Systems & Control Letters 17 (1991) 49–58
North-Holland

Observer design for autonomous discrete-time nonlinear systems

Wonchang Lee and Kwanghee Nam

Department of Electrical Engineering, POSTECH, Pohang, P.O. Box 125, Kyungbuk 790-600, South Korea

Received 31 March 1990
Revised 17 September 1990 and 25 February 1991

Abstract: We obtain a necessary and sufficient condition for a discrete-time nonlinear system to be locally state equivalent to the nonlinear observer form. The result looks similar to the continuous counterpart except for the fact that Ad-operation is utilized instead of ad-operation.

Keywords: Discrete-time nonlinear systems; nonlinear observer; state equivalent; coordinate transformation; observability.

1. Introduction

A dual concept to feedback linearization is the design of a nonlinear observer that has linearizable error dynamics. The nonlinear observer form is defined to be a canonical form for which an observer can be constructed with linearizable error dynamics. If a system can be transformed into the nonlinear observer form, then one can construct an observer for it through the use of coordinate transformation. Krener and Isidori [1] obtained a necessary and sufficient condition for a single-output system to be state equivalent to the nonlinear observer form, and it was extended to the multi-output systems by Krener and Respondek [2] and Xia and Gao [3]. However, all of the previous works were done on continuous-time models.

In this paper, we derive a necessary and sufficient condition for a discrete-time nonlinear system to be state equivalent to the nonlinear observer form in the single-output and multi-output cases, respectively. This work is developed as a parallel result to the continuous-time case [3] using the techniques in [4].

2. Definition and preliminaries

Let M be a smooth n-dimensional manifold. By smooth, we mean infinite differentiability. Let T_xM denote the tangent space at $x \in M$. Given $X(t, x) \in T_xM$ for each $t \in \mathbb{R}$, we denote by $\Phi_t^X(p_0)$ the solution $\Phi(t) \in M$ of

$$\frac{\mathrm{d}\Phi}{\mathrm{d}t} = X(t, \Phi), \quad \Phi(t_0) = p_0.$$

Definition. Let U_1, U_2 be open subsets of a smooth manifold M and $\sigma : U_1 \to U_2$ be a local homeomorphism of class at least C^1. For a vector field X over an open set V of M, we define $\mathrm{Ad}_\sigma X$ to be a vector field on $\sigma(U_1 \cap V)$ such that

$$\mathrm{Ad}_\sigma X(p) = \mathrm{D}\sigma|_{\sigma^{-1}(p)} X(\sigma^{-1}(p)),$$

where $\mathrm{D}\sigma$ means the Jacobian of σ.

0167-6911/91/$03.50 © 1991 – Elsevier Science Publishers B.V. (North-Holland)

Theorem [5]. *Let $\phi: M \to N$ be C^∞. Let X and X_1 be smooth vector fields on M, and let Y and Y_1 be smooth vector fields on N. If $\mathrm{D}\phi X = Y \circ \phi$ and $\mathrm{D}\phi X_1 = Y_1 \circ \phi$, then $\mathrm{D}\phi[X, X_1] = [Y, Y_1] \circ \phi$.*

We denote by $\mathbb{R}$ the real line. We denote by $C^\infty(M, N)$ the set of infinitely many differentiable functions from M to N. Given $F: M \to M$, $h: M \to \mathbb{R}^m$, $h \circ F$ denotes the composite function $h(F(x))$. We also denote by F^n the n-times composite function $F \circ \cdots \circ F$ (n times). We denote by ε_j the unit vector of $\mathbb{R}^n$ whose j-th component is unity. We denote by δ_{ij} the Kronecker delta whose value is 1 if $i = j$ and 0 if $i \neq j$.

The following canonical structure is called nonlinear observer form:

$$z(k+1) = Az(k) + \alpha(y(k)), \tag{1}$$

$$y(k) = Cz(k), \tag{2}$$

where $z(k) \in \mathbb{R}^n$,

$$A = \begin{bmatrix} \begin{matrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{matrix} & & \\ & \ddots & \\ & & \begin{matrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{matrix} \end{bmatrix},$$

$$C = \begin{bmatrix} 0 \cdots 0 \; 1 & & & \\ & 0 \cdots 0 \; 1 & & \\ & & \ddots & \\ & & & 0 \cdots 0 \; 1 \end{bmatrix},$$

$\alpha^{\mathrm{T}} \equiv [\alpha_1, \dots, \alpha_n]$ is a vector of scalar functions of $y(k)$. Once a system is represented by the nonlinear observer form, then we can construct an asymptotic observer in such a way that

$$\hat{z}(k+1) = A\hat{z}(k) + \alpha(y(k)) + L(C\hat{z}(k) - y(k)). \tag{3}$$

Then the state error $e = \hat{z} - z$ satisfies

$$e(k+1) = (A + LC)e(k) \tag{4}$$

and $e(k)$ vanishes as $k \to \infty$ if the eigenvalues of $A + LC$ lie within the unit circle.

