$\left(\frac{d}{dt}I_{\mathrm{n}} - F - LH\right)\hat{\mathbf{x}} = G\mathbf{u} - L\mathbf{y}. \quad (10)$$

This satisfies condition (i) if $F + LH$ is asymptotically stable, and its estimation error dynamics matrix, Δ, coincides with $\xi I_{\mathrm{n}} - F - LH$. Such an observer has a strictly proper rational transfer matrix $\hat{W}(\xi) := (\xi I_{\mathrm{n}} - F - LH)^{-1}[\,G \quad -L\,]$.

4 Proper asymptotic observers

Proposition 3.3 provides a useful parametrization of the asymptotic observers for a system described as in (6)÷(7). This parametrization can be fruitfully exploited to investigate further relevant issues as, in particular, the existence of (strictly) proper asymptotic observers, namely asymptotic observers with (strictly) proper transfer matrix, providing at each time t an estimate $\hat{\mathbf{w}}_2(t)$ which can be obtained by means of a non-anticipating algorithm. Clearly, this is possible if and only if there exists a matrix pair $(X(\xi), Y(\xi))$, with Y Hurwitz, such that $W(\xi) := [Y(\xi)D_2(\xi)]^{-1}[Y(\xi)N_1(\xi) + X(\xi)D_1(\xi)]$ is (strictly) proper rational.

As shown in the following proposition, autonomous behaviors described as in (6)÷(7) always admit proper asymptotic observers.

Proposition 4.1 *Let Σ be a dynamical system, whose behavior $\mathfrak{B}$ is described as in (6)÷(7), with D_2 Hurwitz and D_1 of full row rank d_1. If $\mathfrak{B}$ is autonomous or, equivalently, D_1 is a square matrix, there exists a (strictly) proper estimator of $\mathbf{w}_2$ from $\mathbf{w}_1$.*

PROOF Withoutloss of generality, assume that D_1 is row reduced, with row degrees $\mu_1, \mu_2, \ldots, \mu_{\mathrm{d}_1}$. Let Y

be any Hurwitz matrix such that YD_2 is row reduced with row degrees all greater or equal to $\max_i \mu_i - 1$ ($\max_i \mu_i$). By applying the matrix division algorithm, one can express YN_1 as $Y(\xi)N_1(\xi) = A(\xi)D_1(\xi)+R(\xi)$, with A and R polynomial matrices and R satisfying $\deg i$ th column of $R < \max_i \mu_i$, and hence $\deg i$ th row of $R \leq \deg i$ th row of YD_2, ($\deg i$ th row of $R < \deg i$ th row of YD_2), $\forall\ i$. Consequently, $(YD_2)^{-1}[YN_1 - AD_1] = (YD_2)^{-1}R$ is a proper (strictly proper) estimator of $\mathbf{w}_2$ from $\mathbf{w}_1$. ∎

When $\mathfrak{B}$ is an arbitrary (not necessarily autonomous) behavior, the problem is more involved. In order to solve it, we introduce the (unrestrictive) assumption that the behavior $\mathfrak{B}$ is described as in (1), with $[\,R_2 \quad -R_1\,] \in \mathbb{R}[\xi]^{\mathtt{r}\times(\mathtt{w}_2+\mathtt{w}_1)}$, $\mathtt{r} = \mathtt{w}_1+\mathtt{d}_1$, a row reduced matrix with row degrees $h_1, \ldots, h_{\mathtt{r}}$. The first step toward the solution is given by the following lemma where conditions for the existence of proper/strictly proper observers (not necessarily asymptotic) are given.

Lemma 4.2 *Consider a dynamical system Σ whose behavior $\mathfrak{B}$ is described as in (1), with $\mathbf{w}_2$ detectable from $\mathbf{w}_1$, and let $[\,R_{2hr} \quad -R_{1hr}\,]$ denote the leading row coefficient matrix of $[\,R_2 \quad -R_1\,]$.*

i) A necessary and sufficient condition for the existence of an observer of $\mathbf{w}_2$ from $\mathbf{w}_1$ with proper transfer matrix $W(\xi)$ is that R_{2hr} has full column rank $\mathtt{w}_2$.

ii) A necessary and sufficient condition for the existence of an observer of $\mathbf{w}_2$ from $\mathbf{w}_1$ with strictly proper transfer matrix $W(\xi)$ is that there exists $S \in \mathbb{R}^{\mathtt{w}_2\times(\mathtt{w}_2+\mathtt{d}_1)}$ s.t. $S[\,R_{2hr} \quad -R_{1hr}\,] = [\,I_{\mathtt{w}_2} \quad 0\,]$.

