A Theorem on the Relationship between Companion Form and Observability for Linear Time Invariant Systems

Nathaniel Kroeger, Ian Gravagne*
Department of Electrical Engineering
Baylor University
Nathaniel_Kroeger@Baylor.edu
Ian_Gravagne@Baylor.edu

John Davis
Department of Mathematics
Baylor University
John_M_Davis@Baylor.edu

Abstract

Similarity transformations of linear systems of differential equations are briefly discussed. The companion form of a matrix and of the system is introduced as well as the fundamental control concept of observability. A condition for linear systems that guarantees the existence and uniqueness of a linear transformation that transforms the original system into a companion system is stated. The question is then raised: if there is no way to transform a given system into its companion form, is there at least a way to transform the system matrix into a companion form? The concluding theorem proves that the answer is yes, and furthermore any transformation matrix that will do so must be a full rank observability matrix of the system. This result holds for all systems whose matrix's characteristic polynomial is equal to its minimal polynomial.

Introduction

State space methods for systems and controls are widely used in engineering applications. One of the most prominent ways the state space model is expressed is through the controllable canonical form. This form is particularly desirable for several reasons. As with any canonical form, it highlights certain features of the system so that they can be quickly deduced, and facilitates easy comparison of systems along those specific features. Furthermore, it allows for easy system realization. It also allows for easy determination of the system's transfer function and therefore its poles. Most importantly, the controllable canonical form is an ideal form to control the system with, namely for the ability to arbitrarily assign the system poles using only linear feedback.

*Address all correspondence to this author.

Unfortunately, not every linear time invariant (LTI) system expressed in state variable form can be transformed into controllable canonical form. However, it is possible to transform the system into a "quasi controllable" form by transforming the system matrix (While this result is implied in some instances, e.g. Terrell [1] and Antsaklis and Michel [3], the purpose of this work is to provide a direct constructive proof and example). Although this form may not share the exact same properties as the controllable canonical form, it is still useful for system comparison.

An $n \times n$ linear system of differential equations of the form

$$\mathbf{x}' = A\mathbf{x} + \mathbf{b}u(t), \tag{1}$$

can undergo a nonsingular linear transformation $\mathbf{z} = T\mathbf{x}$ to result in the *companion system*

$$\mathbf{z}' = P\mathbf{z} + \mathbf{e}_n u(t), \tag{2}$$

where $\mathbf{e}_n$ is the n-th standard basis vector $\begin{bmatrix} 0 & 0 & \dots & 1 \end{bmatrix}^T$ and P is the *companion matrix*, sometimes known as the *controllable canonical form* [2, p 793]

$$P = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & 1 \\ -k_0 & -k_1 & -k_2 & \dots & -k_{n-1} \end{bmatrix} \tag{3}$$

such that $P = TAT^{-1}$. The k_i are the coefficients of the characteristic *monic* polynomial of A: $p_A(\lambda) = \lambda^n + k_{n-1}\lambda^{n-1} + \ldots + k_1\lambda + k_0$. However, this transformation exists if and only if 1) the minimal polynomial of A is equal to its characteristic polynomial, and 2) that

$$rank \begin{bmatrix} \mathbf{b} & A\mathbf{b} & A^2\mathbf{b} & \dots & A^{n-1}\mathbf{b} \end{bmatrix} = n \tag{4}$$

has full rank. [1, Theorem 1, p 709]. Additionally, when these requirements are fulfilled, the matrix T that performs the linear transformation is unique. After finding the correct row vector τ, which is the solution to $\tau \begin{bmatrix} \mathbf{b} & A\mathbf{b} & A^2\mathbf{b} & \dots & A^{n-1}\mathbf{b} \end{bmatrix} = \mathbf{e}_n{}^T$ [1, pp.708-709], T has the form

$$T = \begin{bmatrix} \tau \\ \tau A \\ \tau A^2 \\ \vdots \\ \tau A^{n-1} \end{bmatrix} \tag{5}$$

Additionally, if (1) is equipped with a measurement capability

$$y = \mathbf{c}^T \cdot \mathbf{x}, \tag{6}$$

an *observability matrix*, B, of the system can be defined to be

$$B = \begin{bmatrix} \mathbf{c}^T \\ \mathbf{c}^T A \\ \mathbf{c}^T A^2 \\ \vdots \\ \mathbf{c}^T A^{n-1} \end{bmatrix} \tag{7}$$

The system is said to be *observable* if and only if rank$B = n$. Observability is desirable because it allows the reconstruction of the original state ($\mathbf{x}_0$) solely from knowledge of y, $u(t)$, and their derivatives.

