TP3-5:00

NONLINEAR ZERO DISTRIBUTIONS

A J Krener
Department of Mathematics
University of California
Davis, California 95616

A Isidori
Instituto di Automatica
Università di Roma
Via Eudossiana 18
00184 Roma

Introduction

This paper generalizes the concept of a zero from a linear to nonlinear system. We start by briefly reviewing the concept of zero for a linear system. This is quite a complicated idea and many different definitions have been proposed. Of necessity we cannot discuss them all and we refer the reader to the excellent survey paper of MacFarlane and Karcanias [1] for a full treatment. The particular definition that we shall adopt is from Desoer and Schulman [2].

Consider the linear system

(1.1) $\dot{x} = Ax + Bu$

(1.2) $y = Cx$

(1.3) $x(0) = x^o$

where x, u, and y are nx1, mx1 and px1 respectively and A, B and C are sized accordingly. The transfer function

(1.4) $T(s) = C(sI - A)^{-1}B$

is an pxm matrix of strictly proper functions. We assume that T(s) is generally of full rank, i.e., except for a finite subset of the complex plane, the rank of T(s) is equal to min {p,q} and also we assume that (1.1) (1.2) and (1.3) are a minimal realization of T(s).

Definition [2] A point z is a zero of the transfer function T(s) if the rank of T(z) is less than min {p,q}.

Linear Zeroes

In this section we review the implications of the above definition. Most of what is here and a whole lot more can be found in [1] and [2]. We distinguish the two cases, $m \le p$ and $m \ge p$.

If $m \le p$ then z is a zero of T(s) if there exists a nonzero vector f such that

(2.1) $T(z)\, f = 0.$

(Of course f could be complex but we shall proceed as if it were real. The other case is easily handled by combining real and imaginary parts). If we define $x^o = (zI - A)^{-1}Bf$ then it is easily seen that the initial state x^o and input $u(t) = fe^{zt}$ results in zero output.

Moreover

(2.2) $x^o \in \ker C$

and z, x^o are a right eigenpair of A modulo B, i.e.

(2.3) $Ax^o = zx^o - Bf$

On the other hand if the initial state x^o and the input $u(t) = e^{zt}$ results in zero output a straightforward calculation shows that (2.2) holds and that there exists f such that (2.3) holds so z is a zero of T(s) as before (2.1). The vectors f and x^o are called the input and state zero vectors respectively

By linearity if z^i, x^i and f^i satisfy (2.2) and (2.3) for $i = 1,\dots,k$ and x^o is a linear combination of $\{x^1,\dots,x^k\}$

$$x^o = \sum_{i=1}^{k} \alpha_i x^i$$

then there exists an input

$$u(t) = \sum_{i=1}^{k} \alpha_i f^i e^{z^i t}$$

resulting in zero output. Therefore it is natural to consider the space of all state zero directions, i.e. all initial conditions x^o for which there exists an input which zeroes the output. This is precisely $\mathcal{V}^*$, the maximal (A, B) invariant subspace contained in the kernel of C in the sense of Wonham and Morse. Recall a subspace $\mathcal{V}$ is (A,B) invariant if

(2.4) $A\mathcal{V} \subset \mathcal{V} + \mathcal{B}$

where B is the column space of B.

Equivalently there exists mxn matrix F such that

(2.5) $(A + BF)\mathcal{V} \subseteq \mathcal{V}$

Each real zero z corresponds to real state zero direction x^o generating a one dimentional (A,B) invariant subspace. Each complex zero corresponds to a complex conjugate pair of state zero directions x^o and $\bar{x}^o$. The real and imaginary parts span a 2 dim. (A,B) invarient subspace. Furthermore a zero of multiplicity greater than one may correspond to a higher dimensional (A,B) invariant subspace because of the possibility of generalized eigenvectors of A modulo $\mathcal{B}$[1, p52].

A generalized eigenpair of A modulo B is a z and x^i satisfying for some f^i

$$Ax^i = zx^i + x^{i-1} - Bf^i$$

where z and x^{i-1} is an eigenpair or generalized eigenpair of A modulo B.

For any $x^o \in \mathcal{V}^*$ the input which zeroes the output is given by

$$u(t) = F\exp((A + BF)t)x^o,$$

where

$$(A + BF)\mathcal{V}^* \subset \mathcal{V}^*$$

When $m \le p$, zeroes correspond to our intuitive notion of initial states and inputs which result in zero output. Alternatively we can think of F as defining a state feedback law $u = Fx$ which makes the system as unobservable as possible. From this point of view a zero can be thought of as the potential for the loss of observability via state feedback.