PROOF (i) Assume, first, that there exists an observer with proper transfer matrix $W(\xi)$ and let $T(\xi) = [t_{ij}(\xi)]$ be a polynomial matrix s.t. $[Q(\xi) \mid -P(\xi)] := T(\xi)[R_2(\xi) \mid -R_1(\xi)]$ satisfies $W(\xi) = Q^{-1}(\xi)P(\xi)$. It entails no loss of generality assuming that $[Q(\xi) \mid -P(\xi)]$ is row-reduced with row degrees $k_1, k_2, \ldots, k_{\mathtt{w}_2}$ and leading row coefficient matrix $[\,Q_{hr} \quad -P_{hr}\,]$. By the properness assumption on W, Q_{hr} is nonsingular, moreover, the row-properness of $[\,R_2 \quad -R_1\,]$ implies

$$k_i = \max_{h:t_{ih}(\xi)\neq 0}\{\deg t_{ih} + \deg(h\text{th row of } [\,R_2 \quad -R_1\,])\}.$$

Let S be the real matrix whose (i,j)th entry coincides with the leading coefficient of $t_{ij}(\xi)$ when $\deg t_{ij} + \deg(j\text{th row of } [\,R_2 \quad -R_1\,]) = k_i$ and is zero otherwise. Clearly, $Q_{hr} = SR_{2hr}$ thus proving that R_{2hr} has full column rank.

Conversely, suppose that R_{2hr} has full column rank and let $S = [s_{ij}]$ be any of its left inverses. Set $h := \max_i h_i$ and introduce the polynomial matrix $T(\xi) := [s_{ij}z^{h-h_j}]$. Then $[Q(\xi) \mid -P(\xi)] := T(\xi)[R_2(\xi) \mid -R_1(\xi)]$ is a row-reduced matrix (with all row degrees equal to h) and the first $\mathtt{w}_2 \times \mathtt{w}_2$ submatrix of its leading row coefficient matrix coincides with $I_{\mathtt{w}_2}$. Thus, $W(\xi) = Q^{-1}(\xi)P(\xi)$ is a proper rational matrix.

(ii) The proof follows the same lines of the previous one. ∎

Proposition 4.3 *Consider a dynamical system Σ whose behavior $\mathfrak{B}$ is described as in (1), with $\mathbf{w}_2$ detectable from $\mathbf{w}_1$. If there exists a proper observer of $\mathbf{w}_2$ from $\mathbf{w}_1$ then there exists a proper asymptotic observer of $\mathbf{w}_2$ from $\mathbf{w}_1$.*

PROOF Assume, as in the previous lemma, that $[\,R_2 \quad -R_1\,] \in \mathbb{R}[\xi]^{\mathtt{r}\times(\mathtt{w}_2+\mathtt{w}_1)}$ is row reduced with row degrees $h_1, \ldots, h_{\mathtt{r}}$ and leading row coefficient matrix $[\,R_{2hr} \quad -R_{1hr}\,]$. Since there exists a proper observer of $\mathbf{w}_2$ from $\mathbf{w}_1$, R_{2hr} has rank $\mathtt{w}_2$. Set, now,

$$[\,\hat{R}_2 \quad -\hat{R}_1\,] := \begin{bmatrix} \xi^{-h_1} & & \\ & \ddots & \\ & & \xi^{-h_{\mathtt{r}}} \end{bmatrix} [\,R_2 \quad -R_1\,].$$

Clearly, $[\,\hat{R}_2 \quad -\hat{R}_1\,]$ is a polynomial matrix in the negative powers of ξ, and the coefficient matrix of the constant term coincides with $[\,R_{2hr} \quad -R_{1hr}\,]$. Let $U \in \mathbb{R}[\xi^{-1}]^{\mathtt{r}\times\mathtt{r}}$ be a unimodular matrix (in $\mathbb{R}[\xi^{-1}]$) that reduces $\hat{R}_2$ to its Hermite form (still in $\mathbb{R}[\xi^{-1}]$): $U\hat{R}_2 = \begin{bmatrix} \hat{D}_2 \\ 0 \end{bmatrix}$. Clearly, the coefficient matrix of the constant terms in $\hat{D}_2$, $\hat{D}_{20}$, must be nonsingular. Also, by the detectability assumption, $\det \hat{D}_2 \in \mathbb{R}[\xi^{-1}]$ can be expressed as $\det \hat{D}_2 = d_2(\xi)/\xi^K$, with $d_2(\xi) \in \mathbb{R}[\xi]$ a Hurwitz polynomial of degree K.

