Controllability and Observability Conditions of Linear Autonomous Systems

Article · January 1969

1 author:

Malo Hautus
Eindhoven University of Technology

KONINKL. NEDERL. AKADEMIE VAN WETENSCHAPPEN – AMSTERDAM
Reprinted from Proceedings, Series A, 72, No. 5 and Indag. Math., 31, No. 5, 1969

MATHEMATICS

CONTROLLABILITY AND OBSERVABILITY CONDITIONS OF LINEAR AUTONOMOUS SYSTEMS

BY

M. L. J. HAUTUS

(Communicated by Prof. N. G. DE BRUIJN at the meeting of June 28, 1969)

§ 1 *Introduction*

Consider the linear control system

$$(\mathscr{D}) \qquad \dot{x}(t) = Ax(t) + Bu(t) \qquad (t \geqslant 0)$$

where A is an $n \times n$-matrix, B an $n \times m$-matrix, and where the admissible controls are functions $u: [0, \infty) \to R^m$ which are bounded and measurable. The system $(\mathscr{D})$ is called *controllable* (or *completely controllable*) if for each $a, b \in R^n$ there exists an admissible control u and a time $T > 0$, such that for the corresponding solution of (1), with $x(0) = a$ we have $x(T) = b$.

Furthermore, consider the iteration control system

$$(\mathscr{I}) \qquad x_{k+1} = Ax_k + Bu_k \qquad (k = 0, 1, 2, \ldots).$$

Here an admissible control is a sequence $\{u_0, u_1, \ldots\}$ of m-vectors. The system $(\mathscr{I})$ is called *controllable* if for every $a, b \in R^n$ there exists a control $\{u_0, u_1, \ldots\}$ and a number N, such that for the corresponding solution of $(\mathscr{I})$ with $x_0 = a$ we have $x_N = b$.

A necessary and sufficient condition of controllability (due to R. E. Kalman) for $(\mathscr{D})$ is: "The $n \times nm$-matrix $[B, AB, \ldots, A^{n-1}B]$ has rank n". (See [1] p. 81, [2] p. 201, [4] p. 170.) This statement is also true when $(\mathscr{D})$ is replaced by $(\mathscr{I})$. If this condition is satisfied we will call the pair (A, B) *controllable*.

In section 2 of this paper we will give another necessary and sufficient condition for the controllability of (A, B) which is in some cases more tractable. In fact, we will prove the following result:

Theorem 1. (A, B) is controllable if and only if the $n \times (n+m)$-matrix $[A - \lambda I, B]$ has rank n for every eigenvalue λ of A.

If A has Jordan form then theorem 1 is a generalization of a condition given in [2], Theorem 11. This will be shown in theorem 4 of section 2. However, the advantage of theorem 1 is that we can avoid Jordan form manipulation in most cases by applying theorem 1 directly. In section 2 we will also give a number of applications of theorem 1, some of which have been established before, but in a more complicated way.

Now we consider the observed systems

$$(\mathscr{D}_0) \qquad \dot{x} = Ax + Bu, \qquad y = Hx,$$

$$(\mathscr{I}_0) \qquad x_{k+1} = Ax_k + Bu_k, \qquad y_k = Hx_k.$$

Here y is a r-dimensional vector and H is a $r \times n$-matrix. The system $(\mathscr{D}_0)$ is called *(completely) observable* if for each admissible control u and each pair of solutions, x, $\hat{x}$, corresponding to the control u, we have: If $y(t) = Hx(t)$, $\hat{y}(t) = H\hat{x}(t)$, then $y(t) = \hat{y}(t)$ $(t \geqslant 0)$ implies $x(t) = \hat{x}(t)$ $(t \geqslant 0)$. Analogously, system $(\mathscr{I}_0)$ is observable if for each control $u = \{u_0, u_1, \ldots\}$ and each pair of solutions x_k, $\hat{x}_k$, corresponding to u, we have: If $y_k = Hx_k$ and $\hat{y}_k = H\hat{x}_k$, then $y_k = \hat{y}_k$ $(k = 0, 1, 2, \ldots)$ implies $x_k = \hat{x}_k$ $(k = 0, 1, 2, \ldots)$. Then it is well known that $(\mathscr{D}_0)$ (and also $(\mathscr{I}_0)$) is observable if and only if the pair (A', H') is controllable ([4] p. 170 Lemma 4, [1] p. 111, 112). Here the prime denotes transposition. We will call the pair (A, H) observable if (A', H') is controllable. Hence, theorem 1 yields also an observability condition.