If $m > p$ then for any z there always exists an input $u(t) = fe^{zt}$ which zeroes the output when the system is initialized at $x^o = 0$, that is, there always exists m-p linearly independent vectors f satisfying (2.1) for each z.

Presented at 19th CDC, Albuquerque, 1980.

Research supported in part by NSF under MCS-8003263.

0191-2216/80/0000-0665$00.75 © 1980 IEEE

The zeroes are those values of z for which there exists more than $m-p$ linearly independent f's satisfying (2.1). We could continue as before and relate the zeroes to (A,B) invariant subspaces in ker C, but in this case the dimension of V^* exceeds the number of zeroes by the dimension of R^*, the maximal (A,B) controllable subspace in kernel C. We refer the reader to Ansaklis [3] for details.

Instead when $m \geq p$ we switch to a dual approach. From the definition it is immediate that z is a zero if there exists a nonzero $1 \times p$ vector g such that

(2.6) $\quad gT(z) = 0$

We define $\xi^0 = gC(zI - A)^{-1}$ and call g and ξ^0 zero observation and state covectors repectively

Clearly

(2.7) $\quad \xi^0 B = 0$

and z, ξ^0 is a left eigenpair of A modulo C.

(2.8) $\quad \xi^0 A = \xi^0 z - gC$

If we define a linear subspace V

$$V = \{x : \xi^0 x = 0\}$$

then V is a (C,A) invariant subspace containing B because of (2.7). Recall a (C,A) invariant subspace V is one satisfying

(2.9) $\quad A(V \cap \ker C) \subset V$

or equivalently there exists a mxp matrix G such that

(2.10) $\quad (A + GC)V \subset V$

Given z^i, ξ^i and g^i, satisfying (2.7) and (2.8) for $i = 1,\ldots,k$ and $\{\xi^i\}$ linearly independent we can define

$$V = \{x : \xi^i x = 0,\ i = 1,\ldots,k\}$$

then V is a (C,A) invariant subspace in ker C and the codimention of V is k, the number of zeroes.

It is natural at this point to consider V_*, the minimal (C,A) invariant subspace containing $\mathcal{B}$. The codimention of V_* is equal to the number of zeroes of the system. If G is a nxp matrix such that

(2.11) $\quad (A + GC)V_* \subset V_*$

we can view the matrix A + GC as the result of modifying the dynamics (1.1) by output injection.

(2.12) $\quad \dot{x} = Ax + Gy + Bu = (A + GC)x + Bu$

Equation (2.11) guarantees that if $x^0 \in V_*$ then $x(t)$ the solution of (2.12) remains in V_* for all u(t). In other words output injection has resulted in a loss of controllability and the number of uncontrollable modes equals the number of zeroes.

In summary we would like to stress the following point of view. If $m \leq p$ it is convenient to think of the zeroes of T(s) in terms of the right eigenpairs (or generalized eigenpairs) of A mod B, which are contained in the kernel of C. They represent a potential for loss of observability under state feedback. The zero state vectors are the right eigenvectors of A mod B in kernel C and they along with the generalized eigenvectors of A mod B in kernel C generate V^*, the maximal (A,B) invariant subspace in kernel C.

If $m > p$ then V^* contains additional directions, namely R^*, the maximal (A,B) controllable subspace in kernel C so it is convenient to think of the zeroes of T(s) in terms of the left eigenpairs (or generalized eigenpairs) of A mod C which annihilate B. They represent a potential for loss controllability under output injection. The zero state covectors are the left eigenvector of A mod C annihilating B and they and the generalized ones span the annihilator of V_*, the minimal (C,A) invariant subspace containing $\mathcal{B}$.

Pre and Post Processing

The intuitive idea of a zero as initial state and input which zeroes the output is quite natural. Therefore the characterization of zeroes for $m \leq p$ in terms of loss of observability via state feedback is consistant with this intuition. However the characterization of zeroes for $m \geq p$ in terms of loss of controllability via output injection is not so natural and demands further study. Moreover there is no reason to restrict these characterizations to the case $m \leq p$ and $m \geq p$, they are both valid for all values of m and p provided one exercises a little care. This would take us too far afield from the purposes of this note so we shall not do so here. In [2] the authors give loss of observability characterization of zeroes for all m and p.