Observability and the Companion Matrix

It is clear that if (4) does not hold for any given system of the form (1), then it is impossible to linearly transform that system to the form (2). We then ask, if it is impossible to do so, is there at least a linear transformation C that will take A to P such that $P = CAC^{-1}$, and is it easy to find?

Since a similarity transformation is sought, the minimal polynomial of A and P must match, as well as their characteristic polynomials. It can be shown that for all matrices of the form P, the minimal polynomial is equal to its characteristic polynomial. [1, Proposition 1, p 707] Therefore, it is only possible to find C when the minimal polynomial and characteristic polynomial of A are equal. Fortunately, this is not too strict of a condition.

Example 1 Consider two linear systems: the first $\mathbf{x}' = A\mathbf{x} + \mathbf{b}u(t)$ and the second $\mathbf{x}' = A\mathbf{x} + \mathbf{f}v(t)$. Suppose rank $\begin{bmatrix}\mathbf{b} & A\mathbf{b} & A^2\mathbf{b} & \dots & A^{n-1}\mathbf{b}\end{bmatrix} = n$ and rank $\begin{bmatrix}\mathbf{f} & A\mathbf{f} & A^2\mathbf{f} & \dots & A^{n-1}\mathbf{f}\end{bmatrix} < n$. Then a transformation T can be constructed from a unique τ to bring system 1 into companion form with $\mathbf{z} = T\mathbf{x}$, and recall that $P = TAT^{-1}$ and $T\mathbf{b} = \mathbf{e}_n$:

$$\begin{aligned}
\mathbf{z}' &= T\mathbf{x}' && && (8.1)\\
&= T(A\mathbf{x} + \mathbf{b}u(t)) && (\mathbf{x}' = A\mathbf{x} + \mathbf{b}u(t)) && (8.2)\\
&= TAT^{-1}\mathbf{z} + T\mathbf{b}u(t) && (\mathbf{x} = T^{-1}\mathbf{z}) && (8.3)\\
&= P\mathbf{z} + \mathbf{e}_n u(t) && && (8.4)
\end{aligned}$$

No such transformation can be found for system 2. However, if the same T that was used for system 1 is used for system 2, then the process is the same:

$$\mathbf{z}' = T\mathbf{x}' \tag{9.1}$$
$$= T(A\mathbf{x} + \mathbf{f}v(t)) \tag{9.2}$$
$$= TAT^{-1}\mathbf{z} + T\mathbf{f}v(t) \tag{9.3}$$
$$= P\mathbf{z} + T\mathbf{f}v(t) \tag{9.4}$$

where $T\mathbf{f}$ is some arbitrary vector not equal to $\mathbf{e}_n$.

Notice that even though system 2 did not satisfy (4), T still transformed A into P. What makes system 2 fail to transform into a companion system is not A, but rather $\mathbf{f}$.

Example 1 demonstrated that for a system of the form (1) that fails (4), a transformation to a companion matrix is still available if another vector $\mathbf{d}$ can be found such that (4) holds using $\mathbf{d}$ instead of $\mathbf{b}$. In fact, the existence of such a vector $\mathbf{d}$ that satisfies (4) is both a necessary and sufficient condition for transforming A to P (3, pp 131-132). However, it could be difficult to track down such a $\mathbf{d}$, if it even exists. Instead, notice that when (4) is satisfied, the resulting matrix from (5) is merely a special case of the observability matrix (7), where $\tau = \mathbf{c}^T$. As it turns out, surprisingly *any* full-rank observability matrix is enough to transform A into P, provided A is similar to some P to begin with.

Theorem 1. *Given the linear system of the form $\boldsymbol{x}' = A\boldsymbol{x} + \boldsymbol{b}u(t)$, if the characteristic polynomial of A is equal to its minimal polynomial, then every full-rank observability matrix B of the system is a linear transformation that will take A into P such that $P = BAB^{-1}$, and every linear transformation that does so is a full-rank observability matrix.*

Proof. ($\Rightarrow$) If A's minimal and characteristic polynomials are equal to each other, then there exists some P similar to A. Since B is assumed to be full rank, it is invertible and B^{-1} exists. Then it will be enough to show that $BA = PB$.

