

A useful statement of some Schur-Cohn stability criteria for higher order discrete dynamic systems

Raaj K. Sah

IEEE Transactions on Automatic Control, Volume 36, Number 8, August 1991, pages 988-989.

ABSTRACT

The classical Schur-Cohn criteria are an important source of stability conditions for discrete dynamic systems. However, the conventional statement of these criteria is opaque and not of much direct use for an applied analysis of higher order dynamic systems. The contribution of this note is to show how some of these criteria can be stated as useful and easily tested restrictions on the gradient matrices of the original higher order system.

Previous working versions include:

“A useful and easily-tested statement of some Schur-Cohn stability criteria for higher-order discrete dynamic systems.” Yale University, Department of Economics. Economic Growth Center, Center discussion paper number 614. October 1989.

A Useful Statement of Some Schur-Cohn Stability Criteria for Higher Order Discrete Dynamic Systems

Raaj K. Sah

Abstract—The classical Schur-Cohn criteria are an important source of stability conditions for discrete dynamic systems. However, the conventional statement of these criteria is opaque and not of much direct use for an applied analysis of higher order dynamic systems. The contribution of this note is to show how some of these criteria can be stated as useful and easily tested restrictions on the gradient matrices of the original higher order system.

I. INTRODUCTION

Consider autonomous discrete dynamic systems represented by the following set of difference equations:

$$y(T) = f(y(T-1), \dots, y(T-L), \theta) \quad (1)$$

where L is the maximum order of an equation, $y(T)$ is a $(K \times 1)$ vector, f is a $(K \times 1)$ vector function, θ is a parameter, $K \geq 1$, $L \geq 1$, and T is time. For $l = 1, \dots, L$, define $f^l \equiv \partial y(T) / \partial y(T-l)$ as the $(K \times K)$ gradient matrices of (1) with respect to various lagged vectors of variables.

The stability analysis of the higher order system in (1) requires the standard step of transforming it into a first-order system, $z(T) = g(z(T-1))$, where $z(T)$ is a $(KL \times 1)$ vector, and g is a $(KL \times 1)$ vector function. A "companion" matrix to system (1), defined as $A \equiv \partial z(T) / \partial z(T-1)$, is

$$A = \begin{bmatrix} f^1 & f^2 & \dots & f^{L-1} & f^L \\ I_K & O_K & \dots & O_K & O_K \\ \vdots & & & & \\ O_K & O_K & \dots & I_K & O_K \end{bmatrix} \quad (2)$$

where I_K is an identity matrix of order $(K \times K)$, and O_K is a null matrix of order $(K \times K)$. Denote the characteristic polynomial of A by

$$p(\lambda) \equiv |\lambda I_{KL} - A| \equiv \lambda^{KL} + c_1 \lambda^{KL-1} + \dots + c_{KL}. \quad (3)$$

Let a steady state of (1) be denoted by the $(K \times 1)$ vector x ; thus, x satisfies the relationship

$$x = f(x, \dots, x, \theta). \quad (4)$$

Throughout this note, we will assume that A has been evaluated at the steady state in question. Let $r(A)$ denote the spectral radius of A ; that is, the maximum absolute value of the eigenvalues of A . Then, the property

$$r(A) < 1 \quad (5)$$

is important for the following well-known reasons: i) if system (1) is linear, then property (5) is necessary and sufficient for there to be a unique steady state that is asymptotically stable; ii) if system (1) is nonlinear, then (5) is a sufficient condition for the asymptotic stability of a steady state of (1); a necessary condition is $r(A) \leq 1$ [1, p. 38]. When (1) is nonlinear, an asymptotically stable steady state with $r(A) = 1$ can arise in principle (this is called the nonhyperbolic case); however, it is extremely rare in the sense explained in [2, p. 157] and [3, pp. 19-20].

