Smith-McMillan Form for Multivariable Systems

A. Overview

The Smith-McMillan form is used to determine the poles and zeros of the transfer matrix of a system with multiple inputs and/or outputs[1]. The transfer matrix is a matrix of transfer functions between the various inputs and outputs of the system. The poles and zeros that are of interest are the poles and zeros of the transfer matrix itself, not the poles and zeros of the individual elements of the matrix. The locations of the poles of the transfer matrix are available by inspection of the individual transfer functions, but the total number of the poles and their multiplicity is not. The location of system zeros, or even their existence, is not available by looking at the individual elements of the transfer matrix. There are also multiple definitions for system zeros of multivariable systems, which further complicates the issue.

The transfer matrix will be denoted by $G(s)$. The number of rows in $G(s)$ is equal to the number of system outputs; that will be denoted by m. The number of columns in $G(s)$ is equal to the number of system inputs; that will be denoted by p. Thus, $G(s)$ is an $m \times p$ matrix of transfer functions. The normal rank of $G(s)$ is r, with $r \leq \min[p, m]$.

B. The Smith Form

As a first step toward developing the Smith-McMillan form for $G(s)$, we will write the transfer matrix in the following form, which will lead us first of all to the Smith form.

$$G(s) = \frac{1}{d(s)} \cdot P(s) \tag{1}$$

where $d(s)$ is a polynomial that is the *least common multiple* (lcm) of the denominators of all the elements in $G(s)$, and $P(s)$ is a polynomial matrix, that is, a matrix that has a polynomial for each of its entries. $P(s)$ and $G(s)$ have the same dimensions. Thus, the matrix of transfer functions is written as a matrix of polynomials divided by a common denominator polynomial. Formally, $P(s) \in M(R[s])$, where $M(\cdot)$ denotes a matrix and $R[s]$ is the ring of polynomials[2].

Elementary row and column operations are used to transform $P(s)$ into its Smith form. From there it is an easy step to the Smith-McMillan form where the poles and zeros of $G(s)$ can be identified. The three elementary operations for a polynomial matrix are:

- multiplying a row or column by a constant;
- interchanging two rows or two columns; and
- adding a polynomial multiple of a row or column to another row or column.

These operations are carried out on a transfer matrix $G(s)$ by either pre-multiplication or post-multiplication by unimodular polynomial matrices known as elementary matrices. A polynomial matrix is unimodular if its inverse also is a polynomial matrix. A necessary and sufficient condition for this is that the determinant of the polynomial matrix be a constant, that is, a polynomial of zero degree. A unimodular polynomial matrix is an element of $U(R[s])$, the set of units in the ring of polynomial matrices. Pre-multiplication of $G(s)$ by an elementary matrix produces the corresponding row operation, while post-multiplication produces a column operation.

Two polynomial matrices $P(s)$ and $S(s)$ are said to be *equivalent*, denoted by $P(s) \sim S(s)$, if there exists a set of elementary matrices L_i and R_i such that

$$S(s) = L_1(s)L_2(s) \cdots L_{n1}(s)P(s)R_1(s)R_2(s) \cdots R_{n2}(s) \tag{2}$$

By a suitable choice of L_i and R_i matrices, the polynomial matrix $P(s)$ can be transformed into a pseudo-diagonal $m \times p$ matrix $S(s)$, the Smith form, that has the following structure.

$$S(s) = diag\left[\epsilon_1^{'}(s)\ ,\ \epsilon_2^{'}(s)\ ,\ \cdots\ ,\ \epsilon_r^{'}(s)\ ,\ 0\ ,\ 0\ ,\ \cdots\ ,\ 0\right] \tag{3}$$

where r is the normal rank of $G(s)$. The $\epsilon_i'(s)$ in (3) are monic polynomials that are known as the invariant factors of $P(s)$. They have the property that

$$\epsilon_i'(s) \mid \epsilon_{i+1}'(s) \qquad i \in [1,\ r-1] \tag{4}$$

which means that polynomial $\epsilon_i'(s)$ exactly divides polynomial $\epsilon_{i+1}'(s)$. Therefore, each factor of $\epsilon_i'(s)$ also appears as a factor in $\epsilon_{i+1}'(s)$. These polynomials are defined by

