Quick Stability Checks and Matrix Norms[1]

By JOHN CONLISK

Consider the difference equation system $y_t=b+Ay_{t-1}$; where y_t is an $nX1$ variable vector, b is an $nX1$ constant vector and A is an nXn constant matrix. The fundamental necessary and sufficient stability condition for the system is $\mu(A)<1$, where $\mu(A)$ is the modulus of the characteristic root of A of largest modulus. That is, $\mu(A)$ is the *spectral radius* of A, so-called because all roots of A lie on or within a circle of radius $\mu(A)$ in the complex plane.

There exists a class of scalar real-valued functions of matrices called *matrix norms*; and any matrix norm $f(A)$ has the property $\mu(A)\leq f(A)$. Hence a sufficient stability condition for $y_t=b+Ay_{t-1}$ is $f(A)<1$, where $f(.)$ is *any* matrix norm function. The condition is also sufficient for existence and uniqueness of the equilibrium $y_t=(I-A)^{-1}b$. The advantage of demonstrating existence, uniqueness, and stability through a norm $f(A)$ rather than through the spectral radius $\mu(A)$ itself is that many norms are far easier to compute than $\mu(A)$. The disadvantage is that the sufficient condition $f(A)<1$ may be far stronger than necessary.

Though the economic literature has exploited leading special cases of the $f(A)<1$ condition, more emphasis on the general case seems appropriate. Section I reviews matrix norm theory, presents the basic $f(A)<1$ condition, and relates it to existing stability results in the economic literature. The following sections extend the basic result to cover other contexts—structural systems, distributed lags and periodic coefficients.

I. MATRIX NORMS AND STABILITY

We know the system $y_t=b+Ay_{t-1}$ has unique stable equilibrium if $A=0$. Intuitively, we would guess that $y_t=b+Ay_{t-1}$ will still be stable if the "magnitude" of A doesn't depart too far from zero. A matrix norm is a scalar measure of "magnitude" which makes this intuition precise. By definition (Lancaster [4] chapter 6):

> A *matrix norm* is a real-valued scalar function f of a square matrix such that f satisfies the following four axioms, where A and B are any nXn matrices. (i) $f(A)\geq 0$ with equality holding if and only if $A=0$; (ii) $f(cA)=|c|f(A)$ for any scalar c; (iii) $f(AB)\leq f(A)f(B)$; (iv) $f(A+B)\leq f(A)+f(B)$.

[1] This research was supported by National Science Foundation grant number GS-3201. I am indebted to Donald Bear for helpful comments.

The basic norm theorem for our purposes is the following. Throughout, $\mu(.)$ will denote "spectral radius of". Scalars and matrix elements may be complex, unless otherwise noted.

(1) $\mu(A) \leq f(A)$ where $f(.)$ is a matrix norm.

Since (1) is central to the paper, a proof will be reproduced; for the rest, existing proofs will just be cited. Following Lancaster [4, p. 201], let λ be a characteristic root of A of maximum modulus with characteristic vector x; and let A_x be the square matrix $A_x = (x, 0, \ldots, 0)$. Then $\mu(A) = |\lambda|$; $Ax = \lambda x$; $AA_x = \lambda A_x$; and, using axioms (i), (ii) and (iii),

$$\mu(A) = |\lambda| = f(\lambda A_x)/f(A_x) = f(AA_x)/f(A_x)$$
$$\leq f(A)f(A_x)/f(A_x) = f(A).$$

Result (1) immediately implies the basic stability result:

(2) A sufficient condition for the existence, uniqueness and stability of equilibrium of $y_t = b + Ay_{t-1}$ is $f(A) < 1$, where $f(.)$ is any matrix norm.

The usefulness of (2) for dynamic analysis depends on the existence of tractable norm functions. Some prominent examples from the literature follow. (See Lancaster [4] chapter 6). Here $A = [a_{ij}]$ is nXn.

Maximum element norm: $n \max_{i,j} |a_{ij}|$.
Holder norm: $\{\Sigma_i \Sigma_j |a_{ij}|^p\}^{1/p}$ for $1 \leq p \leq 2$.
Euclidean norm: $\{\Sigma_i \Sigma_j |a_{ij}|^2\}^{1/2}$ (Holder norm with $p=2$).
Element sum norm: $\Sigma_i \Sigma_j |a_{ij}|$ (Holder norm with $p=1$).
Column sum norm: $\max_j \Sigma_i |a_{ij}|$.
Row sum norm: $\max_i \Sigma_j |a_{ij}|$.