§ 2 *Proof of theorem* 1 *and applications*

Theorem 1 can be stated in an equivalent way as follows:

Theorem 1'. (A, B) is controllable if and only if for each eigenvalue λ of A and for each (possibly complex) n-dimensional row vector η we have

$$\eta A = \lambda \eta, \ \eta B = 0 \Rightarrow \eta = 0,$$

(i.e., if there is no left eigenvector of A which is a left zero vector of B).

Proof. Suppose there exists $\eta \neq 0$ such that $\eta A = \lambda \eta$, $\eta B = 0$. Then we have $\eta A^k B = \lambda^k \eta B = 0$ $(k = 0, 1, \ldots, n-1)$ and hence rank $[B, AB, \ldots, A^{n-1}B] < n$.

On the other hand, suppose that rank $[B, AB, \ldots, A^{n-1}B] < n$. Then there exists $\zeta \neq 0$ such that

$$(1) \qquad \zeta A^k B = 0 \qquad (k = 0, 1, \ldots, n-1).$$

Let ψ be a polynomial of minimal degree such that $\zeta\psi(A) = 0$. It is clear that such a polynomial exists and has degree d with $1 \leqslant d \leqslant n$. For some λ and some polynomial φ of degree $d-1$ we have $\psi(z) = \varphi(z)(z - \lambda)$. Then we define $\eta = \zeta\varphi(A)$. By definition of ψ we have $\eta \neq 0$, $\eta(A - \lambda I) = 0$. Furthermore, it follows from (1) that $\eta B = 0$, and this completes the proof.

Corollary. The pair (A, H) is observable if and only if

$$(2) \qquad Ax = \lambda x, \quad Hx = 0 \Rightarrow x = 0.$$

Example 1. If (A, B) is controllable, then for every $m \times n$-matrix D we have: $(A + BD, B)$ is controllable. In fact, if $\eta(A + BD) = \lambda\eta$, $\eta B = 0$, then $\eta A = \lambda \eta$, $\eta B = 0$ and therefore $\eta = 0$.

Example 2. If A and H are given by

$$A = \begin{pmatrix} 0 & 1 & 0 \dots 0 \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \\ -a_n & \dots & \dots & -a_1 \end{pmatrix}, \quad H = \begin{pmatrix} b_{11} \dots \dots b_{1n} \\ \vdots \quad \vdots \\ b_{m1} \quad b_{mn} \end{pmatrix}$$

(so A is a companion matrix), then (A, H) is observable if and only if the polynomials $D(z) = z^n + a_1 z^{n-1} + \ldots + a_n$, $N_k(z) = b_{k1} + b_{k2}z + \ldots + b_{kn}z^{n-1}$ $(k = 1, \ldots, m)$ have no common zero. This follows directly from (2) and the fact that the eigenvectors of A are of the form $(1, \lambda, \ldots, \lambda^{n-1})'$, where λ is a zero of $D(z)$. (See also [1] p. 113–116, [4] p. 177–178.)

Example 3. If A is a companion matrix (as in example 2) and B is the column $(0, \ldots, 0, 1)'$ then (A, B) is controllable. For, if we omit the first column of $[A - \lambda I, B]$ we obtain a square triangular matrix with unities on the diagonal.

Example 4. If (A, B) is controllable and $\eta_1, \ldots, \eta_\nu$ are independent left eigenvectors of A corresponding to the same eigenvalue, then the vectors $\eta_1 B, \ldots, \eta_\nu B$ are independent. In fact, if $\alpha_1\eta_1 B + \ldots + \alpha_\nu\eta_\nu B = 0$, then we have $\eta = \alpha_1\eta_1 + \ldots + \alpha_\nu\eta_\nu = 0$ since otherwise η is a left eigenvector of A.

We mention a few notations: If A is an $n \times n$-matrix, then the *spectrum* of A (i.e. the set of eigenvalues) is denoted by $\sigma(A)$. The *order* of an eigenvalue λ of A, denoted by $\omega(\lambda)$, is the number of independent eigenvectors corresponding to λ. Equivalently, $\omega(\lambda) = n - \text{rank } [A - \lambda I]$. Still another characterization of the order of λ is: $\omega(\lambda)$ is the number of Jordan blocks with eigenvalue λ in the Jordan canonical form of A. Furthermore, we define $\omega(A) = \max \{\omega(\lambda) | \lambda \in \sigma(A)\}$. It is easily seen from theorem 1 (and from example 4), that if (A, B) is controllable we have $m \geqslant \omega(A)$. The following theorem is in a sense a converse of this property.

Theorem 2. If (A, B) is controllable, then there exists an $m \times \omega(A)$-matrix C such that (A, BC) is controllable.