Obviously, there is a certain class of systems which are equivalent to (1), (2) under a coordinate change. We consider a discrete-time nonlinear system defined on an n-dimensional smooth manifold M:

$$x(k+1) = F(x(k)), \tag{5}$$

$$y(k) = h(x(k)). \tag{6}$$

We say that the system (5), (6) is (locally) *state equivalent* to the system (1), (2) and vice versa if there exists a (local) diffeomorphism which transforms the system (5), (6) into the system (1), (2).

If the system (5), (6) can be transformed into the nonlinear observer form (1), (2) and the coordinate transformation map is available, then we can reconstruct the state of the system. Specifically, suppose that a (local) diffeomorphism T transforms the system (5), (6) into the nonlinear observer form (1), (2). Then, constructing an observer (3) for (1), (2), the observation error $e(k) = \hat{z}(k) - z(k)$ vanishes. By the continuity of T^{-1}, $\hat{x}(k) \equiv T^{-1}(\hat{z}(k))$ converges to $x(k)$ as $k \to \infty$, so that one can reconstruct the state $x(k)$ through the inverse coordinate transformation map T^{-1}. Therefore, the crucial point in the construction of a nonlinear observer is the transformability of the system (5), (6) into the nonlinear observer form (1), (2).

3. Single-output case

Let $(\mathcal{U}, x)$ be a local chart of M. We assume that F is a local diffeomorphism from a neighborhood $U_{x_e} \subset \mathcal{U}$ of a fixed point x_e (i.e. $x_e = F(x_e)$) onto its image. Let V_0 be a neighborhood of $0 \in \mathbb{R}^n$. We consider a single-output discrete-time nonlinear system of the form

$$x(k+1) = F(x(k)), \tag{7}$$

$$y(k) = h(x(k)), \tag{8}$$

where $h \in C^\infty(U_{x_e}, \mathbb{R})$ and $h(x_e) = 0$. In this section we derive a necessary and sufficient condition for the existence of a local diffeomorphism $T \in C^\infty(U_{x_e}, V_0)$ such that in the new coordinate $z = T(x)$ the system (7), (8) is described by

$$z(k+1) = Az(k) + \alpha(y(k)) \equiv G(z(k)), \tag{9}$$

$$y(k) = cz(k), \tag{10}$$

where

$$A = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}, \qquad c = [0 \quad \cdots \quad 0 \quad 1].$$

Suppose that the system (7), (8) can be transformed into (9), (10) by a coordinate change $z = T(x)$. We also define $\chi \in C^\infty(V_0, U_{x_e})$ to be the inverse map of T, i.e., $\chi(z) \equiv T^{-1}(z)$. Since $x(k+1) = \chi(z(k+1))$, we obtain

$$F(x(k)) = \chi(z(k+1)). \tag{11}$$

Note from the structure of G that $\partial G/\partial z_j = \varepsilon_{j+1}$, where z_j denotes the j-th component of the state z. Differentiating both sides of (11) with respect to z_j yields

$$\left.\frac{\partial F}{\partial x}\right|_{x(k)} \frac{\partial \chi}{\partial z_j}(z(k)) = \left.\frac{\partial \chi}{\partial z}\right|_{z(k+1)} \frac{\partial G}{\partial z_j}(z(k)) = \frac{\partial \chi}{\partial z_{j+1}}(z(k+1)), \quad 1 \leqslant j \leqslant n-1, \tag{12}$$

and

$$\left.\frac{\partial F}{\partial x}\right|_{x(k)} \frac{\partial \chi}{\partial z_n}(z(k)) = \left.\frac{\partial \chi}{\partial z}\right|_{z(k+1)} \frac{\partial \alpha}{\partial z_n}(y(k)). \tag{13}$$

Successive operation yields

$$\frac{\partial \chi}{\partial z_{j+1}}(z(k+j)) = \left.\frac{\partial F}{\partial x}\right|_{x(k+j-1)} \cdots \left.\frac{\partial F}{\partial x}\right|_{x(k)} \frac{\partial \chi}{\partial z_1}(z(k)) = \left.\frac{\partial F^j}{\partial x}\right|_{x(k)} \frac{\partial \chi}{\partial z_1}(z(k)) \tag{14}$$

for $1 \leqslant j \leqslant n-1$ and

$$\left.\frac{\partial \chi}{\partial z}\right|_{z(k+n)} \frac{\partial \alpha}{\partial z_n}(y(k+n-1)) = \left.\frac{\partial F^n}{\partial x}\right|_{x(k)} \frac{\partial \chi}{\partial z_1}(z(k)). \tag{15}$$

