INT. J. SYSTEMS SCI., 1988, VOL. 19, NO. 11, 2147–2155

Algebraic condition for observability of non-linear analytic systems

HEBERTT SIRA-RAMIREZ†

The observability of non-linear analytic systems can be assessed by means of Kalman's rank condition, when the system is represented in terms of the evolution equations of an associated infinite family of tensor powers of the state vector. Carleman's exact linearization provides the necessary link to transform differential-geometric observability conditions into Kalman's algebraic rank condition.

1. Introduction

Kalman's pioneering work on the algebraic nature of concepts such as controllability and observability (Kalman 1960) of linear systems, has given rise to the proposal of a number of extensions of these concepts to include different classes of systems (finite group sequential systems, infinite-dimensional systems, non-linear systems, etc.). Wonham's geometric approach (Wonham 1979) has placed these concepts in a co-ordinate-free framework from which the conceptual solution of long-standing problems in control theory can be properly found. Generalization of the geometric approach, to include the class of non-linear analytic systems, has required extensive use of differential geometry (Brockett 1976, Isidori 1985). Along with this generalization, suitable definitions of controllability and observability have been developed by Hermann and Krenner (1977), Sussmann and Jurdjevic (1972) and Krener and Respondek (1985). The reader is referred to the survey given by Andreev (1982) for a thorough account of the historical and technical details of the differential geometric approach to non-linear systems control theory.

In this note, a Carleman linearization approach is taken as the exact representation of a single-output non-linear analytic system. The observability condition, developed from differential-geometric notions (Krener and Respondek 1985), is then transformed into an algebraic test involving the Kalman rank condition on a pair of time-invariant infinite-dimensional operators. The operators describe, respectively, the evolution of the infinite family of tensor powers associated with the state vector (Brockett 1976, Sira-Ramirez 1984) and the projection of the family onto the scalar output space. Kalman's observability condition is rederived when the results are applied to linear time-invariant systems. The observability condition for linear systems with polynomial and non-linear analytic output maps, exhibits a decomposition into linear subsystem observability conditions.

Section 2 contains definitions and background results on families of tensor powers for vectors and matrices while § 3 presents the main technical result and some illustrative examples. The last section presents the conclusions and suggestions for further research.

Received 1 May 1987.

† Departamento Sistemas de Control, Escuela Ingeniería de Sistemas, Universidad de Los Andes, Mérida, Venezuela.

2. Notation, definitions and basic background results

In this section some definitions regarding tensor powers of vectors, matrices and their infinitesimal versions are presented (Brockett 1973, Sira-Ramirez 1984).

Let x be an n-vector with components $x_1, x_2, \ldots, x_n$ then the vector $x^{[p]}$ denotes the $N(n, p)$-dimensional vector

$$N(n,p) := \binom{n+p-1}{n}$$

of homogeneous p-forms in the components of x. By convention we set $x^{[0]} = 1$. The elements of the vector $x^{[p]}$ are of the form $\prod_{i=1}^{m} x_i^{p_i}$ with $\sum p_i = p$, $p_i > 0$. This 'power' of the vector x is usually addressed as the pth tensor power of x.

If $y = Ax$ then $y^{[p]} = A^{[p]}x^{[p]}$ is verified. The matrix $A^{[p]}$ is called the pth tensor power of the matrix A. The infinitesimal version of $A^{[p]}$ is denoted by $A_{[p]}$ and defined as the constant matrix satisfying

$$\frac{d}{dt} x^{[p]} = A_{[p]} x^{[p]}$$

whenever

$$\frac{d}{dt} x = Ax$$

Let $\bar{x} := \begin{bmatrix} 1 \\ x \end{bmatrix}$, then for any integer p, $\bar{x}^{[p]} := [1 \quad x^{\mathrm{T}} \quad (x^{[2]})^{\mathrm{T}} \quad \ldots \quad (x^{[p]})^{\mathrm{T}}]^{\mathrm{T}}$ defines the $\binom{n-p}{p}$-dimensional vector known as the pth family of tensor powers of x. If $p = \infty$ then $\bar{x}^{[\infty]}$ denotes the infinite family of tensor powers of x. Extension of these definitions to matrices is straightforward and results in

$$\bar{A}^{[p]} = \operatorname{diag}[1 \quad A \quad A^{[2]} \quad \ldots \quad A^{[p]}]$$

and

$$\bar{A}^{[\infty]} = \operatorname{diag}[1 \quad A \quad A^{[2]} \quad \ldots \quad A^{[k]} \quad \ldots]$$