Proof. Let $\sigma(A) = \{\lambda_1, \ldots, \lambda_\nu\}$. For each k there are $n - \omega(\lambda_k)$ columns of $A - \lambda_k I$ and $\omega(\lambda_k)$ columns of B which form together a base in R^n. There exists an $m \times \omega(A)$-matrix C_k, the entries of which are zeros and ones and such that BC_k contains the afore-mentioned columns of B. Then we have rank $[A - \lambda_k I, BC_k] = n$. If we have constructed matrices $C_1, \ldots, C_\nu$ in this way, we try to find C as a linear combination of them: $C(\alpha) = \alpha_1 C_1 + \ldots + \alpha_\nu C_\nu$, where $\alpha = (\alpha_1, \ldots, \alpha_\nu)$ is a real ν-vector. For each k we have rank $[A - \lambda_k I, BC(\alpha)] = n$ for some α. So there is an $n \times n$-determinant in this matrix, $D_k(\alpha)$ say, which is not identically zero. Since $D_k(\alpha)$ is a polynomial it vanishes only in a closed nowhere dense set in R^ν. Therefore, there is an $\bar{a} \in R^{\nu}$, such that $D_k(\bar{a}) \neq 0$ $(k=1, \ldots, \nu)$. Hence, rank $[A-\lambda_k I, BC(\bar{a})]=n$ $(k=1, \ldots, \nu)$.

Remark. For $\omega(A)=1$ this theorem was proved in [1] (p. 86–90). Actually, for this case a proof can be given which is somewhat simpler than the one given here, using the formulation of theorem 1′.

If A is an $n \times n$-matrix and $\varphi(z)$ is a complex function analytic in a neighborhood of $\sigma(A)$, then $\varphi(A)$ can be defined by the integral representation $\varphi(A)=(2\pi i)^{-1} \int\limits_C \varphi(z)(A-zI)^{-1}dz$, where C consists of one or more contours, the interiors of which are disjoint and include together $\sigma(A)$. Also, $\varphi(A)$ can be expressed by means of the Jordan canonical representation of A. Sometimes $\varphi(A)$ can be defined by a power series (e.g., $e^{tA}=\sum\limits_{n=0}^{\infty} t^n A^n/n!$). It is well known that $\sigma(\varphi(A))=\varphi(\sigma(A))$; i.e., the eigenvalues of $\varphi(A)$ are of the form $\varphi(\lambda)$ with $\lambda \in \sigma(A)$. Furthermore, eigenvectors of A are also eigenvectors of $\varphi(A)$. However, an eigenvector of $\varphi(A)$ is not necessarily an eigenvector of A (see [5] p. 99–101, [6] Ch. V).

Theorem 3. *If A is an $n \times n$-matrix and φ is a function analytic on $\sigma(A)$, which satisfies*

i) $\varphi'(\lambda) \neq 0 \qquad (\lambda \in \sigma(A))$,

ii) $\varphi(\lambda) \neq \varphi(\mu) \qquad (\lambda, \mu \in \sigma(A), \lambda \neq \mu)$,

then $(\varphi(A), B)$ is controllable if and only if (A, B) is controllable.

Proof. We show that left eigenvectors of $\varphi(A)$ are also left eigenvectors of A. Suppose that η is a left eigenvector of $\varphi(A)$. Then we have $\eta\varphi(A)=\varphi(\lambda)\eta$ for some $\lambda \in \sigma(A)$. Let ψ be defined by $\psi(z)=(\varphi(z)-\varphi(\lambda))/ /(z-\lambda)$ $(z \neq \lambda)$ and $\psi(\lambda)=\varphi'(\lambda)$. Then $\psi(z)$ is analytic on $\sigma(A)$ and by i) and ii) $\psi(z) \neq 0$ $(z \in \sigma(A))$, that is, $\psi(A)$ is non-singular. Furthermore, we have $\varphi(z)=\varphi(\lambda)+(z-\lambda)\psi(z)$, and hence $\eta\varphi(A)=\varphi(\lambda)\eta+\eta(A-\lambda I)\psi(A)$. This implies $\eta A=\lambda\eta$.

Remark. An alternate proof of theorem 3 can be obtained by observing that conditions i) and ii) imply that φ has an analytic inverse, say χ, on a neighborhood of $\sigma(A)$. Since the eigenvectors of $\varphi(A)$ are also eigenvectors of $\chi(\varphi(A))=A$ the result follows immediately.