The classical Schur-Cohn criteria provide a set of necessary and sufficient conditions for (5) to hold. Among these are the following two necessary criteria, which are of interest for the present note:

$$p(1) > 0 \text{ and } (-1)^{KL} p(-1) > 0. \quad (6)$$

Manuscript received April 21, 1989; revised October 27, 1989.
The author is with the University of Chicago, Chicago, IL 60637.
IEEE Log Number 9100296.

See [1, p. 27] for the other necessary criteria. These other criteria, along with those in (6), are together sufficient for (5) to hold.

It is apparent from (2) and (3) that it is nearly impossible to discern a transparent connection between the gradient matrices f' and the criteria in (6). Also, the criteria in (6) are tedious to verify.

The contribution of this note is as follows. The theorem in the next section shows that the criteria in (6) can be stated directly in terms of the gradient matrices f' of the original higher order system. This alternative statement is not only easier to test than (6), but it also permits us to see the restrictions that (6) imposes on the gradient matrices f' . Moreover, the theorem yields some useful information for the sensitivity study of the stable steady states of (1). These aspects, concerning the practical value of the theorem, are illustrated in Section III.

The proof of the theorem, as well as of the lemma on which it depends, is straightforward. However, based on an extensive literature search, I believe that these or similar results have not previously been reported.

II. RESULTS

Lemma:

$$|\lambda I_{KL} - A| = \lambda^{KL} |I_K - \sum_I \lambda^{-I} f^I|, \quad \text{for } \lambda \neq 0. \quad (7)$$

Proof: From (2), $|\lambda I_{KL} - A|$

$$= \begin{vmatrix} \lambda I_K - f^1 & \cdots & -f^{L-1} & -f^L \\ -I_K & \cdots & O_k & O_k \\ \vdots & & & \\ O_k & \cdots & \lambda I_K & O_k \\ O_k & \cdots & -I_K & \lambda I_K \end{vmatrix}. \quad (8)$$

In (8), multiply each of the last K columns by $1/\lambda$, and add these columns respectively to the preceding K columns. That is, for $k = 1$ to K : i) multiply column $(L-1)K + k$ by $1/\lambda$, and ii) add the resulting column to column $(L-2)K + k$. This yields $|\lambda I_{KL} - A|$

$$= \lambda^K \begin{vmatrix} \lambda I_K - f^1 & \cdots & -(f^{L-1} + \frac{1}{\lambda} f^L) & -\frac{1}{\lambda} f^L \\ -I_K & \cdots & O_k & O_k \\ \vdots & & & \\ O_k & \cdots & \lambda I_K & O_k \\ O_k & \cdots & O_k & I_K \end{vmatrix} \quad (9)$$

$$= \lambda^K \begin{vmatrix} \lambda I_K - f^1 & \cdots & -f^{L-2} & -(f^{L-1} + \frac{1}{\lambda} f^L) \\ -I_K & \cdots & O_k & O_k \\ \vdots & & & \\ O_k & \cdots & \lambda I_K & O_k \\ O_k & \cdots & -I_K & \lambda I_K \end{vmatrix} \quad (10)$$

(To obtain (10) from (9), expand the determinant in (9) along the last row. The last row and the last column drop out. Repeat this $K-1$ times.)

A repetition of all preceding steps, $L-1$ times, yields the desired result: $|\lambda I_{KL} - A| = \lambda^{K(L-1)} |\lambda I_K - \sum_I \lambda^{-I} f^I| = \lambda^{KL} |I_K - \sum_I \lambda^{-I} f^I|$. \square

Theorem: The criteria in (6) can be stated respectively as

$$\left| I_K - \sum_I f^I \right| > 0 \quad \text{and} \quad \left| I_K - \sum_I (-1)^I f^I \right| > 0. \quad (11)$$

Proof: Since $p(\lambda) \equiv |\lambda I_{KL} - A|$, (6) can be rewritten as

$$|I_{KL} - A| > 0 \quad \text{and} \quad (-1)^{KL} | -I_{KL} - A | > 0. \quad (12)$$

Next, substitute $\lambda = 1$ and -1 , respectively, into (7). The resulting expressions imply that (12) can be restated as (11).