These are all far easier to compute than $\mu(A)$; hence the usefulness of the $f(A) < 1$ stability condition.

The following theorem (Lancaster [4] p. 198) indicates how to generate an infinite number of new matrix norms, given one norm.

(3) If $f(A)$ is a matrix norm, then $g(A) = f(P^{-1}AP)$ is a matrix norm for any non-singular matrix P.

Many practical applications of (3) involve a diagonal P, which keeps computations simple. Using a diagonal P amounts to changing the units of measure of y_t in $y_t = b + Ay_{t-1}$ (Fisher [3]). When P is not diagonal, (3) is interesting because it shows that norms need not be functions only of the absolute values of elements of A, as in all the preceding examples.

We can now interpret some prominent stability results in the economic literature as special cases of the $f(A) < 1$ stability condition (2). The column (row) sum stability condition was developed in the economic literature by Solow, Fisher [2, 3], and others; and it can be found in the mathematical literature (only slightly disguised) in some matrix texts (Varga [6, chapter 2] for example). The column (row) sum condition states that $y_t = b + Ay_{t-1}$ is stable if, for some diagonal P, the matrix $P^{-1}AP$ has all column (row) sums (of element absolute values) less than one. Thus, the column (row) sum stability condition is equivalent to the norm condition $f(A)<1$ when $f(A)=g(P^{-1}AP)$ and $g(.)$ is the column (row) sum norm. In the literature, this column (row) sum stability condition is usually stated in the context of a non-negative A, in which case the condition is necessary as well as sufficient (Fisher [3] or Varga [6, Theorem 2.2]).

The dominant diagonal stability condition, which is prominent in the economic literature (Bear [1] for example), is also a special case of the norm condition $f(A)<1$; since, as Fisher [3, p. 447] points out, the dominant diagonal condition and column (row) sum condition are equivalent.

An important procedure for generating matrix norms is to "induce" them from "vector norms". By definition:

> A *vector norm* is a real-valued, scalar function F of a vector such that F satisfies the following three axioms, where x and y are any $nX1$ vectors. (a) $F(x)\geq 0$ with equality holding if and only if $x=0$; (b) $F(cx)=|\,c\,|\,F(x)$ for any scalar c; (c) $F(x+y)\leq F(x)+F(y)$.

The usefulness of this for "inducing" matrix norms is indicated by the theorem (Lancaster [4], p. 208):

(4) For a square matrix A and vector norm $F(x)$, the function $f(A)=\max_x F(Ax)/F(x)$ is a matrix norm (said to be induced by F).

For example, the Euclidean vector norm $F(x)=(x^*x)^{1/2}$ induces the matrix norm $f(A)=[\max_x(x^*A^*Ax)/(x^*x)]^{1/2}$, which expression is known to equal $\mu(A^*A)^{1/2}$. (Here an asterisk indicates a conjugate transpose; so $A^*=A'$ for real A.) Thus, we have:

> *Spectral norm* (*of* A): $\mu(A^*A)^{1/2}$.

A sufficient condition from the economic literature for the existence, uniqueness and stability of equilibrium of $y_t=b+Ay_{t-1}$ (with A real) is that $I-A'A$ be positive definite (Samuelson [5] p. 438 for example). Since $f(A)=\mu(A'A)^{1/2}<1$ if and only if $I-A'A$ is positive definite, then this condition is another special case of the $f(A)<1$ condition.

When the matrix A partitions conveniently, the following result is useful. It was not found in the literature, but its proof is straightforward enough to be omitted.

(5) Let $f(.)$ be a matrix norm having the property $f(A)\leq f(B)$ whenever $0\leq A\leq B$; let $g(.)$ be any matrix norm; and let A partition into square blocks A_{ij}. Then $h(A)$ is a matrix norm, where

$$h(A)=f\begin{pmatrix} g(A_{11})\ldots & g(A_{1k}) \\ \vdots & \vdots \\ g(A_{k1})\ldots & g(A_{kk}) \end{pmatrix}.$$

II. Application to Structural Systems

Often it is more convenient to deal with a structural system $y_t = d + By_t + Cy_{t-1}$ than with the reduced form system $y_t = b + Ay_{t-1}$, where $b = (I-B)^{-1}d$ and $A = (I-B)^{-1}C$. The following result handles this case.