Example 5. (Sampling). Consider the observable system $(\mathcal{D}_0)$. For a given control u we want to determine the initial state $x(0)$ from the observed function y by equidistant sampling, that is, from the knowledge of $y(t_k)$ $(k=1, \ldots, n)$ where $t_k=k\tau$ and τ is a given positive number. We have $x(t_k)=e^{k\tau A}x(0)+g(t_k)$, where $g(t)=\int\limits_0^t e^{(t-s)A}Bu(s)ds$ is a known function. Hence, $y(t_k)=He^{k\tau A}x(0)+Hg(t_k)$, from which $x(0)$ can be solved uniquely if rank $[H', e^{\tau A'}H', \ldots, e^{(n-1)\tau A'}H']=n$; i.e., if $(e^{\tau A}, H)$ is observable. According to theorem 3 a sufficient condition for this property is

$$\tau(\lambda-\mu) \not\equiv 0 \pmod{2\pi i} \qquad (\lambda, \mu \in \sigma(A), \lambda \neq \mu) \tag{3}$$

(see also [1], p. 383).

Example 6. Consider system ($\mathscr{D}$) and assume that it is controllable. Now suppose that u has to be constant on the time intervals $t_k \leqslant t < t_{k+1}$ ($k=0, 1, \ldots$) where $t_k = k\tau$ and $\tau > 0$. Then ($\mathscr{D}$) is actually a discrete system:

$$x(t_{k+1}) = e^{\tau A}x(t_k) + \gamma(A)Bu(t_k) \tag{$\mathscr{D}_s$}$$

where $\gamma(z) = \int_0^\tau e^{\sigma z} d\sigma$. The problem which arises is whether ($\mathscr{D}_s$) is controllable; i.e., whether we have not destroyed controllability by sampling. Since $\gamma(A)$ and $e^{\tau A}$ commute, it is easily seen that $(e^{\tau A}, \gamma(A)B)$ is controllable if and only if $\gamma(A)$ is non-singular and $(e^{\tau A}, B)$ is controllable. If A is a real matrix, then (3) is easily seen to be a sufficient condition for both of these properties. (For another proof, see [2] theorem 12.)

Now we are going to express theorem 1′ in terms of Jordan forms. If A is a Jordan canonical form, the left eigenvectors of A are easily determined. Suppose $A = \text{diag}\,(J_1, \ldots, J_p)$, where J_k is a Jordan block of dimension ν_k (so that we have $\sum \nu_k = n$). For $q = 1, \ldots, p$ we have that $\eta_q = (0, \ldots, 0, 1, 0, \ldots, 0)$ is a left eigenvector of A where the non-zero component of η_q is in the n_q-th place with $n_q = \nu_1 + \ldots + \nu_q$. If eigenvalues corresponding to different blocks are different, then these are the only left eigenvectors (except for scalar multiplication). If A has blocks with equal eigenvalues, then a (nontrivial) linear combination of the corresponding η_q's is also an eigenvector and each eigenvector can be obtained in this way. Now we can restate theorem 1′ for this case. Suppose that $\beta_1, \ldots, \beta_n$ are the rows of B.

Theorem 4. With the notation just given, (A, B) is controllable if and only if the rows $\beta_{\nu_1}, \beta_{\nu_1+\nu_2}, \ldots, \beta_n$ do not vanish and the ones which correspond to blocks with equal eigenvalues are independent (compare example 4).

For $m=1$ this criterion is given in [2] theorem 11.

We conclude with some results which can easily be derived from theorems 1′ and 3.

7) If A is a real $n \times n$-matrix, then (A, B) is controllable for every real $B \neq 0$ if and only if either $n=1$ or $n=2$ and A has complex eigenvalues (Th. 1). (See also [3].)
8) If (A, B) is controllable and A is a stability matrix, then (A^2, B) is controllable (Th. 3). (See [1], p. 420 Ex. 8.)

9) It is possible to sample even the discrete system ($\mathscr{I}$) (or ($\mathscr{I}_0$)) by controlling (or observing) only at each N^{th} step. In order that the system is still controllable (or observable) a sufficient condition is: A is non-singular and $\lambda^N \neq \mu^N$ ($\lambda, \mu \in \sigma(A), \lambda \neq \mu$).

Technological University,
Eindhoven

REFERENCES

1. Lee, E. B. and L. Markus, "Foundations of optimal control theory". Wiley (1967).
2. Kalman, R. E., Y. C. Ho and K. S. Narendra, "Controllability of linear dynamical systems". Contrib. Diff. Equations 1, No. 2, 189–213 (1963).
3. Chu, H., "A remark on complete controllability". J. Siam Control 3, 439–442 (1966).
4. Kalman, R. E., "Mathematical description of linear dynamical systems". J. Siam Control. 1, 152–192 (1963).
5. MacDuffee, C. C., "The Theory of Matrices". Chelsea (1956).
6. Gantmacher, F. R., "Matrizenrechnung I". V.E.B. Deutscher Verlag der Wissenschaften (1965).