Similarly

$$\bar{A}_{[p]} = \operatorname{diag}[0 \quad A \quad A_{[2]} \quad \ldots \quad A_{[p]}]$$

and

$$A_{[\infty]} = \operatorname{diag}[0 \quad A \quad A_{[2]} \quad \ldots \quad A_{[k]} \quad \ldots]$$

Consider the non-linear analytic system

$$\frac{d}{dt} x = f(x) \tag{2.1}$$

A Taylor expansion for f is guaranteed to converge in some neighbourhood of the initial state $x_0 := 0$, thus allowing the representation of $f(x)$ as

$$\begin{aligned} f(x) &= f(0) + \sum_{p=1}^{\infty} \sum_{i_1+i_2+\ldots+i_n=p} \left| \frac{1}{p!} \frac{\partial^{i_1+\ldots+i_n}}{\partial x_1^{i_1} \ldots \partial x_n^{i_n}} f(x) \right|_{x=0} x_1^{i_1} x_2^{i_2} \ldots x_n^{i_n} \\ &=: [F_{(0)}^1 \quad F_{(1)}^1 \quad \ldots \quad F_{(k)}^1 \quad \ldots] \bar{x}^{[\infty]} \end{aligned} \tag{2.2}$$

with $F^1_{(0)} = f(0)$, henceforth assumed to be also zero. The matrix $F^1_{(k)}$ has dimensions $n \times N(n, k)$ and is given by

$$F^1_{(k)} = \sum_{i_1+i_2+\ldots+i_n=k} \left| \frac{1}{k!} \frac{\partial^{i_1+\ldots+i_n} f(x)}{\partial x_1^{i_1} \ldots \partial x_n^{i_n}} \right|_{x=0} \tag{2.3}$$

Lemma 1

The evolution of $x^{[\infty]}$ is governed by

$$\frac{d}{dt}\bar{x}^{[\infty]} = A\bar{x}^{[\infty]} \tag{2.4}$$

with A an infinite-dimensional matrix operator having the following structure:

$$A = \begin{bmatrix} 0 & 0 & 0 & \ldots & 0 & \ldots \\ 0 & F^1_{(1)} & F^1_{(2)} & \ldots & F^1_{(k)} & \ldots \\ 0 & 0 & F^2_{(2)} & \ldots & F^2_{(k)} & \ldots \\ 0 & 0 & 0 & \ldots & F^3_{(k)} & \ldots \\ \vdots & \vdots & \vdots & & \vdots & \\ 0 & 0 & 0 & \ldots & F^{k-1}_{(k)} & \ldots \\ 0 & 0 & 0 & \ldots & F^k_{(k)} & \ldots \\ \vdots & \vdots & \vdots & & \vdots & \end{bmatrix} \tag{2.5}$$

where $F^j_{(i)}$ is an $N(n,j) \times N(n,i)$ matrix $(i \geqslant j)$ obtained by means of linear operations on $F^{j-1}_{(i-1)}$. The matrices $F^j_{(i)} = 0$ for $i \leqslant j$.

Proof

With minor notational variations, and a thorough exposition of the finite approximation properties of the scheme, the result can be found in the work by Loparo and Blankenship (1978). Also, the results given by Krener (1974) directly generate (2.5) when no control inputs are considered. □

Let $h(x)$ be a scalar analytic function of x representing the system output as

$$y = h(x) \tag{2.6}$$

A Taylor series expansion of $h(x)$ results in

$$y = [h_{(0)} \quad h_{(1)} \quad h_{(2)} \quad \ldots \quad h_{(k)} \quad \ldots]\bar{x}^{[\infty]} := c\bar{x}^{[\infty]} \tag{2.7}$$

where

$$h_{(k)} = \sum_{i_1+i_2+\ldots+i_n=k} \left[\frac{1}{k!} \frac{\partial^{i_1+\ldots+i_n} h(x)}{\partial x_1^{i_1} \ldots \partial x_n^{i_n}} \right]_{x=0} \tag{2.8}$$

It will be assumed henceforth that $h_{(0)} = h(0) = 0$.