Corresponding to $\hat{T} := [\,I_{\mathtt{w}_2} \quad 0\,]U$, the matrix pair $[\hat{Q} \mid -\hat{P}] := \hat{T}[\hat{R}_2 \mid -\hat{R}_1]$ provides a left matrix fraction description (over $\mathbb{R}[\xi^{-1}]$) of the proper transfer matrix $W = \hat{Q}^{-1}\hat{P}$, and therefore there exists some nonsingular diagonal matrix D, with all monomial entries, such that the $\mathtt{w}_2 \times (\mathtt{w}_2 + \mathtt{w}_1)$ polynomial matrix

$$[\,Q \quad -P\,] = D\hat{T} \begin{bmatrix} \xi^{-h_1} & & \\ & \ddots & \\ & & \xi^{-h_{\mathtt{r}}} \end{bmatrix} [\,R_2 \quad -R_1\,]$$

represents a proper asymptotic observer. ∎

REMARK As a consequence of the above proposition and lemma, once we assume that $[\,R_2 \quad -R_1\,]$ is row reduced, the existence of a proper asymptotic observer can be easily checked by simply verifying that R_{2rh} has full column rank.

Consider a state space model, described as in (4), with (F, H) a detectable pair. Also in this case we can look for (strictly) proper asymptotic state estimators, namely observers described by

$$Q(\frac{d}{dt})\hat{\mathbf{x}} = P_u(\frac{d}{dt})\mathbf{u} + P_y(\frac{d}{dt})\mathbf{y}, \qquad (11)$$

with i) $Q(\xi)$ nonsingular Hurwitz; ii) $[Q(\xi) \ \ -P_u(\xi) \ \ -P_y(\xi)] = [Y(\xi) \ \ X(\xi)] \begin{bmatrix} \xi I_n - F & -G & 0 \\ H & J & -I_p \end{bmatrix}$; iii) $Q^{-1}[P_u \ |P_y]$ (strictly) proper rational.

As previously remarked, the Luenberger observer exhibits a strictly proper rational transfer matrix. So, the existence of a proper state observer is not an issue. We are interested, instead, in determining what matrices Q, and hence Δ, possibly appear in the description of the estimation error dynamics. The first step toward the solution is given by the following lemma, which proves that condition (iii) is satisfied if and only if $Q^{-1}P_y$ is proper rational.

Lemma 4.4 *An asymptotic state observer described as in (11), with matrices Q, P_u and P_y satisfying conditions (i) and (ii), is (strictly) proper if and only if $Q^{-1}P_u$ is (strictly) proper.*

PROOF As matrix $\begin{bmatrix} \xi I_n - F & -G & 0 \\ H & J & -I_p \end{bmatrix}$ is full row rank, the pair (X, Y) is uniquely determined by the observer matrices Q, P_u and P_y as $X(\xi) = P_y(\xi)$ and $Y(\xi) = (Q(\xi) - P_y(\xi)H)(\xi I_n - F)^{-1}$. Consequently, $P_u(\xi) = Y(\xi)G - X(\xi)J = Q(\xi)(\xi I_n - F)^{-1}G - P_y(\xi)\left(H(\xi I_n - F)^{-1}G + J\right)$, and hence $Q^{-1}P_u = (\xi I_n - F)^{-1}G - Q^{-1}(\xi)P_y(\xi)[H(\xi I_n - F)^{-1}G + J]$. So, it is easily seen that $Q^{-1}P_u$ is (strictly) proper whenever $Q^{-1}P_y$ is.

The converse is obvious. ■

As an immediate consequence of the above lemma, the search for (strictly) proper asyntotic state observers is equivalent to the problem of determining proper asyntotic state observers for the autonomous system described as

$$\begin{bmatrix} \frac{d}{dt} I_n - F \\ H \end{bmatrix} \mathbf{x} = \begin{bmatrix} 0 \\ I_p \end{bmatrix} \mathbf{y}. \qquad (12)$$

This allows us to reduce ourselves to the special case of autonomous behaviors, previously analysed, thus getting the following result.