Suppose we apply state feedback $u = Fx + v$ the system (1.1) (1.2) (1.3) so that the result is

(3.1) $\quad \dot{x} = (A + BF)x + Bv$

(3.2) $\quad y = Cx$

(3.3) $\quad x(0) = x^0$

The new transfer funtion is

$$T_F(s) = C(sI - A - BF)^{-1}B$$

which can be factored as

$$T_F(s) = C(sI - A)^{-1}B(I - F(sI - A)^{-1}B)^{-1}$$

The second factor $(I - F(sI - A)^{-1}B)^{-1}$ can be thought of as a preprocessor and is equal to

$$I + F(sI - A - BF)^{-1}B$$

which is realised by

(3.4) $\quad \dot{z} = (A + BF)z + Bv$

(3.5) $\quad u = Fz + v$

(3.6) $\quad z(0) = z^0$

From this we see that if (3.4) (3.5) (3.6) is used to preprocess v before (1.1) (1.2) (1.3) the resulting output is the same as that for the feedback system (3.1) (3.2) (3.3) provided $x^0 = z^0$. Moreover a straightforward calculation shows that when original system is preprocessed

$$\dot{x} - \dot{z} = A(x - z)$$

so $x(t) = z(t)$ for all t and v(t).

Now suppose V^* and F are as before and $x^0 = z^0 \in V^*$ then $z(t) \in V^*$ for all t so the output of the preprocessed system is $Cx(t) = Cz(t) = 0$ and of course the same is true for the feedback system. If $x^0 = z^0 \notin V^*$ then the output need not be zero but it will be the same over all initial conditions in $x^0 + V^*$.

Suppose instead we use output injection resulting in

$$(3.7) \quad \dot{x} = (A + GC)x + Bu$$

$$(3.8) \quad w = Cx$$

$$(3.9) \quad x(0) = x^0$$

The reader might object at this point that output injection is not physically realizable. But it is certainly no less realizable than state feedback. Of course we have just seen how state feedback can be theoretically realized via preprocessing if the initial state is known. It turns out that output injection can be theoretically realized via postprocessing under the same assumption.

The transfer function of (3.9) (3.10) (3.11) is

$$_G T(s) = C(sI - A - GC)^{-1}B$$

which can be factored as

$$_G T(s) = (I - C(sI - A)^{-1}G)^{-1}C(sI - A)^{-1}B$$

The first factor can be thought of as a postprocessor and is equal to

$$I + C(sI - A - GC)^{-1}G$$

which is realized by

$$(3.10) \quad \dot{z} = (A + GC)z - Gy$$

$$(3.11) \quad w = -Cz + y$$

$$(3.12) \quad z(0) = z^0$$

Therefore (1.1) (1.2) (1.3) followed by postprocessor (3.10) (3.11) (3.12) is equivalent to the output injected system (3.7) (3.8) (3.9) provided $x^0 = z^0$.

For the original system postprocessed we have

$$\dot{x} - \dot{z} = (A + GC)(x - z) + Bu.$$

Let V^* and G be as before then $x^0 - z^0 = 0 \in V^*$ and so $x(t) - z(t) \in V^*$ for all inputs $u(t)$ therefore we can track $x(t)$ mod V^* via postprocessing. In fact we can do better than that for from w we can recover those components of $x(t) - z(t)$ which are not in kernel C since $w = C(x - z)$ and so we can track $x(t)$ mod$(V^* \cap \ker C)$. Even if $x^0 - z^0 \neq 0$ then $x(t) - z(t)$ modulo V^* is independent of $u(t)$. From this discussion we see the importance of system zeroes for tracking and filtering.

Nonlinear Zero Distributions

Frequency domain methods breakdown when dealing with nonlinear systems so it does not appear to be possible to directly generalize the concept of a zero frequency. However many geometric notions of linear systems theory do carry over.

For example basic tools for geometric multivariable systems theory such as (A,B) and (C,A) invariant subspaces have been generalized and used to solve problems of decoupling [4]. We use these same tools to extend the notions of state zero vector and covector. This is closely related to [5] in which the loss of controllability and observability of cascaded nonlinear systems is discussed. Such loss of controllability and observability can be thought of as the matching of a zero of the first (second) system with a mode of the first (second) The nonlinear systems which we are considering are

$$(4.1) \quad \dot{x} = f(x) + g(x)u$$

$$(4.2) \quad y = h(x)$$

$$(4.3) \quad x(0) = x^0$$

where x denotes local coordinates on an m dimensional C^∞ manifold M and as before $u \in R^m$, $y \in R^p$. The function $h: M \to R^m$ and the vector fields f and the m columns of g are assumed to be C^∞.