Remark 1

Even though (2.4) and (2.7) constitute a linear system of differential equations of the form:

$$\frac{d}{dt}\xi = A\xi, \quad y = c\xi$$

the vector ξ can only be considered formally as a state due to its inherent redundancy. Also, the initial values for ξ cannot be chosen freely because of the non-linear interdependence among its infinite components. The set of possible values for ξ are constrained to an algebraic variety embedded in the underlying infinite-dimensional euclidean space of ξ. The direct use of Kalman's rank condition on the matrix $[c \quad cA \quad cA^2 \quad \ldots \quad cA^{n-1}]$ must be rigorously justified by means other than the usual derivation (Kailath 1980). A crucial argument of the standard observability-condition derivation requires that ξ be considered as a free vector, which is not true in this case. For this reason, the local differential-geometric definition of observability (Krener and Respondek 1985) will be used to obtain such a justification.

For reference purposes the following standard definitions are introduced (Isidori 1985). Let h be an analytic function. The differential of h is the row vector

$$dh := \left[\frac{\partial h}{\partial x_1} \quad \frac{\partial h}{\partial x_2} \quad \cdots \quad \frac{\partial h}{\partial x_n}\right] = \frac{\partial h}{\partial x} \tag{2.9}$$

Let f be an analytic vector field. The Lie derivative of h with respect to f is given by the analytic scalar function

$$L_f h := \frac{\partial h}{\partial x_1} f_1 + \frac{\partial h}{\partial x_2} f_2 + \ldots + \frac{\partial h}{\partial x_n} f_n = \langle dh, f\rangle = \frac{\partial h}{\partial x} f \tag{2.10}$$

The Lie derivative of the gradient of h with respect to f is defined as

$$L_f(dh) = d(L_f h) \tag{2.11}$$

Let $w(x)$ be an m-dimensional analytic vector field. The jacobian of w is an analytic matrix with each row vector being the differential of the corresponding component of $w(x)$, i.e.

$$dw(x) = \begin{vmatrix} \frac{\partial w_1}{\partial x_1} & \frac{\partial w_1}{\partial x_2} & \cdots & \frac{\partial w_1}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial w_m}{\partial x_1} & \frac{\partial w_m}{\partial x_2} & \cdots & \frac{\partial w_m}{\partial x_n} \end{vmatrix} \tag{2.12}$$

Let c be a constant row vector and $cw(x) = \langle c^T, w(x)\rangle$ a scalar analytic function. The differential of the function in terms of the jacobian of w is simply

$$d(cw(x)) = c\, dw(x) \tag{2.13}$$

With the obvious extension of the jacobian to infinite-dimensional vector fields, such as those represented by $x^{[\infty]}$, the following proposition holds true.

Proposition

$$\text{rank } d\bar{x}^{[\infty]} = n \tag{2.14}$$

Proof

From the definition of $x^{[\infty]}$ and using (2.12), it is easy to see that the second 'block' component of the jacobian is the identity matrix in $\mathbb{R}^n$. □

The following lemma is a well-known result about the rank of a matrix product (Kailath 1980).

Lemma 2

Let A be an $n \times p$ matrix and B a $p \times n$ matrix $p > n$. Then

$$\operatorname{rank}(AB) \leqslant \min[\operatorname{rank} A, \operatorname{rank} B]$$

If B is rank n then, $\operatorname{rank}(AB) = \operatorname{rank} A$. The result holds even if $p = \infty$.

Remark 2

Lemma 2 allows us to say that with rank $B = n$, then rank $(AB) = n$ if and only if rank $A = n$.

