View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

Comp. & Maths. with Appls., Vol. 5, pp. 141–144
© Pergamon Press Ltd., 1979. Printed in Great Britain
0097-4943/79/0601-0141/$02.00/0

RELIABILITY FOR LINEAR DIFFERENTIAL EQUATIONS WITH NOISY COEFFICIENTS

MICHAEL P. WINDHAM
Mathematics Department, Utah State University, Logan, UT 84322, U.S.A.

Communicated by Robert B. Kelman

(*Received September* 1978; *in revised form October* 1978)

Abstract—Suppose $\dot{\mathbf{X}} = (A + \mathbf{W})\mathbf{X}$ is a system of stochastic differential equations, where A is a matrix of constants and $\mathbf{W}$ is a matrix of white noises. We say the system is reliable if the variance-covariance matrix of the states asymptotically approaches zero. We give conditions in terms of measures of the coefficient matrix and a matrix whose entries are standard deviation parameters of the coefficient noises which will insure that the system is reliable.

The linear constant coefficient system of differential equations is a widely used mathematical model. If a situation can be described in terms of such a system, the solution provides predictions for the future states. For example, Lorenzen *et al.*[3], have proposed a model for the phosphorus concentration in a lake. The model consists of a linear system of differential equations where the states are the phosphorus concentrations of the lake water and the lake bed as functions of time. The predictions provided by the model are predicated on the assumption that the coefficients of the system are constant. These coefficients involve such physical parameters as the flow of water into the lake, and the volume and surface area of the lake. Parameters such as these may be subject to essentially unpredictable flucuations which could cast some doubt on the reliability of the predictions of the model. One approach to dealing with this problem is to introduce a stochastic noise term in the coefficients. The states then become stochastic processes. The expected values of the states can be used as a basis for predictions and their variances as a measure of the reliability of these predictions.

The state variances are functions of the coefficients and their noise variances. One is naturally concerned with how much noise in the coefficients can be tolerated before the state predictions become unreliable, in that the state variances become large. We give here bounds on the variance parameters of the coefficient noises which will insure the state variances asymptotically approach zero.

Suppose we have an n-dimensional system of differential equations

$$\dot{\mathbf{X}} = (A + \mathbf{W})\mathbf{X} \tag{1}$$

where A is a constant matrix and $\mathbf{W}$ is a matrix of white noises satisfying $E(\mathbf{W}) = 0$ and $E(\mathbf{W}_{ik}(t)\mathbf{W}_{jl}(s)) = c_{ikjl}\delta(t-s)$. The c_{ikjl}'s are covariance parameters and c_{ijij} is the variance parameter for the i,jth coefficient noise.

This description is somewhat heuristic, but the system can be reformulated as an Ito equation of the form

$$\mathrm{d}\mathbf{X} = f(\mathbf{X}, t)\,\mathrm{d}t + g(\mathbf{X}, t)\,\mathrm{d}\mathbf{B}\,. \tag{2}$$

For our case, $f(\mathbf{X}, t) = A\mathbf{X}$, $g(\mathbf{X}, t)$ is an $n \times n^2$ matrix whose diagonal entries are $\mathbf{X}^T$ and off diagonal entries are zero matrices. Also, $\mathbf{B}$ is an n^2-dimensional vector of second order stochastic processes with independent increments. For $\Delta\mathbf{B} = \bar{\mathbf{B}}(t + \Delta t) - \mathbf{B}(t)$ we assume that $E(\Delta\mathbf{B}) = 0$ and $E(\Delta\mathbf{B}\Delta\mathbf{B}^T) = C\Delta t$ for a symmetric matrix, C. The relationship between this formulation and equation (1) is that one thinks of $\mathrm{d}\mathbf{B}$ as being $[\mathbf{W}_{11} \ldots \mathbf{W}_{1n} \ldots \mathbf{W}_{21} \ldots \mathbf{W}_{2n} \ldots \mathbf{W}_{nn}]$. Also, the entries of C are the c_{ikjl}'s.