(6) A sufficient condition for the existence, uniqueness and stability of equilibrium of the system $y_t = d + By_t + Cy_{t-1}$ is $f(B) + f(C) < 1$, where $f(.)$ is any matrix norm such that $f(I) = 1$.

The condition $f(I) = 1$ is fairly restrictive. For example, it holds for the row sum, column sum and spectral norms; it does not hold for the maximum element and Holder norms, for which $f(I)$ may exceed one [$f(I) \geq 1$ always]. For the simpler norms, the condition $f(B) + f(C) < 1$ can be checked almost at a glance. For proof of (6), it suffices to show that $f(B) + f(C) < 1$ implies $f[(I-B)^{-1}C] < 1$. This can be done as follows, using the series inverse formula $(I-B)^{-1} = \sum_{r=0}^{\infty} B^r$ and axioms (iii) and (iv).

$$\begin{aligned} f[(I-B)^{-1}C] &\leq f(I + B + B^2 + \cdots)f(C) \\ &\leq [1 + f(B) + f(B)^2 + \cdots]f(C) \\ &= [1 - f(B)]^{-1}f(C) = 1 - [1 - f(B)]^{-1}[1 - f(B) - f(C)] < 1. \end{aligned}$$

III. Applications to Distributed Lags

The main result of this section is as follows.

A sufficient condition for the existence, uniqueness and stability of equilibrium of the system

(7) $$y_t = b + A_1 y_{t-1} + \cdots + A_k y_{t-k}$$

is $f(A_1) + \cdots + f(A_k) < 1$, where $f(.)$ is a matrix norm such that $f(I) = 1$.

Again, the stability condition is easily checked for the simpler norms. It is known that the distributed lag model can be rewritten as a first order system with matrix

$$A = \begin{bmatrix} A_1 & A_2 \cdots A_{k-1} & A_k \\ I & 0 \cdots 0 & 0 \\ 0 & I \cdots 0 & 0 \\ \vdots & \vdots \quad\; \vdots & \vdots \\ 0 & 0 \cdots I & 0 \end{bmatrix}.$$

Thus, to prove result (7), it suffices to show that $\sum_j f(A_j) < 1$ implies $\mu(A) < 1$.

To establish this implication, construct the function $h(A) = g(P^{-1}XP)$, where P is a kXk diagonal matrix, g is the row sum norm, and X is the

kXk matrix gotten by replacing the k^2 blocks of A by their f-norms. By (3) and (5), $h(A)$ is a norm. Since $f(I)=1$, $P^{-1}XP$ writes out as

$$P^{-1}XP=\begin{bmatrix} f(A_1) & f(A_2)p_1/p_2 \cdots & f(A_{k-1})p_1/p_{k-1} & f(A_k)p_1/p_k \\ p_2/p_1 & 0 & \cdots 0 & 0 \\ 0 & p_3/p_2 & \cdots 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots p_k/p_{k-1} & 0 \end{bmatrix}.$$

We are given that $\sum_j f(A_j)<1$. It is possible to choose $p_1>p_2>\cdots> p_k>0$ such that $\sum_j f(A_j)p_1/p_j<1$. With such a choice, it can be seen that the row sums of $P^{-1}XP$ are all less than one. Thus, $h(A)=g(P^{-1}XP)<1$, which implies $\mu(A)<1$, which completes the proof.

A useful way of looking at result (7) is in terms of the scalar distributed lag equation

(8) $$x_t=f(A_1)x_{t-1}+\cdots+f(A_k)x_{t-k}.$$

It is well known that this equation is stable if and only if $f(A_1)+\cdots+ f(A_k)<1$ (Theorem 1 in Bear [1] for example). Hence result (7) tells us that we can establish the stability of the n-equation system by reducing it to a non-negative-coefficient scalar equation (8) and checking the stability of this reduced equation.