3. Main result

In this section the consequence of the observability property, in a non-linear system of the form (2.1) and (2.6) is explored in terms of the infinite-dimensional representation (2.4) and (2.7).

Definition (Krener and Respondek 1985)

The system

$$\left.\begin{aligned} \frac{d}{dt}x = f(x), &\quad x \in \mathbb{R}^n \\ y = h(x), &\quad y \in \mathbb{R} \end{aligned}\right\} \tag{3.1}$$

is observable at x_0 if there exists a neighbourhood U of x_0 such that the set of n row vectors

$$\{L_f^{j-1}(dh), \quad j = 1, 2, \ldots, n\} \tag{3.2}$$

is a linearly independent set.

Theorem

The non-linear analytic system (f, h) is observable if and only if the matrix

$$C = \begin{vmatrix} c \\ cA \\ cA^2 \\ \vdots \\ cA^{n-1} \end{vmatrix} \tag{3.3}$$

has rank n.

Proof

It will be proved by induction that

$$L_f^j(dh) = cA^j \, d\bar{x}^{[\infty]} \tag{3.4}$$

From (2.11) we have $dh = d(c\bar{x}^{[\infty]}) = c\, d\bar{x}^{[\infty]}$, then

$$\begin{aligned} L_f(dh) &= dL_f h = d\left(\frac{\partial h}{\partial x} f\right) = d\left(\frac{\partial h}{\partial x} f\right) = d\left(\frac{\partial h}{\partial x}\dot{x}\right) = d\left(\frac{dh}{dt}\right) \\ &= d\left(\frac{dy}{dt}\right) = d\left(c\frac{d}{dt}\bar{x}^{[\infty]}\right) = d(cA\bar{x}^{[\infty]}) \\ &= cA\, d\bar{x}^{[\infty]} \end{aligned}$$

thus (3.4) is valid for $j = 1$. Assume (3.4) is true for $j = k$. It then follows that

$$\begin{aligned} L_f^{k+1}(dh) &= L_f(cA^k\, d\bar{x}^{[\infty]}) = L_f[d(cA^k\bar{x}^{[\infty]})] = dL_f[cA^k\bar{x}^{[\infty]}] \\ &= dL_f\left(\frac{d^k}{dt^k}y\right) = d\left[\frac{\partial}{\partial x}\left(\frac{d^k}{dt^k}y\right)f\right] = d\left[\frac{\partial}{\partial x}\left(\frac{d^k}{dt^k}y\right)\dot{x}\right] \\ &= d\left[\left(\frac{d^{k+1}}{dt^{k+1}}y\right)\right] = d(cA^{k+1}\bar{x}^{[\infty]}) = cA^{k+1}\, d\bar{x}^{[\infty]} \end{aligned}$$

which proves that (3.4) is true for all $j = 0, 1, \ldots, n-1$. Rewriting (3.4) in matrix form

$$\begin{vmatrix} dh \\ L_f\, dh \\ L_f^2\, dh \\ \vdots \\ L_f^{n-1}\, dh \end{vmatrix} = \begin{vmatrix} c \\ cA \\ cA^2 \\ \vdots \\ cA^{n-1} \end{vmatrix} d\bar{x}^{[\infty]} = C\, d\bar{x}^{[\infty]} \tag{3.5}$$

Since $d\bar{x}^{[\infty]}$ has rank n, it follows from Lemma 2 in § 2 that the matrix in the left-hand side of (3.5) has rank n if and only if C has rank n. □

Example 1

Consider the linear system $\dot{x} = Fx$, $y = hx$, then in this case, $c = [0 \quad h \quad 0^T \quad \ldots \quad 0^T \quad \ldots]$ and

$$A = \text{diag}\,[0 \quad F \quad F_{[2]} \quad \ldots \quad F_{[k]} \quad \ldots]$$

The matrix C is given by

$$C = \begin{vmatrix} c \\ cA \\ cA^2 \\ \vdots \\ cA^{n-1} \end{vmatrix} = \begin{vmatrix} 0 & h & 0 & \ldots \\ 0 & hF & 0 & \ldots \\ 0 & hF^2 & 0 & \ldots \\ \vdots & \vdots & \vdots & \\ 0 & hF^{n-1} & 0 & \ldots \end{vmatrix} \tag{3.6}$$

Then C has tank n if and only if the set of vectors $\{hF^j : j = 0, 1, 2, \ldots, n-1\}$ are linearly independent. This rederives the well known Kalman observability condition for linear time-invariant systems (Kalman 1960).