On the other hand, since $h(x(k)) = h \circ \chi(z(k)) = cz(k)$,

$$\left\langle \frac{\partial h}{\partial x}(\chi(z(k))), \frac{\partial \chi}{\partial z_j}(z(k)) \right\rangle = \delta_{jn}, \quad 1 \leqslant j \leqslant n. \tag{16}$$

Using (14), we can rewrite (16) as

$$\left\langle \frac{\partial h}{\partial x}\bigg|_{x(k+j-1)} \frac{\partial F^{j-1}}{\partial x}(x(k)), \frac{\partial \chi}{\partial z_1}(z(k)) \right\rangle = \left\langle \frac{\partial (h \circ F^{j-1})}{\partial x}(x(k)), \frac{\partial \chi}{\partial z_1}(z(k)) \right\rangle = \delta_{jn}. \tag{17}$$

Then, defining the observability matrix O by

$$O(x) = \begin{bmatrix} \dfrac{\partial h}{\partial x} \\ \dfrac{\partial (h \circ F)}{\partial x} \\ \vdots \\ \dfrac{\partial (h \circ F^{n-1})}{\partial x} \end{bmatrix}(x),$$

we obtain

$$O(x(k))\frac{\partial \chi}{\partial z}(z(k)) = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ \vdots & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \times \\ 0 & \cdot^{\cdot^{\cdot}} & \times & \vdots \\ 1 & \times & \cdots & \times \end{bmatrix}, \tag{18}$$

where $\times$ denotes an arbitrary function. The nonsingularity of the right hand side of (18) and $\partial\chi/\partial z$ implies that $O(x_e)$ is full-rank.

The right hand side of (14) becomes

$$\frac{\partial F^j}{\partial x}\bigg|_{x(k)} \frac{\partial \chi}{\partial z_1}(\chi^{-1}(F^{-j}(x(k+j)))) = \frac{\partial F^j}{\partial x}\bigg|_{F^{-j}(x(k+j))} \frac{\partial \chi}{\partial z_1}\bigg|_{z(k)} (F^{-1}(x(k+j)))$$

for $1 \leqslant j \leqslant n-1$. Thus, we can rewrite (14) as

$$\frac{\partial \chi}{\partial z_{j+1}}(z(k+j)) = \mathrm{Ad}_F^j\left(\frac{\partial \chi}{\partial z_1}\right)(z(k+j)), \quad 0 \leqslant j \leqslant n-1. \tag{19}$$

Since $\partial\chi/\partial z_1, \partial\chi/\partial z_2, \ldots, \partial\chi/\partial z_n$, i.e., $\partial\chi/\partial z_1$, $\mathrm{Ad}_F(\partial\chi/\partial z_1), \ldots, \mathrm{Ad}_F^{n-1}(\partial\chi/\partial z_1)$ are push-forwarded vectors of $\partial/\partial z_1, \partial/\partial z_2, \ldots, \partial/\partial z_n$, it must follow from the theorem that

$$\left[\mathrm{Ad}_F^i\left(\frac{\partial \chi}{\partial z_1}\right), \mathrm{Ad}_F^j\left(\frac{\partial \chi}{\partial z_1}\right)\right] = 0, \quad 0 \leqslant i, \ j \leqslant n-1. \tag{20}$$

Commutativity of vector fields (20) is a necessary condition that the system (7), (8) is state equivalent to the nonlinear observer form (9), (10). This implies that the integral curves of the vector fields $\partial\chi/\partial z_1$, $\mathrm{Ad}_F(\partial\chi/\partial z_1), \ldots, \mathrm{Ad}_F^{n-1}(\partial\chi/\partial z_1)$ can be used to form the coordinate lines of coordinate system. Necessary condition (20) seems to be uncheckable, since it utilizes the unknown map χ. But, from (18), we can check the following identity:

$$\frac{\partial \chi}{\partial z_1}(z(k)) = O^{-1}(x(k))\varepsilon_n.$$

This means that we can directly check the transformability of the system (7), (8) into (9), (10) by the commutativity of the vector fields generated by F and $O^{-1}(x)\varepsilon_n$.

Proposition 1. *There exists a local diffeomorphism which transforms* (7), (8) *into the nonlinear observer form* (9), (10) *if and only if*

(i) $O(x_e)$ *is full-rank*,

(ii) $[\mathrm{Ad}_F^i g, \mathrm{Ad}_F^j g] = 0$, $0 \leq i,\ j \leq n-1$, *where* $g(x) \equiv O^{-1}(x)\varepsilon_n$.