Proposition 4.5 *Let $D_\ell^{-1}N_\ell$ be a left coprime matrix fraction description over $\mathbb{R}[\xi]$ of $H(\xi I_n - F)^{-1}$, with D_ℓ row-reduced with row indices $h_1, \ldots, h_n$. For every polynomial pair $(\bar{X}(\xi), \bar{Y}(\xi))$ such that $Q(\xi) := \bar{Y}(\xi)(\xi I_n - F) + \bar{X}(\xi)H$ is row reduced with row degrees lower bounded by $\max_i h_i - 1$ ($\max_i h_i$), there exists a new pair $(X(\xi), Y(\xi))$ such that $Y(\xi)(\xi I_n - F) + X(\xi)H = Q(\xi)$ and $Q(\frac{d}{dt})\hat{\mathbf{x}} = X(\frac{d}{dt})\mathbf{y}$ is a proper (strictly proper) state observer for (12).*

PROOF It is sufficient to observe that the set of possible solutions of the matrix equation $Q(\xi) = Y(\xi)(\xi I_n - F) + X(\xi)H$ is expressed as $[Y(\xi) \mid X(\xi)] = [\bar{Y}(\xi) \mid \bar{X}(\xi)] + T(\xi)[-N_\ell(\xi) \mid D_\ell(\xi)]$, as T varies in $\mathbb{R}[\xi]^{n \times p}$, and to apply the same division algorithm of Proposition 4.1 to $\bar{X}(\xi)$ and D_ℓ, thus getting $\bar{X}(\xi) = -T(\xi)D_\ell(\xi) + X(\xi)$, with all row degrees of X smaller than $\max_i h_i$. ■

Notice that as the row indices $h_1, h_2, \ldots, h_n$ are the well-known *observability indices* [1, 2, 3], the previous proposition states that it is always possible to obtain a state observer (11) with Q row reduced with row degrees lower bounded by the maximum of the observability indices. So, these indices somehow provide a lower bound on the complexity of the asymptotic state observers. This situation reminds of an analogous one for the classical output feedback compensator, where the reachability indices are involved [2, 3].

5 Latent variables models

Up to now we have devoted our attention to ordinary dynamical systems, where all quantities involved in the system description can be thought of as *manifest variables*, by this meaning that they generally correspond to variables of interest in the system. In several situations, however, when writing down the laws which govern the system evolution one has to include also auxiliary variables, that play an "additional" role in the system decription, as the essence of the system behavior is already captured by the manifest variables. As a consequence, we assume that latent variables are not measurable.

Consider a dynamical system with latent variables [4] Σ_ℓ whose behavior $\mathfrak{B}_f$ is described by the differential equation

$$R(\frac{d}{dt})\mathbf{w} = L(\frac{d}{dt})\ell, \qquad (13)$$

with $\dim\ell = 1$, and assume that the vector $\mathbf{w}$ of external variables splits into two subvectors: $\mathbf{w} = \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \end{bmatrix}$, $\dim\mathbf{w}_i = \mathtt{w}_i$, (and R is conformably partitioned as $R(\xi) = [R_2(\xi) | -R_1(\xi)]$), with $\mathbf{w}_1$ measurable, and $\mathbf{w}_2$ not accessible. Also in this framework we can introduce the notions of observability and detectability.

Definition For a dynamical system with latent variables Σ_ℓ, with behavior $\mathfrak{B}_f$ described by (13), $\mathbf{w}_2$ is

- *observable* from $\mathbf{w}_1$, if $(\mathbf{w}_1, \mathbf{w}_2, \ell), (\mathbf{w}_1, \bar{\mathbf{w}}_2, \bar{\ell}) \in \mathfrak{B}_f$ implies $\mathbf{w}_2 = \bar{\mathbf{w}}_2$;

- *detectable* from $\mathbf{w}_1$, if $(\mathbf{w}_1, \mathbf{w}_2, \ell), (\mathbf{w}_1, \bar{\mathbf{w}}_2, \bar{\ell}) \in \mathfrak{B}_f$ implies $\mathbf{w}_2(t) - \bar{\mathbf{w}}_2(t) \xrightarrow[t\to\infty]{} 0$.