Bear [1] has pointed out the importance to economic theory of knowing whether the stability of a model holds up under an equilibrium-invariant retiming of the lags. For example, suppose the system $y_t= b+A_1y_{t-1}+A_2y_{t-2}$ is stable. The question is whether systems like the following, which have the same equilibrium $y_t=(I-A_1-A_2)^{-1}b$, are stable:

$$y_t=(A_1+A_2)y_{t-1},\ y_t=A_2y_{t-1}+A_1y_{t-3},\ y_t=(A_1/2)y_{t-2}+A_2y_{t-4}+(A_1/2)y_{t-6}.$$

In general the answer is no (see Bear). However, if the original system satisfies the stability condition $f(A_1)+f(A_2)<1$, then the answer is yes. This idea develops as a straightforward corollary to (7):

(9) Let $\{E_j\}$ be any set of nXn matrices obeying $\sum_j f(E_j)<1$, where $f(.)$ is a matrix norm with the property $f(I)=1$. Then any distributed lag system $y_t=b+\sum_{i=1}^{k}A_iy_{t-i}$, where the A_i are gotten by summing disjoint subsets of the E_j, is stable.

In practice, economic theory might give us a dynamic model with equilibrium $y_t=(I-\sum_jE_j)^{-1}b$, where the E_j satisfy $\sum_jf(E_j)<1$. Then, as far as stability is concerned, we need not worry about the timing of lags in putting the E_j together into $y_t=b+\sum_iA_iy_{t-i}$; stability is assured.

By combining the analysis of the last two sections, the following structural form version of (7) can be proven straightforwardly.

A sufficient condition for the existence, uniqueness and stability of equilibrium of the system

$$y_t = d + By_t + C_1 y_{t-1} + \cdots + C_k y_{t-k} \tag{10}$$

is $f(B) + f(C_1) + \cdots + f(C_k) < 1$, where $f(.)$ is any matrix norm such that $f(I) = 1$.

The extension of result (9) to this structural form is also straightforward, with B and the C_i (instead of the A_i) formed by summing disjoint subsets of the E_j.

IV. Applications to Periodic Coefficient Models

Consider a system $y_t = b_t + A_t y_{t-1}$, where the coefficient b_t and A_t vary with time subject to $b_t = b_{t-r}$ and $A_t = A_{t-r}$ for all t, where r is a positive integer. The coefficients are said to be *periodic*; the period is r. For example, if t is measured in quarters and $r = 4$, then b_t and A_t may vary from quarter to quarter within a year, but a given quarter has the same coefficients every year. This is a natural way to handle seasonal effects; it corresponds to the use of seasonal dummies in econometric studies. Similarly, t might be in months with $r = 12$. Or the model might be inter-generational (father, son, grandson, etc.) with r the number of sub-periods within an individual generation's lifetime.

Since $b_t = b_{t-r}$ and $A_t = A_{t-r}$, there are only r separate b's and A's—$(b_1, \ldots, b_r)$ and $(A_1, \ldots, A_r)$. Noting this and restating y_t as $y(T, q)$, where T dates the overall period and q the sub-period, allows restatement of $y_t = b_t + A_t y_{t-1}$ as

$$\begin{aligned} y(T, 1) &= b_1 + A_1 y(T-1, r) \\ y(T, 2) &= b_2 + A_2 y(T, 1) \\ &\vdots \\ y(T, r) &= b_r + A_r y(T, r-1). \end{aligned} \tag{11}$$

Making $r-1$ lag substitutions in each of these turns them into

$$\begin{aligned} y(T, 1) &= b_1^0 + A_1^0 y(T-1, 1) \\ &\vdots \\ y(T, r) &= b_r^0 + A_r^0 y(T-1, r), \end{aligned} \tag{12}$$

where the b_j^0 are vector functions of the b_j and A_j, and where

$$\begin{aligned} A_2^0 &= A_1 A_r A_{r-1} \ldots A_3 A_2 \\ A_2^0 &= A_2 A_1 A_r \ldots A_4 A_3 \\ &\vdots \\ A_r^0 &= A_r A_{r-1} \ldots A_2 A_1. \end{aligned} \tag{13}$$

The equations of (12) are separate systems, one for each sub-period. If $\mu(A_j^0) < 1$ for all j, the systems will all have unique stable equilibria; but the equilibria may differ. For example, overall equilibrium in a quarterly model may mean that y_t is the same for a given quarter every year, but different from quarter to quarter within a year. The following theorem suggests a simple norm test of the stability conditions $\mu(A_j^0)<1$ for all j.