Proof. The necessary part was proved in the above. For sufficiency, we define a map $\chi \in C^\infty(V_0, U_{x_e})$ by

$$\chi(z_1, \ldots, z_n) = \Phi_{z_1}^{g} \circ \Phi_{z_2}^{\mathrm{Ad}_F g} \cdots \circ \Phi_{z_n}^{\mathrm{Ad}_F^{n-1} g}(x_e). \tag{21}$$

From condition (ii) we obtain

$$\frac{\partial \chi}{\partial z_j}(z(k)) = \mathrm{Ad}_F^{j-1} g(\chi(z(k))), \quad 1 \leq j \leq n,$$

for a $z(k)$ in V_0. Thus,

$$\frac{\partial \chi}{\partial z}(z(k)) = [g \quad \mathrm{Ad}_F g \quad \cdots \quad \mathrm{Ad}_F^{n-1} g](\chi(z(k))). \tag{22}$$

From (22) and $g(x(k)) \equiv O^{-1}(x(k))\varepsilon_n$ we obtain the relation (18). Further, since $O(x_e)$ is full-rank, $(\partial\chi/\partial z)(0)$ is nonsingular. Hence, χ is a local diffeomorphism from 0 to x_e. Choosing $T = \chi^{-1}$ as a coordinate transformation map, we obtain $z(k+1) = \tilde{F}(z(k))$, $y(k) = h \circ \chi(z(k))$, where $\tilde{F}(z(k)) = \chi^{-1} \circ F \circ \chi(z(k))$. Then, from (22), we obtain

$$\begin{aligned}
\frac{\partial \tilde{F}}{\partial z}(z(k)) &= [g \ \cdots \ \mathrm{Ad}_F^{n-1} g]^{-1}(x(k+1)) \left.\frac{\partial F}{\partial x}\right|_{x(k)} [g \ \cdots \ \mathrm{Ad}_F^{n-1} g](x(k)) \\
&= [g \ \cdots \ \mathrm{Ad}_F^{n-1} g]^{-1}(x(k+1)) \left.\frac{\partial F}{\partial x}\right|_{F^{-1}(x(k+1))} [g \ \cdots \ \mathrm{Ad}_F^{n-1} g](F^{-1}(x(k+1))) \\
&= [g \ \cdots \ \mathrm{Ad}_F^{n-1} g]^{-1}(x(k+1)) [\mathrm{Ad}_F g \ \cdots \ \mathrm{Ad}_F^{n} g](x(k+1)).
\end{aligned}$$

This implies that

$$\frac{\partial \tilde{F}}{\partial z_j}(z(k)) = \varepsilon_{j+1}, \quad 1 \leq j \leq n-1.$$

Thus, we deduce that $\tilde{F}_j$ for $2 \leq j \leq n$ is a linear function of z_{j-1}, while all components of $\tilde{F}$ are nonlinear functions of z_n. Hence, we obtain

$$\tilde{F}(z(k)) = Az(k) + \alpha(z_n(k)).$$

Further, we obtain

$$\frac{\partial(h \circ \chi)}{\partial z_j}(z(k)) = \left\langle \frac{\partial h}{\partial x}(x(k)), \mathrm{Ad}_F^{j-1} g(x(k)) \right\rangle = \delta_{jn}, \quad 1 \leq j \leq n.$$

Therefore, we obtain $y(k) = cz(k)$. □

Remark. Obviously, if $O(x_e)$ is full-rank, the system (7), (8) is observable in an arbitrary small neighborhood of x_e. But, with condition (ii), the observability condition is much more strengthened. That is, the region of observability can be extended to the place where the diffeomorphic condition of χ collapses. Further, if the maps F and χ are global diffeomorphisms, then the system (7), (8) is globally state equivalent to the system (9), (10).

4. Extension to multi-output systems

In this section, we consider the problem of transforming the multi-output system (5), (6) into the nonlinear observer from (1), (2). Suppose that m-tuple of integers $k_1, \ldots, k_m$ with $k_1 \geqslant k_2 \geqslant \cdots \geqslant k_m$ and $\Sigma_{i=1}^{m} k_i = n$ are observability indices. We assume that the components z_{ij} of z are ordered lexographically. Suppose that the system (5), (6) can be transformed into (1), (2) by a coordinate change $z = T(x)$. Let $\chi \in C^\infty(V_0, U_{x_e})$ to be the inverse map of T, i.e., $\chi(z) \equiv T^{-1}(z)$. Then, similarly to the single-output case, we obtain for $1 \leqslant i \leqslant m$,

$$\frac{\partial \chi}{\partial z_{ij+1}}(z(k+j)) = \left.\frac{\partial F^j}{\partial x}\right|_{x(k)} \frac{\partial \chi}{\partial z_{i1}}(z(k)), \quad 0 \leqslant j \leqslant k_i - 1, \tag{23}$$

and

$$\left.\frac{\partial \chi}{\partial z}\right|_{z(k+k_i)} \frac{\partial \alpha}{\partial z_{ik_i}}(y(k+k_i-1)) = \left.\frac{\partial F^{k_i}}{\partial x}\right|_{x(k)} \frac{\partial \chi}{\partial z_{i1}}(z(k)). \tag{24}$$