As observability/detectability in the presence of ℓ requires, in particular, observability/detectability corresponding to $\ell = 0$, $\mathfrak{B}_f$ can always be described as follows

$$\begin{array}{rcl} D_2(\frac{d}{dt})\mathbf{w}_2 & = & N_1(\frac{d}{dt})\mathbf{w}_1 + L_A(\frac{d}{dt})\ell, \\ 0 & = & D_1(\frac{d}{dt})\mathbf{w}_1 + L_B(\frac{d}{dt})\ell, \end{array}$$

with D_2 Hurwitz and $[D_1|L_B]$ full row rank.

Proposition 5.1 *Let $\Sigma_\ell = (\mathbb{R}, \mathbb{R}^{\mathtt{w}_1+\mathtt{w}_2+1}, \mathfrak{B}_f)$ be a dynamical system with latent variables, whose behavior $\mathfrak{B}_f$ is described as in (5). Then i) $\mathbf{w}_2$ is observable from $\mathbf{w}_1$ if and only if D_2 is unimodular and*

$$\ker(L_B(\frac{d}{dt})) \subseteq \ker(L_A(\frac{d}{dt})); \qquad (14)$$

ii) $\mathbf{w}_2$ is detectable from $\mathbf{w}_1$ if and only if D_2 is Hurwitz and (14) holds true.

Proof (i) Assume, first, that $\mathbf{w}_2$ is observable from $\mathbf{w}_1$. Clearly, D_2 is unimodular, otherwise for every nonzero $\mathbf{v} \in \ker(D_2(\frac{d}{dt}))$ we would have $(\mathbf{w}_1, \mathbf{w}_2, \ell) \in \mathfrak{B}_f \Rightarrow (\mathbf{w}_1, \mathbf{w}_2 + \mathbf{v}, \ell) \in \mathfrak{B}_f$, thus contradicting the observability assumption. On the other hand, if $\ker(L_B(\frac{d}{dt}))$ would not be included in $\ker(L_A(\frac{d}{dt}))$, there would be some sequence $\tilde{\mathbf{v}}$ such that $L_B(\frac{d}{dt})\tilde{\mathbf{v}} = 0$ but $L_A(\frac{d}{dt})\tilde{\mathbf{v}} \neq 0$, and hence we would have $(\mathbf{w}_1, \mathbf{w}_2, \ell) \in \mathfrak{B}_f$ and $(\mathbf{w}_1, \tilde{\mathbf{w}}_2, \ell + \tilde{\mathbf{v}}) \in \mathfrak{B}_f$, with $\tilde{\mathbf{w}}_2 \neq \mathbf{w}_2$. The proof of the converse follows the same reasoning.

(ii) Suppose, now, that $\mathbf{w}_2$ is detectable from $\mathbf{w}_1$. If D_2 would not be Hurwitz, ker $(D_2(\frac{d}{dt}))$ would include some sequence $\mathbf{v}$ which does not asymptotically estinguish, and we would have $(\mathbf{w}_1, \mathbf{w}_2, \ell) \in \mathfrak{B}_f \Rightarrow (\mathbf{w}_1, \mathbf{w}_2 + \mathbf{v}, \ell) \in \mathfrak{B}_f$, thus contradicting the detectability assumption. As in (i), if (14) would not be verified, there would be sequences $(\mathbf{w}_1, \mathbf{w}_2, \ell) \in \mathfrak{B}_f$ and $(\mathbf{w}_1, \tilde{\mathbf{w}}_2, \ell + \tilde{\mathbf{v}}) \in \mathfrak{B}_f$, with $\tilde{\mathbf{w}}_2 - \mathbf{w}_2$ that does not go to zero asymptotically, thus contradicting the detectability assumption. The converse part is proved by reversing the above arguments. ∎

In the sequel, condition (14) will be replaced by the equivalent statement [4] that $L_A = ML_B$ for some suitable polynomial matrix M. Of course, as for ordinary dynamical systems, one may look for conditions guaranteeing the existence of asymptotic observers of $\mathbf{w}_2$ from $\mathbf{w}_1$. As before, we assume that the observer is described by some differential equation $Q(\frac{d}{dt})\hat{\mathbf{w}}_2 = P(\frac{d}{dt})\mathbf{w}_1$,with P and Q polynomial matrices of suitable dimensions, and assume that the set $\mathfrak{B}_e$ of admissible estimation errors is a stable autonomous behavior, i.e. $\mathfrak{B}_e = \ker\Delta(\frac{d}{dt})$, for some Hurwitz matrix Δ.