[1] *Multivariable Feedback Design*, J.M. Maciejowski, Addison-Wesley, Reading, MA, 1989, Chap. 2.
[2] See Appendix.

$$\epsilon_i'(s) = \frac{D_i(s)}{D_{i-1}(s)} \ , \qquad D_0(s) \triangleq 1 \tag{5}$$

where $D_i(s)$ is the *greatest common divisor* (gcd) of all $i \times i$ minors of $P(s)$. If $x(s)$ and $y(s)$ are polynomials, the greatest common divisor of $x(s)$ and $y(s)$ is the polynomial $w(s)$ that has the properties

$$\begin{aligned} & w(s) \mid x(s) \text{ and } w(s) \mid y(s) \\ & \text{For any } w_1(s) \text{ such that } w_1(s) \mid x(s) \text{ and } w_1(s) \mid y(s) \ , \ \text{then } w_1(s) \mid w(s) \end{aligned} \tag{6}$$

Once the $D_i(s)$ polynomials are determined, and the $\epsilon_i'(s)$ polynomials are computed, it is a short step to the Smith-McMillan form of the transfer matrix.

C. Smith-McMillan Form

Just as the polynomial matrix $P(s)$ is equivalent to a pseudo-diagonal polynomial matrix $S(s)$, so also the transfer matrix $G(s)$ is equivalent to a pseudo-diagonal matrix of transfer functions, that is, $G(s) \sim M(s)$. The matrix $M(s)$ is known as the Smith-McMillan form of the transfer matrix $G(s)$. The matrix has the following structure:

$$M(s) = diag\left[\frac{\epsilon_1(s)}{\psi_1(s)} \ , \ \frac{\epsilon_2(s)}{\psi_2(s)} \ , \ \cdots \ , \ \frac{\epsilon_r(s)}{\psi_r(s)} \ , \ 0 \ , \ 0 \ , \ \cdots \ , \ 0\right] \tag{7}$$

where

$$\frac{\epsilon_i(s)}{\psi_i(s)} = \frac{\epsilon_i^{'}(s)}{d(s)} \tag{8}$$

with all possible cancellations between $\epsilon_i^{'}(s)$ and $d(s)$ being performed. The polynomials $\epsilon_i(s)$ and $\psi_i(s)$ are coprime and have the following characteristics.

$$\begin{aligned} \epsilon_i(s) \mid \epsilon_{i+1}(s) \qquad & i \ \in \ [1, \ r-1] \\ \psi_{i+1}(s) \mid \psi_i(s) \qquad & i \ \in \ [1, \ r-1] \end{aligned} \tag{9}$$

The poles and zeros of the transfer matrix $G(s)$ can be found from the elements of $M(s)$. The *pole polynomial* is defined as

$$p(s) = \prod_{i=1}^{r} \psi_i(s) = \psi_1(s) \ \psi_2(s) \ \cdots \ \psi_r(s) \tag{10}$$

The roots of $p(s) = 0$ are the poles of the transfer matrix $G(s)$. The total number of poles in the system is given by $\deg\left[p(s)\right]$, which is known as the *McMillan degree*. It is the dimension of a minimal state-space representation of $G(s)$. A state-space representation of $G(s)$ may be of higher order than the McMillan degree, indicating pole-zero cancellations in the system. Repeated poles can also be identified by inspection of $p(s)$. More is said about this in the next section.

In similar fashion, the *zero polynomial* is defined as

$$z(s) = \prod_{i=1}^{r} \epsilon_i(s) = \epsilon_1(s) \ \epsilon_2(s) \ \cdots \ \epsilon_r(s) \tag{11}$$

The roots of $z(s) = 0$ are known as the *transmission zeros* of $G(s)$. From (11), it can be seen that any transmission zero of the system must be a factor in at least one of the $\epsilon_i(s)$ polynomials. The normal rank of both $M(s)$ and $G(s)$ is r. It is clear from (7) that if any $\epsilon_i(s) = 0$, then the rank of $M(s)$ drops below r. Therefore, since the ranks of $M(s)$ and $G(s)$ are always equal, a transmission zero of the transfer matrix $G(s)$ is defined to be any value of s for which $G(s)$ loses rank. These transmission zeros are not in general the zeros of the individual transfer functions that are elements of $G(s)$.