Example 2

Consider the class of linear time-invariant systems with non-linear analytic output maps. In this case the infinite row vector c, defining the scalar output, can be partitioned in correspondence with the dimensions of the successive tensor power components of $x^{[\infty]}$, i.e.

$$y = [h_0 \quad h_{(1)} \quad h_{(2)} \quad \dots \quad h_{(p)} \quad \dots]\bar{x}^{[\infty]} \tag{3.7}$$

while A is as in the previous example. In this case C is given by

$$C = \begin{vmatrix} h_0 & h_{(1)} & \dots & h_{(p)} & \dots \\ 0 & h_{(1)}F & \dots & h_{(p)}F_{[p]} & \dots \\ \vdots & \vdots & & \vdots & \\ 0 & h_{(1)}F^{n-1} & \dots & h_{(p)}F_{[p]}^{n-1} & \dots \end{vmatrix} \tag{3.8}$$

Using standard facts about the rank of partitioned matrices, it is concluded that for the system $\dot{x} = Fx$ and $y = h(x)$ to be observable, it is sufficient that the finite-dimensional pair $(h_{(i)}, F_{[i]})$ generates a rank n observability matrix for some $i = 1, 2, 3, \dots$. In particular, if for some i the pair $(h_{(i)}, F_{[i]})$ is observable, then the system (h, F) is also observable. The converse is not necessarily true. For this class of systems a finite expansion of the output map in a Taylor series may suffice to establish observability.

As a corollary to the previous example, linear systems with polynomial output maps

$$y = [h_{(0)} \quad h_{(1)} \quad \dots \quad h_{(p)} \quad 0^{\mathrm{T}} \quad 0^{\mathrm{T}} \quad \dots]\bar{x}^{[\infty]} = \sum_{i=0}^{p} h_{(i)} x^{[i]}$$

are observable if for some $i = 1, 2, \dots, p$ the pair $(h_{(i)}, A_{(i)})$ generates a rank n observability matrix.

Example 3

For the non-linear system $\dot{x} = f(x)$ with a linear output map $y = cx$ a sufficient condition for observability is readily obtained after computation of the matrix C ('*' denotes a non-zero matrix)

$$C = \begin{vmatrix} 0 & c & 0 & 0 & \dots \\ 0 & cF_{(1)}^1 & cF_{(2)}^1 & * & \dots \\ 0 & c[F_{(1)}^1]^2 & * & * & \dots \\ \vdots & \vdots & \vdots & \vdots & \\ & c[F_{(1)}^1]^{n-1} & * & * & \dots \end{vmatrix} \tag{3.9}$$

It follows that for the non-linear system to be observable it is sufficient that the linearized system $(c, F^1_{(1)})$ is observable. Notice that even if the linearized system is non-observable, the matrix C may still have rank n, hence the sufficiency of the condition.

Example 4

The result of Example 3 can be applied to well-known non-linear systems such as robot manipulators with multiple links and linear scalar output maps (such as joint position, joint velocity or linear combinations of such variables). The closed-loop system of differential equations, in terms of the vector of joints positions x, is of the form $\ddot{x} = f(x, \dot{x})$. A state-space representation is obtained by letting $x_1 = x : x_2 = \dot{x}$

$$\dot{x}_1 = x_2, \quad \dot{x}_2 = f(x_1, x_2), \quad y = c_1 x_1 + c_2 x_2 \tag{3.10}$$

The system is observable if the linearized pair

$$\left((c_1, c_2), \begin{bmatrix} 0 & I \\ F^{11}_{(1)} & F^{21}_{(1)} \end{bmatrix} \right) \tag{3.11}$$

is observable, with

$$F^{11}_{(1)} = \frac{\partial f}{\partial x_1} : F^{21}_{(1)} = \frac{\partial f}{\partial x_2}$$