On the other hand, since $h(x(k)) = h(\chi(z(k))) = Cz(k)$, it should follow that for $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant k_i$, $1 \leqslant l \leqslant m$,

$$\left\langle \frac{\partial h_l}{\partial x}(\chi(z(k))), \frac{\partial \chi}{\partial z_{ij}}(z(k)) \right\rangle = \delta_{il}\delta_{jk_i}. \tag{25}$$

Using (23), we obtain

$$\left\langle \left.\frac{\partial h_l}{\partial x}\right|_{x(k+j-1)} \frac{\partial F^{j-1}}{\partial x}(x(k)), \frac{\partial \chi}{\partial z_{i1}}(z(k)) \right\rangle = \left\langle \frac{\partial \left(h_l \circ F^{j-1}\right)}{\partial x}(x(k)), \frac{\partial \chi}{\partial z_{i1}}(z(k)) \right\rangle = \delta_{il}\delta_{jk_i}, \tag{26}$$

for $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant k_i$, $1 \leqslant l \leqslant m$. We define

$$O(x) = \begin{bmatrix} O_1(x) \\ \vdots \\ O_m(x) \end{bmatrix}, \quad O_i(x) = \begin{bmatrix} \dfrac{\partial h_i}{\partial x} \\ \dfrac{\partial (h_i \circ F)}{\partial x} \\ \vdots \\ \dfrac{\partial \left(h_i \circ F^{k_i - 1}\right)}{\partial x} \end{bmatrix}(x), \quad 1 \leqslant i \leqslant m.$$

Then, we obtain

$$O(x(k)) \frac{\partial \chi}{\partial z}(z(k)) = \begin{bmatrix} \begin{matrix} 0 & \cdots & 0 & 1 \\ \vdots & \iddots & \iddots & \times \\ 0 & \iddots & \times & \vdots \\ 1 & \times & \cdots & \times \end{matrix} & & \\ & \ddots & \\ & & \begin{matrix} 0 & \cdots & 0 & 1 \\ \vdots & \iddots & \iddots & \times \\ 0 & \iddots & \times & \vdots \\ 1 & \times & \cdots & \times \end{matrix} \end{bmatrix}. \tag{27}$$

The nonsingularity of the right hand side of (27) and $\partial\chi/\partial z$ implies that $O(x_e)$ is full-rank. Further, we obtain

$$\frac{\partial\chi}{\partial z_{ij+1}}(z(k+j)) = \mathrm{Ad}_F^j\left(\frac{\partial\chi}{\partial z_{i1}}\right)(z(k+j)), \quad 1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant k_i - 1,$$

and

$$\left[\mathrm{Ad}_F^p\left(\frac{\partial\chi}{\partial z_{i1}}\right), \mathrm{Ad}_F^q\left(\frac{\partial\chi}{\partial z_{j1}}\right)\right] = 0, \quad 1 \leqslant i,\ j \leqslant m,\ 0 \leqslant p \leqslant k_i - 1,\ 0 \leqslant q \leqslant k_j - 1.$$

Let $g_i(x(k)) = (\partial\chi/\partial z_{i1})(z(k))$. Then, from (26), we obtain

$$\left\langle \frac{\partial\left(h_l \circ F^{j-1}\right)}{\partial x}(x(k)),\ g_i(x(k)) \right\rangle = \delta_{il}\delta_{jk_i}, \tag{28}$$

for $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant k_i$, $1 \leqslant l \leqslant m$. Differently from single-output systems, we need the following lemma for multi-output systems. Lemma 1 is a discrete-time system version of Theorem 2.3 in [3]. We put the proof of Lemma 1 in the Appendix.

Lemma 1. *The discrete-time nonlinear system* (5), (6) *is transformed into the nonlinear observer form* (1), (2) *only if there exist m-tuple of integers* $k_1, \ldots, k_m$, $k_1 \geqslant k_2 \geqslant \cdots \geqslant k_m$, *and* $\sum_{i=1}^m k_i = n$, *such that*

(i) *the set* $\mathcal{Q} = \{(\partial(h_i \circ F^j)/\partial x)(x)\colon 1 \leqslant i \leqslant m,\ 0 \leqslant j \leqslant k_i - 1\}$ *spans n-dimensional space in* U_{x_e},

(ii) *if we define* $\mathcal{Q}_i = \{(\partial h_j \circ F^k)/\partial x)(x)\colon 1 \leqslant j \leqslant m, 0 \leqslant k \leqslant k_i - 1\} - \{(\partial(h_i \circ F^{k_i-1})/\partial x)(x)\}$, $1 \leqslant i \leqslant m$, *then* span $\mathcal{Q}_i = \mathrm{span}\{\mathcal{Q} \cap \mathcal{Q}_i\}$, $1 \leqslant i \leqslant m$, *where* $A - B = A \cap \bar{B}$ *for any two sets* A, B.