A transfer matrix may or may not have any transmission zeros, even though individual transfer functions do have zeros. Let the state-space representation of $G(s)$ be given by $[A, B, C, D]$. If the system outputs contain direct information about each of the states, then there are no transmission zeros[3]. This would be the case if $C = I, \ D = 0$, for example. For square systems with $m = p$ inputs and outputs and n states, limits on the number of transmission zeros are:

[3] *Multivariable Feedback Control: Analysis and Design*, S. Skogestad and I. Postlethwaite, John Wiley & Sons, Chichester, UK, 1996, Chapter 4.

$$\begin{aligned} D &\neq 0: \quad \text{At most } n-m+\text{rank}(D) \text{ zeros} \\ D &= 0: \quad \text{At most } n-2m+\text{rank}(CB) \text{ zeros} \\ D &= 0 \text{ and rank } (CB)=m: \quad \text{Exactly } n-m \text{ zeros} \end{aligned} \tag{12}$$

D. Repeated Poles and Multiplicity of Poles

In the pole polynomial $p(s)$, there may be repeated factors such as $(s-p_1)^{n_1}$. This indicates that there are n_1 poles located at $s=p_1$. The number n_1 is known as the McMillan degree of the pole p_1. For single-input, single-output (SISO) systems, the number of repeated poles at a certain location is known as the multiplicity of the pole. Therefore, for a SISO system, a term such as $(s-p_1)^{n_1}$ indicates that the multiplicity of p_1 is n_1. However, for multivariable systems, there is not such as obvious relation between the number of the poles at a specific location and the multiplicity. This is due to the concept of input and output directions for multivariable systems.

Since $p(s)=\psi_1(s)\ \psi_2(s)\ \cdots\ \psi_r(s)$, a pole that is repeated in $p(s)$ may appear in more than one $\psi_i(s)$ polynomial. The repeated pole may be written as

$$(s-p_1)^{n_1} = (s-p_1)^{k_1}\ \ (s-p_1)^{k_2}\ \ \cdots\ \ (s-p_1)^{k_r}\ , \qquad \sum_{i=1}^{r} k_i = n_1 \tag{13}$$

The factor k_i is the multiplicity of the pole p_1 for the i^{th} input/output of the Smith-McMillan matrix $M(s)$. This multiplicity controls the partial fraction expansion for that input/output pair. For example,

$$\frac{1}{(s-p_1)^{k_1}} \Rightarrow a_1 e^{p_1 t} + a_2 t e^{p_1 t} + \cdots + a_{k_1} t^{(k_1-1)} e^{p_1 t} \tag{14}$$

When analyzing the strong stabilizability of a system with repeated poles, it is the total number of poles at a specific location, the McMillan degree n_1, that is used when applying the parity interlacing property[4].

The transfer matrix $G(s)$ is given in terms of its state-space representation by

$$G(s) = C\,(sI-A)^{-1}\,B + D = \frac{C\cdot \text{adj}\,(sI-A)\cdot B + D\cdot |sI-A|}{|sI-A|} \tag{15}$$

The dimension of the A matrix is the number of states in the system, denoted by n. The set of poles of $G(s)$, given by the factors of the pole polynomial $p(s)$ is a subset of the set of eigenvalues of the A matrix. Therefore, $n \geq \deg[p(s)]$. If $n > \deg[p(s)]$, then the system has lost either full state controllability or full state observability, or both. If $n=\deg[p(s)]$, then the system is both fully controllable and fully observable, and $[A,B,C,D]$ is a minimal realization of the system.