Let Δ be a <u>distribution</u> on M, that is, for each $x \in M$, $\Delta(x)$ is a linear subspace of T_xM the tangent space to M at x. We confuse Δ with space of vector fields which are pointwise in it. Following [4] we say that Δ is (f,g) invariant if there exists $\alpha(x)$ and $\beta(x)$ such that

$$(4.4) \quad [\tilde{f}, \Delta] \subset \Delta$$

$$(4.5) \quad [\tilde{g}, \Delta] \subset \Delta$$

where $\tilde{f} = f + g\alpha$ and $\tilde{g} = g\beta$.

Such distributions correspond to (A,B) invariant subspaces. Of course every subspace V of R^n induces an equivalence relation mod V and the quotient is the vector space R^n/V. Not every distribution Δ induces an equivalence relation on M in such a way that the quotient M/Δ is a nice C^∞ manifold. Those that do are called regular [4, p6] and henceforth we shall assume all distributions discussed in this paper regular. Necessary conditions for regularity are that Δ be of constant dimension and involutive but they are not sufficient.

A distribution Δ is a <u>null observable distribution</u> if it is (f,g)invariant and contained in kernel dh. It is not always true that a maximal such distribution exists, if it does we denote is by Δ^*. Given such a Δ, corresponding to each initial condition x^0 there is an input $u(t; x^0)$ defined as follows. Let $x(t; x^0)$ be the solution of $\dot{x} = \tilde{f}(x)$, $x(0; x^0) = x^0$ and $u(t; x^0) = \alpha(x(t; x^0))$. The system (4.1) (4.2) (4.3) is initialled at x^0 and driven by $u(t; x^0)$ the output $y(t; x^0)$ will be same over the Δ equivalence class of x^0. More generally we can allow an open loop control $v(\cdot)$ by defining $x(t; x^0, v(\cdot))$ as the solution of $\dot{x} = \tilde{f}(x) + \tilde{g}(x)v$, $x(0; x^0, v(\cdot)) = x^0$ and $u(t; x^0, v(\cdot)) = \alpha(x(t; x^0, v(\cdot)) + \beta(x(t; x^0, v(\cdot)))v(t)$. For eached fixed $v(\cdot)$, the output from x^0, $u(t; x^0, v(\cdot))$ is constant over the Δ equivalent class of x^0. That the output is constant over equivalence classes is the nonlinear generalization of the output being zero. The above construction is the nonlinear version of loss of observability via state feedback.

We turn now to the second way of characterizing zeroes, via loss of controllability. A distribution Δ is (h,f) invariant if

$$(4.6) \quad [f, \Delta \cap \ker dh] \subset \Delta.$$

$$(4.7) \quad [g, \Delta \cap \ker dh] \subset \Delta$$

Such a distribution is a <u>null controllable distribution</u> if it is regular and the columns of g(x) are in $\Delta(x)$ for each x. It is not always true that a minimal such distribution exists but if it does we denote it by Δ_*. It is the nonlinear analog of V_*

Given a controllability zero distribution Δ we can construct a postprocessor which accepts the output y and initial state x^0 and calculates the equivalence class of x(t) modulo ($\Delta \cap$ ker dh) without knowledge of the input u(t). This is the nonlinear analog of loss of controllability via output injection. The details are in [4].

In conclusion, while the frequency domain concept of a zero does not seem to generalize to nonlinear system, we have discussed how the geometric version of a zero does. The usefullness of this concept has already been demonstrated in nonlinear decoupling and tracking [4] and is the study of controllability and observability of cascaded nonlinear systems [5]

[1] MacFarlane, A.G.J. and Karcanias, N. Poles and zeroes of linear multivariable system: a survey of the algebraic, geometric and complex variable theory. Int. J. Con. 24 (1976), pp 33-74

[2] Desoer, C.A. and Schulman, J.D. Zeroes and poles of matrix transfer function and their dynamical interpretation. IEEE Trans. Cir. Sys. 21(1974), pp 3-8

[3] Antsaklis, P.J. Maximal order reduction and supremal (A,B) - invariant and controllability subspace. IEEE Trans. Auto. Con. 25(1980) pp 44-49

[4] Isidori, A., Krener, A. J., Gori-Giorgi, C., and Monaco, S. Nonlinear decoupling via feedback: a differential geometric approach. IEEE Trans. Auto. Con. to appear.

[5] Isidori, A., Gori-Giorgi, C. and Monaco, S. The observability of cascade connected nonlinear systems, Proceedings IFAC/81 Congress, to appear