Remark. Condition (i) of Lemma 1 implies that the system (5), (6) is observable in U_{x_e}. Condition (ii) implies that elements in $\mathcal{Q}_i - \{\mathcal{Q} \cap \mathcal{Q}_i\}$ belong to $\mathrm{span}\{\mathcal{Q} \cap \mathcal{Q}_i\}$, since $\mathrm{span}\,\mathcal{Q}_i \supset \mathrm{span}\{\mathcal{Q} \cap \mathcal{Q}_i\}$ always holds. If condition (ii) does not hold, there is a chance that an element in $\mathcal{Q}_i - \{\mathcal{Q} \cap \mathcal{Q}_i\}$ may belong to $\mathrm{span}\{\mathcal{Q}_i \cup \{\partial(h_i \circ F^{k_i-1})/\partial x\}\} = \mathrm{span}\{\partial(h_j \circ F^k)/\partial x\colon 1 \leqslant j \leqslant m,\ 0 \leqslant k \leqslant k_i - 1\}$. Then, condition (28) can not be satisfied. In other words, the solvability of (28) is guaranteed by a consequence of Lemma 1.

Proposition 2. *There exists a local diffeomorphism which transforms the system* (5), (6) *into the nonlinear observer form* (1), (2) *if and only if*

(a) *conditions* (i) *and* (ii) *of Lemma* 1 *hold,*

(b) *there exist vector fields* $g_1, \ldots, g_m$ *satisfying*

$$\left\langle \frac{\partial\left(h_l \circ F^{j-1}\right)}{\partial x}(x),\ g_i \right\rangle = \delta_{il}\delta_{jk_i}, \quad 1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant k_i,\ 1 \leqslant l \leqslant m, \tag{29}$$

such that

$$\left[\mathrm{Ad}_F^p g_i, \mathrm{Ad}_F^q g_j\right] = 0, \quad 1 \leqslant i,\ j \leqslant m,\ 0 \leqslant p \leqslant k_i - 1,\ 0 \leqslant q \leqslant k_j - 1. \tag{30}$$

Remark. Notice that Proposition 2 becomes Proposition 1 when $m = 1$. In the multi-output case, we do not choose g_i as a column of the matrix $O^{-1}(x)$. If we do, it gives us only sufficiency as was pointed out in Example 2 in [3]. But, if the observability indices are identical, i.e., $k_1 = k_2 = \cdots = k_m$, then $\mathcal{Q} \supset \mathcal{Q}_i$, for $1 \leqslant i \leqslant m$, so that condition (ii) of Lemma 1 is satisfied automatically. Further, in this case, the g_i's are determined uniquely.

Proof. Since necessary part was proved in the above, it remains to prove sufficiency. We define a map $\chi \in C^\infty(V_0, U_{x_e})$ by

$$\chi(z_{11}, \ldots, z_{mk_m}) = \Phi_{z_{11}}^{g_1} \circ \Phi_{z_{12}}^{\mathrm{Ad}_F g_1} \circ \cdots \circ \Phi_{z_{1k_1}}^{\mathrm{Ad}_F^{k_1-1} g_1} \circ \cdots \circ \Phi_{z_{m1}}^{g_m} \circ \Phi_{z_{m2}}^{\mathrm{Ad}_F g_m} \circ \cdots \circ \Phi_{z_{mk_m}}^{\mathrm{Ad}_F^{k_m-1} g_m}(x_e).$$

From the commutativity (30) we obtain

$$\frac{\partial \chi}{\partial z_{ij}}(z(k)) = \mathrm{Ad}_F^{j-1} g_i(\chi(z(k))), \quad 1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant k_i,$$

for a z in V_0. Thus,

$$\frac{\partial \chi}{\partial z}(z(k)) = \left[g_1 \quad \mathrm{Ad}_F g_1 \quad \cdots \quad \mathrm{Ad}_F^{k_1-1} g_1 \quad \cdots \quad g_m \quad \mathrm{Ad}_F g_m \quad \cdots \quad \mathrm{Ad}_F^{k_m-1} g_m \right](\chi(z(k))). \tag{31}$$