E. Example

To illustrate the computations involved in forming the Smith-McMillan matrix and the interpretation of those results, consider the following example of a system with $m=3$ outputs and $p=2$ inputs. The transfer matrix is shown below; the normal rank of $G(s)$ is $r=2$.

$$G(s) = \begin{bmatrix} \dfrac{1}{(s+1)\,(s+2)} & \dfrac{-1}{(s+1)\,(s+2)} \\ \dfrac{s^2+s-4}{(s+1)\,(s+2)} & \dfrac{2s^2-s-8}{(s+1)\,(s+2)} \\ \dfrac{s-2}{(s+1)} & \dfrac{2s-4}{(s+1)} \end{bmatrix} \tag{16}$$

- The first step in the procedure is to determine $d(s)$, the least common multiple of all the denominators in $G(s)$.

$$d(s) = \text{lcm}\left\{[(s+1)\,(s+2)]^4\,(s+1)^2\right\} = (s+1)\,(s+2) \tag{17}$$

- The next step is to determine the polynomial matrix $P(s)$ that will yield $G(s)$ when divided by $d(s)$. This matrix is

$$P(s) = \begin{bmatrix} 1 & -1 \\ s^2+s-4 & 2s^2-s-8 \\ (s-2)\,(s+2) & (2s-4)\,(s+2) \end{bmatrix} \tag{18}$$

[4] *A Tutorial on Robust Control: Design and Analysis*, G.O. Beale, Notes for ECE 720, Fall 2005, Chapter 5.

- The third step is to find the set of $D_i(s)$ polynomials that are the greatest common divisors of all the $i \times i$ minors of $P(s)$. For this $G(s)$, we will need to find $D_1(s)$ and $D_2(s)$. $D_0(s) = 1$.

$$D_1(s) = \gcd\left[1, -1, \left(s^2+s-4\right), \left(2s^2-s-8\right), \left(s^2-4\right), \left(2s^2-8\right)\right] = 1 \tag{19}$$

$$\begin{aligned} D_2(s) &= \gcd\left[\left|\begin{array}{cc} 1 & -1 \\ s^2+s-4 & 2s^2-s-8 \end{array}\right|, \left|\begin{array}{cc} 1 & -1 \\ \left(s^2-4\right) & \left(2s^2-8\right) \end{array}\right|, \left|\begin{array}{cc} s^2+s-4 & 2s^2-s-8 \\ \left(s^2-4\right) & \left(2s^2-8\right) \end{array}\right|\right] \\ &= \gcd\left[(s+2)(s-2),\ (s+2)(s-2),\ s(s+2)(s-2)\right] \\ &= (s+2)(s-2) \end{aligned} \tag{20}$$

- Now the $\epsilon_i'(s)$ polynomials can be found. They are

$$\begin{aligned} \epsilon_1'(s) &= \frac{D_1(s)}{D_0(s)} = \frac{1}{1} = 1 \\ \epsilon_2'(s) &= \frac{D_2(s)}{D_1(s)} = \frac{(s+2)(s-2)}{1} = (s+2)(s-2) \end{aligned} \tag{21}$$

The Smith form of $P(s)$ is

$$S(s) = \left[\begin{array}{cc} 1 & 0 \\ 0 & (s+2)(s-2) \\ 0 & 0 \end{array}\right] \tag{22}$$

- The Smith-McMillan form of the transfer matrix can now be formed. The entries in $M(s)$ will be

$$\begin{aligned} \frac{\epsilon_1(s)}{\psi_1(s)} &= \frac{\epsilon_1'(s)}{d(s)} = \frac{1}{(s+1)(s+2)} \\ \frac{\epsilon_2(s)}{\psi_2(s)} &= \frac{\epsilon_2'(s)}{d(s)} = \frac{(s+2)(s-2)}{(s+1)(s+2)} = \frac{(s-2)}{(s+1)} \end{aligned} \tag{23}$$

and the transfer matrix $G(s)$ is equivalent to

$$G(s) \sim M(s) = \left[\begin{array}{cc} \dfrac{1}{(s+1)(s+2)} & 0 \\ 0 & \dfrac{(s-2)}{(s+1)} \\ 0 & 0 \end{array}\right] \tag{24}$$