Let r_i be the ith standard row vector $\begin{bmatrix}0 & 0 & \dots & 1 & \dots & 0\end{bmatrix}$, where the 1 is in the ith position, $i = 1, 2, \dots, n$. The ith row of B, denoted by b_i, can be written as $\begin{bmatrix}\tau A^{i-1}\end{bmatrix}$, $i = 1, 2, \dots, n$. Notice that $b_i A = b_{i+1}$ for $i = 1, 2, \dots, (n-1)$.

Then, for $i = 1, 2, \dots, (n-1)$,

$$r_i BA = b_i A \qquad\qquad r_i PB = r_{i+1} B \tag{10.1}$$
$$= b_{i+1} \qquad\qquad = b_{i+1} \tag{10.2}$$

This shows that the first $n-1$ rows of BA and PB are equal. All that is left to show is that the bottom rows are the same. To do this, recall the Cayley-Hamilton

Theorem: $A^n + a_{n-1}A^{n-1} + \ldots + a_1 A + a_0 I = 0$, where the a_i are the coefficients of the monic characteristic polynomial of A. Then,

$$A^n = \sum_{i=0}^{n-1} -a_i A^i \tag{11}$$

Proceed to front multiply BA and PB by r_n:

$$\begin{aligned} r_n BA &= (r_n B)A && r_n PB = (r_n P)B && (12.1)\\ &= (\tau A^{n-1})A && = \begin{bmatrix} -k_0 & -k_1 & \ldots & -k_{n-1} \end{bmatrix} B && (12.2)\\ &= \tau A^n && = (-k_0 r_1 + -k_1 r_2 + \ldots + -k_{n-1} r_n)B && (12.3)\\ &= \tau \sum_{i=0}^{n-1} -a_i A^i && = -k_0(r_1 B) + -k_1(r_2 B) + \ldots + -k_{n-1}(r_n B) && (12.4)\\ &= -a_0 \tau I + -a_1 \tau A + \ldots + -a_{n-1} \tau A^{n-1} && = -k_0 \tau I + -k_1 \tau A + \ldots + -k_{n-1} \tau A^{n-1} && (12.5) \end{aligned}$$

The first half of the proof is complete once it is shown that $a_i = k_i$ for $i = 0, 1, \ldots, (n-1)$. However, this is guaranteed because A and P are assumed to be similar, and so their monic characteristic polynomials must be equal, so every individual coefficient must also be equal.
($\Leftarrow$) Suppose a transformation D existed that did not take the form of B, so that $P = DAD^{-1}$, so $DA = PD$. Let d_i denote the ith row of D. Without loss of generality, suppose that $d_2 \neq d_1 A$. Then the first rows of DA and PD would not be the equal because

$$\begin{aligned} r_1 DA &= (r_1 D)A && r_1 PD = r_2 D && (13.1)\\ &= d_1 A && = d_2 && (13.2) \end{aligned}$$

However, it was assumed that $d_2 \neq d_{row1} A$, which is required for $DA = PD$, a contradiction.

$\square$

Example 2 Consider the linear system $\mathbf{x}' = A\mathbf{x} + \mathbf{b}u(t)$, with

$$A := \begin{bmatrix} 35 & 28 & -73 & -33 & 14 \\ 52 & 35 & -95 & -49 & 18 \\ 33 & 23 & -62 & -31 & 12 \\ 17 & 15 & -38 & -16 & 7 \\ 21 & 15 & -40 & -19 & 8 \end{bmatrix} \qquad \text{and} \qquad \mathbf{b} := \begin{bmatrix} -1 \\ -3 \\ -1 \\ -1 \\ 1 \end{bmatrix} \tag{14}$$

Condition (4) does not hold, as

$$\begin{bmatrix} \mathbf{b} & A\mathbf{b} & A^2\mathbf{b} & A^3\mathbf{b} & A^4\mathbf{b} \end{bmatrix} = \begin{bmatrix} -1 & 1 & 3 & 1 & -1 \\ -3 & 5 & 9 & 1 & -3 \\ -1 & 3 & 5 & 1 & -1 \\ -1 & -1 & 1 & 1 & -1 \\ 1 & 1 & 3 & 3 & 1 \end{bmatrix} \tag{15}$$