> Let $f_1(.),\ldots,f_r(.)$, and $f(.)$ be any matrix norms. A sufficient condition for existence, uniqueness, and stability of the sub-period equilibria of the model of (11)–(13) is
>
> $$f_1(A_1^0),\ldots,f_r(A_r^0)<1, \tag{14}$$
>
> for which it is in turn sufficient that
>
> $$f(A_1)f(A_2)\ldots f(A_r)<1.$$

The first condition follows from (2) applied to (12); and the second follows from the first via axiom (iii).

The model for this last theorem allows only first order lags, which is very restrictive when the time interval is short enough to allow seasonal effects. Thus, consider the more general periodic model

$$\begin{aligned} &y_t=b_t+A_{t1}y_{t-1}+\ldots+A_{tk}y_{t-k}\\ &\text{where } b_t=b_{t-r} \text{ and } A_{tj}=A_{t-r,j} \text{ for all } j. \end{aligned} \tag{15}$$

This model restates as

$$Y_t=B_t+A_tY_{t-1}(B_t=B_{t-r},A_t=A_{t-r}) \tag{16}$$

where $Y_t'=(y_t',\ldots,y_{t-k+1}')$, $B_t'=(b_t',0',\ldots,0')$, and

$$A_t=\begin{bmatrix} A_{t1} & A_{t2} & \ldots & A_{t,k-1} & A_{tk}\\ I & 0 & \ldots & 0 & 0\\ 0 & I & \ldots & 0 & 0\\ \vdots & \vdots & & \vdots & \vdots\\ 0 & 0 & \ldots & I & 0 \end{bmatrix}. \tag{17}$$

Letting $A_1^0,\ldots,A_r^0$ be defined as in (13), the following generalizes the preceding theorem to this distributed lag case.

> Let $f_1(.),\ldots,f_r(.),f(.)$, and $g(.)$ be matrix norms such that $g(I)=1$. A sufficient condition for the existence, uniqueness, and stability of the sub-period equilibria of the model of (15)–(17) is
>
> $$f_1(A_1^0),\ldots,f_r(A_r^0)<1, \tag{18}$$
>
> for which it is in turn sufficient that
>
> $$f(A_1)f(A_2)\ldots f(A_r)<1,$$
>
> for which it is in turn sufficient that
>
> $$[\Sigma_jg(A_{1j})][\Sigma_jg(A_{2j})]\ldots[\Sigma_jg(A_{rj})]<1.$$

For proof, note that the first two conditions follow immediately from the preceding theorem (14) applied to (16). It remains to show that the third condition—which is the easiest to apply since it is stated in terms of the original matrices A_{tj} of (15)—implies the second.

To show this, let $f(A_t)$ in the second condition of (18) be defined by

$$f(A_t)=J\begin{bmatrix} g(A_{t1}) & g(A_{t2})p_1/p_2 & \dots & g(A_{t,k-1})p_1/p_{k-1} & g(A_{tk})p_1/p_k \\ p_2/p_1 & 0 & \dots & 0 & 0 \\ 0 & p_3/p_2 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & p_k/p_{k-1} & 0 \end{bmatrix}$$

for $t=1, \dots, r$. Here $J(.)$ is the row sum norm and the p_j are weights. By (3) and (5), $f(A_t)$ is indeed a norm. In view of the third condition of (18), it is possible to choose $p_1>p_2>\dots>p_k>0$ such that $f(A_1)\dots f(A_r)<1$, which completes the proof.

University of California, San Diego

REFERENCES

[1] Bear, D. V. T., "Distributed Lags and Economic Theory", *Review of Economic Studies*, vol. 33 (1966), pp. 235–43.

[2] Fisher, F. M., "An Alternate Proof and Extension of Solow's Theorem on Non-negative Square Matrices", *Econometrica*, vol. 30 (1962), pp. 349–50.

[3] ——, "Choice of Units, Column Sums and Stability in Linear Dynamic Systems with Non-negative Square Matrices", *Econometrica*, vol. 33 (1965), pp. 445–50.

[4] Lancaster, P., *Theory of Matrices*, New York, 1969.

[5] Samuelson, P., *Foundations of Economic Analysis*, New York, 1965.

[6] Varga, R. S., *Matrix Iterative Analysis*, Englewood Cliffs, 1962.