In order to analyse this problem, we introduce some (unrestrictive) assumptions on the structure of $[D_1 \ \ L_B]$, namely we assume $[D_1 \ \ L_B] = \begin{bmatrix} D_{11} & L_1 \\ D_{21} & 0 \end{bmatrix}$, with L_1 and D_{21} full row rank matrices. Consequently, $\mathfrak{B}_f$ will be described by the following set of differential equations:

$$\begin{array}{rcl} D_2(\frac{d}{dt})\mathbf{w}_2 & = & N_1(\frac{d}{dt})\mathbf{w}_1 + L_A(\frac{d}{dt})\ell, \\ 0 & = & D_{11}(\frac{d}{dt})\mathbf{w}_1 + L_1(\frac{d}{dt})\ell, \\ 0 & = & D_{21}(\frac{d}{dt})\mathbf{w}_1, \end{array}$$

with D_2 Hurwitz, L_1 and D_{21} full row rank.

Proposition 5.2 *Given a dynamical system with latent variables Σ_ℓ, whose behavior $\mathfrak{B}_f$ is described as in (5), a necessary and sufficient condition for the existence of an asymptotic observer of $\mathbf{w}_2$ from $\mathbf{w}_1$ is that $\mathbf{w}_2$ is detectable from $\mathbf{w}_1$.*

Proof If $\mathbf{w}_2$ is detectable from $\mathbf{w}_1$, then D_2 is nonsingular Hurwitz and $L_A = ML_B = [M_1 \ \ M_2]\begin{bmatrix} L_1 \\ 0 \end{bmatrix} = M_1 L_1$ for some suitable polynomial matrix M_1. We aim to show that $D_2(\frac{d}{dt})\hat{\mathbf{w}}_2 = (N_1 - M_1 D_{11})(\frac{d}{dt})\mathbf{w}_1$ is an asymptotic observer of $\mathbf{w}_2$ from $\mathbf{w}_1$. By applying (5) we get $D_2(\frac{d}{dt})(\mathbf{w}_2 - \hat{\mathbf{w}}_2) = N_1(\frac{d}{dt})\mathbf{w}_1 + L_A(\frac{d}{dt})\ell - (N_1 - M_1 D_{11})(\frac{d}{dt})\mathbf{w}_1 = (M_1 L_1)(\frac{d}{dt})\ell + (M_1 D_{11})(\frac{d}{dt})\mathbf{w}_1 = 0$, thus proving that the estimation error belongs to some stable autonomous behavior. The proof of the converse follows the same lines of that in Proposition 3.1. ∎

As in section 4, it can be easily proved that matrix Q can always be assumed nonsingular square. Under this assumption, by applying the same kind of reasonings adopted to prove Proposition 3.3, we get the following parametrization of asymptotic observers.

Proposition 5.3 *Consider a dynamical system with latent variables Σ_ℓ, with behavior $\mathfrak{B}_f$ described as in (5), D_2 Hurwitz and $L_A = M_1 L_1$ for some polynomial matrix M_1. If P and Q are polynomial matrices, with Q full row rank, then $Q(\frac{d}{dt})\hat{\mathbf{w}}_2 = P(\frac{d}{dt})\mathbf{w}_1$, is an asymptotic observer of $\mathbf{w}_2$ from $\mathbf{w}_1$ if and only if there exists $Y \in \mathbb{R}[\xi]^{\mathtt{w}_2 \times \mathtt{w}_2}$ Hurwitz and $X \in \mathbb{R}[\xi]^{\mathtt{w}_2 \times \mathtt{d}_1}$ s.t.*

$$\begin{array}{rcl} Q(\xi) & = & Y(\xi) D_2(\xi) \\ P(\xi) & = & Y(\xi)[N_1(\xi) - M_1(\xi) D_{11}(\xi)] + X(\xi) D_{21}(\xi) \end{array}$$

References

[1] T. Kailath. *Linear Systems.* Prentice Hall, 1980.

[2] V. Kučera. *Discrete linear control: the polynomial equation approach.* J.Wiley & Sons, NY, 1979.

[3] V. Kučera. *Analysis and design of discrete linear control systems.* Academia, Praga, 1991.

[4] J.C. Willems. Paradigms and puzzles in the theory of dynamical systems. *IEEE Trans. Aut.Contr.*, vol.36, 1991, pp. 259–294.

[5] J.C. Willems. On interconnections, control, and feedback. *IEEE Trans. Aut. Contr.*, vol. AC-42, 1997, pp. 326–339.