The first and last columns of the matrix are the same, so the above matrix cannot have full rank. Therefore, the usual method to transform A into its canonical form will not work. One way to find a transformation is to take a detour through the Jordan form. We need to find the Jordan decomposition of A, where $A = VJV^{-1}$. After some computation,

$$J := \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \qquad \text{and} \qquad V := \begin{bmatrix} 1 & 1 & 3 & 1 & 1 \\ 2 & 4 & 1 & 3 & 0 \\ 1 & 2 & 1 & 2 & 0 \\ 1 & 0 & 2 & 0 & 1 \\ 1 & 0 & 0 & 2 & 0 \end{bmatrix} \tag{16}$$

This calculation did several things. Because A is similar to J, their characteristic polynomials, minimal polynomials, and eigenvalues are the same. Looking at J, it is clear that the characteristic polynomial for both is:

$$(\lambda + i)(\lambda - i)(\lambda + 2)(\lambda - 1)^2 = \lambda^5 - 2\lambda^3 + 2\lambda^2 - 3\lambda + 2 \tag{17}$$

Further, there is only one Jordan block for each distinct eigenvalue, so the characteristic polynomial must be equal to the minimal polynomial. Therefore, A must be similar to a canonical control matrix P (before this observation, it was still unknown as to whether A was even similar to such a P). Specifically, A is similar to P, where

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ -2 & 3 & -2 & 2 & 0 \end{bmatrix} \tag{18}$$

The bottom row entries were found by looking at the characteristic polynomial of A, as discussed in the beginning of the paper.

Because P is similar to A, it must also be similar to J, so there must be a matrix S such that $P = SJS^{-1}$. S can be found by performing the Jordan decomposition of P, and

$$S = \begin{bmatrix} 1 & 0 & 1 & 1 & -4 \\ 0 & 1 & -2 & 1 & -3 \\ -1 & 0 & 4 & 1 & -2 \\ 0 & -1 & -8 & 1 & -1 \\ 1 & 0 & 16 & 1 & 0 \end{bmatrix} \tag{19}$$

Finally, we can now chain the similarity relationships together to obtain the desired result: $P = SJS^{-1} = S(V^{-1}AV)S^{-1} = (SV^{-1})A(VS^{-1}) = (SV^{-1})A(SV^{-1})^{-1}$ The transformation matrix T that brings A into P is $T = SV^{-1}$

$$T = \begin{bmatrix} -36 & -27 & 72 & 32 & -13 \\ -17 & -12 & 33 & 14 & -6 \\ -18 & -17 & 43 & 16 & -8 \\ 9 & 10 & -25 & -10 & 6 \\ -34 & -33 & 83 & 34 & -16 \end{bmatrix} \tag{20}$$

This T was obtained unconventionally, as the standard method for finding T did not apply. Nevertheless, the theorem presented still holds: T is a full-rank observability matrix. Choosing $\mathbf{c}^T := \begin{bmatrix} -36 & -27 & 72 & 32 & -13 \end{bmatrix}$ (the first row of T), the matrix described in (7),

$$B = \begin{bmatrix} \mathbf{c}^T \\ \mathbf{c}^T A \\ \mathbf{c}^T A^2 \\ \vdots \\ \mathbf{c}^T A^{n-1} \end{bmatrix}$$

results in the *exact same* matrix as T!

As a consequence, using the output

$$y = \mathbf{c}^T \cdot \mathbf{x}$$

makes the system completely observable.

Conclusion

Linear systems of differential equations are briefly reviewed. A condition for transforming a linear system into its control canonical form is stated, and the observability

matrix is introduced. Although not every linear system can be transformed into its control canonical form, the system matrix can be transformed into its canonical form provided that its characteristic polynomial is equal to its minimal polynomial. The concluding theorem proves that any full rank observability matrix will transform the system matrix into its canonical form and that all such transformation matrices are full rank observability matrices.

References

---

1. William J. Terrell. "Some Fundamental Control Theory I: Controllability, Observability, and Duality". *The American Mathematical Monthly*, vol 106, 1999, pp 705-719.

2. Lathi, B.P. *Signal Processing and Linear Systems.* New York: Oxford University Press, 1998. Print.

3. Antsaklis, Panos J., and Anthony N. Michel. *Linear Systems.* New York: McGraw-Hill, 1997. Print.