- The pole and zero polynomials can be identified by applying (10) and (11) to (24). The pole polynomial is

$$p(s) = [(s+1)(s+2)](s+1) = (s+1)^2(s+2) \tag{25}$$

so the McMillan degree for the system is 3. Two poles are located at $s=-1$, so the McMillan degree of that pole is 2. However, the multiplicity of the pole is only 1, since it appears in each of the two $\psi_i(s)$ polynomials raised to the first power. The other pole is located at $s=-2$.
The zero polynomial is given by

$$z(s) = s-2 \tag{26}$$

so there is a transmission zero located at $s=2$. For this system, the transmission zero does correspond in location to the zero that appears in two of the transfer functions in $G(s)$. There are four other zeros in the elements of $G(s)$ that are not transmission zeros. They are located at $s=-2.5616,\ 1.5616,\ 2.2656,\ -1.7656$. At $s=2$ the rank of $G(s)$ drops from 2 to 1 since both elements of the third row of $G(s)$ are zero, and the determinant formed from the first two rows is $3(s-2)/\left[(s+1)^2(s+2)\right]$.

F. Conclusions

To fully understand the characteristics of a multivariable system, it is necessary to know the locations and multiplicities of the poles and zeros. Since a multivariable system has input and output directions, the poles and zeros are not always apparent by inspection, particularly the zeros. The Smith-McMillan form provides a way to determine those locations and multiplicities of the poles and zeros. The stability of the system depends on the pole locations in the same way for both multivariable and SISO systems. In order to be bounded-input, bounded-out (BIBO) stable, all roots of $p(s) = 0$ must have strictly negative real parts. The presence of transmission zeros implies the blocking of certain input signals.

G. Appendix: Rings[5]

Definition 1: A ring is a nonempty set R closed under two binary operations, addition and multiplication, that satisfy the following axioms:

- Under the operation of addition, R is a commutative group, which means that

$$a + (b + c) = (a + b) + c, \ \forall \ a, b, c \in R \tag{27}$$
$$a + b = b + a, \ \forall \ a, b \in R \tag{28}$$

There exists an element $0 \in R$ such that

$$a + 0 = 0 + a = a, \ \forall \ a \in R \tag{29}$$

For every element $a \in R$ there exists a corresponding element $-a \in R$ such that

$$a + (-a) = (-a) + a = 0 \tag{30}$$

- Under the operation of multiplication, R is a semigroup, which means that

$$a \cdot (b \cdot c) = (a \cdot b) \cdot c, \ \forall \ a, b, c \in R \tag{31}$$

- Multiplication is distributive over addition, which means that

$$a \cdot (b + c) = a \cdot b + a \cdot c, \ \forall \ a, b, c \in R \tag{32}$$
$$(a + b) \cdot c = a \cdot c + b \cdot c, \ \forall \ a, b, c \in R \tag{33}$$

The element $0 \in R$ is known as the additive identity of the ring, and the element $-a \in R$ is known as the additive inverse. Note that when an element a is added to its additive inverse $-a$, the result is the additive identity. Also note that there is no multiplicative identity or multiplicative inverse defined for the (basic) ring. ⧫

Definition 2: A ring R is said to be Abelian or commutative if $ab = ba, \ \forall \ a, b \in R$. A ring is said to have an identity if there exists an element $1 \in R$ such that $a \cdot 1 = 1 \cdot a = a, \ \forall \ a \in R$. Note that this identity is a multiplicative identity. In a ring with identity, an element $x \in R$ is said to be a *unit* of R if there exists an element $y \in R$ such that $x \cdot y = y \cdot x = 1$. The element y is known as the (multiplicative) inverse of x and is denoted by x^{-1}. As for addition, note that multiplication of an element by its inverse yields the multiplicative identity. A ring must be with identity in order for there to exist a multiplicative inverse. The multiplicative identity of the ring (1) and its additive inverse (-1) are units; there may or may not be other units in the ring. ⧫

Commutativity and multiplicative identity are independent characteristics. A ring can have either one or the other, neither, or both. Note that multiplication of a unit by its inverse in a ring with identity does commute. This does not imply that commutativity holds for multiplying other elements in the ring. For example, the ring of all $n \times n$ matrices does have a multiplicative identity I_n, and multiplicative inverses exist for some elements, but the order of multiplication of elements in the ring does not commute. A ring that is both commutative and with identity is known as a commutative ring with identity.

[5] *Control System Synthesis: A Factorization Approach*, M. Vidyasagar, MIT Press, Cambridge, MA, 1987, Appendix A