4. Conclusions and suggestions for further research

In this paper, an algebraic condition has been derived for the observability of non-linear analytic systems of a general form. The algebraic condition is a rank condition of Kalman's type on a pair of infinite-dimensional maps. One of the maps describes the dynamic evolution of the infinite family of tensor powers generated by the state vector; the second map represents the projection of this family onto the output space. Carleman's linearization constitutes the key procedure to establish a link among the differential-geometric definition of observability and the algebraic rank condition of non-linear systems observability. This connection reveals the conceptual simplicity and generality of Kalman's cornerstone contribution.

The generalization here explored can be used in computer calculations for the assessment of non-linear systems observability. Whether this condition is easier to implement in software subroutines than its differential-geometric counterpart (Krener and Respondek 1985) remains to be demonstrated.

We point out that similar results can be obtained using the cartesian tensor formalism. In this case, Kalman's condition remains valid in a space of tensor operators acting on the tensor space of infinite copies of the Lebesgue space ℓ^2. The work of Banks (1986), Banks and Ashtiani (1985), and Banks and Yew (1985) contains sufficient information for this task.

Interesting connections might arise for the class of exactly output-linearizable non-linear systems— via diffeomorphic state co-ordinate transformation and output injections (Krener and Respondek 1985)—and the Carleman exact linearization approach. A crucial task in this problem is to establish the role of the Kronecker observability indices in the context of Carleman's exact linearization.

Acknowledgments

This work was partially supported by the Joint Services Electronics Program under contract no. N00014-C-0149 and by the Consejo de Desarrollo Cientifico, Humanístico y Technológico of the Universidad de Los Andes, Mérida, Venezuela.

This work was initiated while the author was a visiting Professor at the Coordinated Science Laboratory of the University of Illinois at Urbana-Champaign.

References

ANDREEV, Y. N., 1982, *Automn remote Control*, **43,** 1231.

BANKS, S. P., 1986, *Systems Control Lett.*, **6,** 337.

BANKS, S. P., and ASHTIANI, A., 1985, *Int. J. Systems Sci.*, **16,** 841.

BANKS, S. P., and YEW, M. K., 1985, *Systems Control Lett.*, **5,** 327.

BROCKETT, R. W., 1973, Lie algebras and Lie groups in control theory. *Geometric Methods in Systems Theory*, edited by R. W. Brockett and D. Q. Mayne (The Netherlands: Reidel); 1976, *Proc. Inst. elect. electron. Engrs*, **64,** 61.

HERMANN, R., and KRENNER, A., 1977, *I.E.E.E. Trans. autom. Control*, **22,** 728.

ISIDORI, A., 1985, *Nonlinear Control Systems: An Introduction*, Lecture Notes in Control and Information Sciences, Vol. 72 (New York: Springer-Verlag).

KAILATH, T., 1980, *Linear Systems* (Englewood Cliffs, NJ: Prentice-Hall).

KALMAN, R., 1960, On the general theory of control systems. *Proc. IFAC 1st International Congress*, Vol. 1, pp. 481–492.

KRENER, A., 1974, Linearization and bilinearization of control systems. *Proc. 12th Allerton Conf. on Circuits and System Theory*, University of Illinois at Urbana-Champaign, pp. 834–843.

KRENER, A. J., and RESPONDEK, W., 1985, *SIAM J. Control Optim.*, **23,** 197

LOPARO, K., and BLANKENSHIP, G., 1978, *I.E.E.E. Trans. autom. Control*, **23,** 602.

SIRA-RAMIREZ, H., 1984, A bilinear observer approach for a class of nonlinear state reconstruction problems. *Proc. IFAC 9th Triennial World Congress*, Budapest, Hungary, pp. 423–428.

SUSSMANN, H. J., and JURDJEVIC, V. J., 1972, *J. diff. Eqns*, **12,** 95.

WONHAM, M. W., 1979, *Linear Multivariable Control: A Geometric Approach* (New York: Springer-Verlag).