III. REMARKS AND EXAMPLES

It is clear that though the conditions in (11) are the exact equivalents of the Schur-Cohn criteria in (6), the former are transparent and directly observable restrictions on the gradient matrices f' .

Two brief examples are presented below. The first example illustrates how the theorem presented above makes it easier to test the criteria in (6). The second example shows how this theorem yields useful information for the sensitivity analysis of the stable steady states of (1). In these examples, the elements of the vectors $y(T)$ and x are denoted, respectively, as $y(T) \equiv (y_1(T), \dots, y_K(T))$ and $x \equiv (x_1, \dots, x_K)$.

Example 1: Consider the following nonlinear system, with $K = L = 2$:

$$\begin{aligned} y_1(T) &= \{y_1(T-1)\}^2 + y_2(T-1) + y_2(T-2), \text{ and} \\ y_2(T) &= y_1(T-1) + \beta y_2(T-2) \end{aligned} \quad (13)$$

where β is a real constant. The gradient matrices of (13), evaluated at a steady state x , are:

$$f^1 = \begin{bmatrix} 2x_1 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad f^2 = \begin{bmatrix} 0 & 1 \\ 0 & \beta \end{bmatrix}. \quad (14)$$

It is easily ascertained that system (13) has only two steady states; they are $x = (0, 0)$, and $x = (-(1 + \beta)/(1 - \beta), -(1 + \beta)/(1 - \beta)^2)$.

Suppose we are interested in determining the values of β for which the steady state $x = (0, 0)$ satisfies the criteria in (6). A direct use of (6) would first require constructing a (4×4) matrix A , as defined in (2), and then solving for the values of β that satisfy the inequalities in (6). In contrast, the theorem presented earlier allows the following simpler calculation. From (14),

$$\begin{aligned} I_2 - \sum_I f^I &= \begin{bmatrix} 1 - 2x_1 & -2 \\ -1 & 1 - \beta \end{bmatrix} \text{ and} \\ I_2 - \sum_I (-1)^I f^I &= \begin{bmatrix} 1 + 2x_1 & 0 \\ 1 & 1 - \beta \end{bmatrix}. \end{aligned}$$

It then follows by inspection that the conditions in (11) are satisfied at $x = (0, 0)$ if and only if $\beta < -1$. It can be similarly verified that at $x = (-(1 + \beta)/(1 - \beta), -(1 + \beta)/(1 - \beta)^2)$, the corresponding restriction is $-1/3 > \beta > -1$.

Example 2: Consider the sensitivity analysis of an asymptotically stable steady state x with respect to a scalar parameter θ . Recalling the reasons noted in Section I, we assume here that (5) is satisfied at x . From (4), the effect of a small change in θ is described by $dx/d\theta = [I_K - F]^{-1}(\partial f/\partial\theta)$, where $F = \sum_I f^I$. This yields $dx/d\theta = D(\partial f/\partial\theta) / |I_K - F|$, where D denotes the adjoint matrix of $I_K - F$. From the first inequality in (11), therefore, $\text{sgn}\{dx/d\theta\} = \text{sgn}\{D(\partial f/\partial\theta)\}$. Thus, by the theorem presented earlier, we need not be concerned about the sign of $|I_K - F|$ when evaluating the signs of the vector $\frac{dx}{d\theta}$. This useful information is clearly unavailable from the criteria in (6).

ACKNOWLEDGMENT

The author wishes to thank the anonymous referee for many valuable suggestions, and A. Ascencios and N. Zellner for their assistance.

REFERENCES

- [1] J. P. LaSalle, "The stability and control of discrete processes," in *Applied Mathematical Sciences*, Vol. 62. New York: Springer-Verlag, 1986.
- [2] M. W. Hirsch and S. Smale, *Differential Equations, Dynamical Systems and Linear Algebra*. New York: Academic, 1974.
- [3] R. Seydel, *From Equilibrium to Chaos*. New York: Elsevier, 1988.