From (29), (31) we obtain the relation (27). Since $O(x_e)$ is full-rank, $(\partial\chi/\partial z)(0)$ is nonsingular. Therefore, χ is a local diffeomorphism from 0 to x_e. We can choose $T = \chi^{-1}$ as a coordinate transformation map such that $z(k+1) = \tilde{F}(z(k))$, $y(k) = h \circ \chi(z(k))$, where $\tilde{F}(z(k)) = \chi^{-1} \circ F \circ \chi(z(k))$. Then, we obtain

$$\frac{\partial \tilde{F}}{\partial z}(z(k)) = \left.\frac{\partial \chi^{-1}}{\partial x}\right|_{x(k+1)} \left[\mathrm{Ad}_F g_1 \quad \cdots \quad \mathrm{Ad}_F^{k_1} g_1 \quad \cdots \quad \mathrm{Ad}_F g_m \quad \cdots \quad \mathrm{Ad}_F^{k_m} g_m\right](x(k+1)).$$

Using (31), we obtain

$$\frac{\partial \tilde{F}_{pq}}{\partial z_{ij}}(z(k)) = \delta_{pi}\delta_{qj+1}, \quad 1 \leqslant i \leqslant m,\ 1 \leqslant j \leqslant k_i - 1,$$

where $\tilde{F}_{pq}$ denotes the pq-th component of $\tilde{F} = [\tilde{F}_{11}, \ldots, \tilde{F}_{1k_1}, \ldots, \tilde{F}_{m1}, \ldots, \tilde{F}_{mk_m}]^{\mathrm{T}}$. Thus, we deduce that $\tilde{F}_{ij}$ for $1 \leqslant i \leqslant m$, $2 \leqslant j \leqslant k_i$ is a linear function of z_{ij-1}, while all components of $\tilde{F}$ are nonlinear functions of $z_{1k_1}, z_{2k_2}, \ldots, z_{mk_m}$. Hence, we obtain

$$\tilde{F}(z(k)) = Az(k) + \alpha\left(z_{1k_1}(k), z_{2k_2}(k), \ldots, z_{mk_m}(k)\right).$$

Further, from (29), we obtain

$$\frac{\partial(h_l \circ \chi)}{\partial z_{ij}}(z(k)) = \left\langle \frac{\partial h_l}{\partial x}(x(k)), \mathrm{Ad}_F^{j-1} g_i(x(k)) \right\rangle = \delta_{il}\delta_{jk_i}$$

for $1 \leqslant l \leqslant m$, $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant k_i$. Therefore, we obtain $y(k) = Cz(k)$. □

Example. Let $x = (x_{11}, x_{12}, x_{21})$ and $z = (z_{11}, z_{12}, z_{21})$. Consider

$$\begin{bmatrix} x_{11}(k+1) \\ x_{12}(k+1) \\ x_{21}(k+1) \end{bmatrix} = \begin{bmatrix} x_{12}(k) \\ x_{21}(k) + (x_{12}(k))^2 \\ x_{11}(k) \end{bmatrix} = F(x(k)), \tag{32}$$

$$y(k) = \begin{bmatrix} x_{11}(k) \\ x_{21}(k) \end{bmatrix} = \begin{bmatrix} h_1(x(k)) \\ h_2(x(k)) \end{bmatrix}. \tag{33}$$

Since

$$\frac{\partial h_1}{\partial x}(x) = [1 \ 0 \ 0], \quad \frac{\partial(h_1 \circ F)}{\partial x}(x) = [0 \ 1 \ 0], \quad \frac{\partial h_2}{\partial x}(x) = [0 \ 0 \ 1],$$

and $k_1 = 2$, $k_2 = 1$, condition (i) of Lemma 1 is satisfied. Also, since $\mathcal{Q}_1 - \{\mathcal{Q} \cap \mathcal{Q}_1\} = \{(\partial h_2 \circ F)/\partial x)(x) = [1 \ 0 \ 0]\}$ and $\mathcal{Q}_2 - \{\mathcal{Q} \cap \mathcal{Q}_2\} = \emptyset$, where $\emptyset$ denotes the empty set, condition (ii) of Lemma 1 holds. Let $g_1(x(k)) = [0 \ 1 \ 0]^{\mathrm{T}}$ and $g_2(x(k)) = [0 \ 0 \ 1]^{\mathrm{T}}$. Then, we obtain

$$\mathrm{Ad}_F g_1(x(k)) = \left.\frac{\partial F}{\partial x}\right|_{F^{-1}(x(k))} g_1(F^{-1}(x(k))) = [1 \quad 2x_{12}(k-1) \quad 0]^{\mathrm{T}} = [1 \quad 2x_{11}(k) \quad 0]^{\mathrm{T}}.$$

Thus, $[g_1, g_2] = [g_1, \mathrm{Ad}_F g_1] = [g_2, \mathrm{Ad}_F g_1] = 0$. Hence, from (31), we obtain

$$\frac{\partial \chi}{\partial z} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 2x_{11} & 0 \\ 0 & 0 & 1 \end{bmatrix}_{x=\chi(z)}.$$

Integrating this partial differential equation, we obtain

$$x = \chi(z) = \begin{bmatrix} z_{12} \\ z_{11} + (z_{12})^2 \\ z_{21} \end{bmatrix} \quad \text{and} \quad z = T(x) = \begin{bmatrix} x_{12} - (x_{11})^2 \\ x_{11} \\ x_{21} \end{bmatrix}$$

transforming the system (32), (33) into the nonlinear observer form

$$z(k+1) = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} z(k) + \begin{bmatrix} z_{21}(k) \\ (z_{12}(k))^2 \\ z_{12}(k) \end{bmatrix}, \quad y(k) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} z(k).$$

5. Appendix

Proof of Lemma 1. Condition (i) is already proved. Obviously span $\mathscr{Q}_i \supset \operatorname{span}\{\mathscr{Q} \cap \mathscr{Q}_i\}$, $1 \leqslant i \leqslant m$. So if we show that $\dim\operatorname{span} \mathscr{Q}_i \leqslant \dim\operatorname{span}\{\mathscr{Q} \cap \mathscr{Q}_i\}$, then condition (ii) is satisfied. By (i) it is clear that $\dim\operatorname{span}\{\mathscr{Q} \cap \mathscr{Q}_i\} = i_{ki} + k_{i+1} + \cdots + k_m - 1$, $1 \leqslant i \leqslant m$. On the other hand, the following one forms are in $\mathscr{Q}_i$:

$$\begin{array}{lllll}
\partial h_1/\partial x & \partial(h_1 \circ F)/\partial x & \cdots & \partial(h_1 \circ F^{k_i-2})/\partial x & \partial(h_1 \circ F^{k_i-1})/\partial x \\
\vdots & \vdots & & \vdots & \vdots \\
\partial h_{i-1}/\partial x & \partial(h_{i-1} \circ F)/\partial x & \cdots & \partial(h_{i-1} \circ F^{k_i-2})/\partial x & \partial(h_{i-1} \circ F^{k_i-1})/\partial x \\
\partial h_i/\partial x & \partial(h_i \circ F)/\partial x & \cdots & \partial(h_i \circ F^{k_i-2})/\partial x & \\
\partial h_{i+1}/\partial x & \partial(h_{i+1} \circ F)/\partial x & \cdots & \partial(h_{i+1} \circ F^{k_i-2})/\partial x & \partial(h_{i+1} \circ F^{k_i-1})/\partial x \\
\vdots & \vdots & & \vdots & \vdots \\
\partial h_m/\partial x & \partial(h_m \circ F)/\partial x & \cdots & \partial(h_m \circ F^{k_i-2})/\partial x & \partial(h_m \circ F^{k_i-1})/\partial x.
\end{array}$$

As for the annihilating vector fields of the first row of $\mathscr{Q}_i$, we can find from (25), $\{\partial\chi/\partial z_{1l};\ 1 \leqslant l \leqslant k_1 - k_i\}$ and for the $(i-1)$-th row of $\mathscr{Q}_i$, we can find $\{\partial\chi/\partial z_{i-1l};\ 1 \leqslant l \leqslant k_{i-1} - k_i\}$. Finally, $\{\partial\chi/\partial z_{i1}\}$ annihilates the one forms in the i-th row of $\mathscr{Q}_i$. Notice that no vector field other than

$$\mathscr{V} \equiv \{\partial\chi/\partial z_{rl};\ 1 \leqslant r \leqslant i-1,\ 1 \leqslant l \leqslant k_r - k_i\} \cup \{\partial\chi/\partial z_{i1}\}$$

can annihilate all one forms in $\mathscr{Q}_i$. Since vector fields in $\mathscr{V}$ are linearly independent and

$$\dim\operatorname{span}\mathscr{V} = (k_1 - k_i) + \cdots + (k_{i-1} - k_i) + 1,$$

we obtain $\dim\operatorname{span}\mathscr{Q}_i \leqslant n - \dim\operatorname{span}\mathscr{V} = ik_i + k_{i+1} + \cdots + k_m - 1$. $\square$

References

[1] A.J. Krener and A. Isidori, Linearization by output injection and nonlinear observers, *Systems Control Lett.* **3** (1983) 47–52.

[2] A.J. Krener and W. Respondek, Nonlinear observers with linearizable error dynamics, *SIAM J. Control Optim.* **23** (1985) 197–216.

[3] Xiao-Hua Xia and Wei-Bin Gao, Non-linear observer design by observer error linearization, *SIAM J. Control Optim.* **27** (1989) 199–216.

[4] K. Nam, Linearization of discrete-time nonlinear systems and a canonical structure, *IEEE Trans. Automat. Control* **34** (1989) 119–122.

[5] F.W. Warner, *Foundations of Differentiable Manifolds and Lie Groups* (Springer-Verlag, New York, 1983).