A second order, Markov solution exists and is unique for a given initial condition which is independent of the increments of $\mathbf{B}$, ([2], Theorem 4.5; or [4], Theorem 5.2.4). Furthermore, if h is a function from $\mathbf{R}^{n+1}$ to $\mathbf{R}$ which has continuous second derivatives and $\mathbf{X}(t)$ is a solution to

(2), then

$$\frac{\mathrm{d}}{\mathrm{d}t}E(h)\mathbf{X},t)) = \Sigma E\left(f_k\frac{\partial h}{\partial x_k}\right) + E\left(\frac{\partial h}{\partial t}\right) + \frac{1}{2}\Sigma E\left((gCg^T)_{kl}\frac{\partial^2 h}{\partial x_k \partial x_l}\right)$$

(see [5], Section 7.4.1 (c)). By applying this result to appropriate choices of h we obtain differential equations for the mean and second moments of X. In particular, if we let $x(t) = E(\mathbf{X}(t))$ and $S(t) = E(\mathbf{X}\mathbf{X}^T)$ we have

$$\dot{x} = Ax \tag{3}$$

$$\dot{S} = AS + SA^T + [\Sigma_{k,l}c_{ikjl}S_{kl}]. \tag{4}$$

When the variance-covariance matrix for X approaches zero asymptotically, we say the system $\dot{\mathbf{X}} = (A + \mathbf{W})\mathbf{X}$ is *reliable.* Since this matrix is given by $S - xx^T$, in order to insure that the system is reliable it suffices to provide conditions which imply that the zero solutions to equations (3) and (4) are asymptotically stable.

Both systems are linear with constant coefficients. In fact, (4) can be written as an n^2-dimensional system whose coefficient matrix has entries of the form $a_{ik}\delta_{jl} + \delta_{ik}a_{jl} + c_{ikjl}$. A particular row of this matrix corresponds to a fixed value for i and j and a column to fixed values of k and l. A sufficient condition for the reliability of the system is that all the eigenvalues of the coefficient matrices in (3) and (4) have negative real parts. The matrix A is assumed to be known, so, obtaining its eigenvalues would be a reasonable approach to dealing with (3). However, this may involve a great deal of computation, with its associated numerical difficulties. For (4), the problem is compounded by the fact that the coefficient matrix is $n^2 \times n^2$. This size could be reduced to $n(n+1)/2$, since S is symmetric. In any case, finding eigenvalues for the coefficient matrix in (4) assumes that the c_{ikjl}'s are known. This does not allow one to establish, in advance, tolerance levels for the noise which will insure reliable predictions by the model.

An alternate approach is provided by using matrix measurements. We define the *measure* of an $n \times n$ matrix, A, by

$$\mu(A) = \sup\{\mathrm{Re}\,(a_{ii}) + \Sigma_{k\neq i}|a_{ik}|\}\,.$$

Letting $|x| = \sup|x_i|$ for a vector x in $\mathbf{C}^n$, the measure of A satisfies

$$\lim_{h\to 0^+}(|x + hAx| - |x|)/h \leq \mu(A)|x|. \tag{5}$$

Furthermore, if x is an eigenvector of A with eigenvalue, λ, then the limit is equal to $\mathrm{Re}\,(\lambda)|x|$. So, the measure of a matrix is at least as large as the real part of any of its eigenvalues. This means that solutions to a system of differential equations are asymptotically stable if the measure of the coefficient matrix is negative.

This measure is one of a family of measures which have found wide application in the study of stability. For any norm on $\mathbf{C}^n$, an associated matrix measure can be defined to be the smallest number satisfying the inequality (5) for all vectors in $\mathbf{C}^n$ (see [1], Chap. 2). We have chosen the measure induced by the sup-norm because it can be computed by inspection and provides significant simplification in dealing with noise covariances. In fact, our first result involves only $\mu(A)$ and $\mu(D)$ where D is the $n \times n$ matrix whose entries are $d_{ij} = (c_{ijij})^{1/2}$. The entries of D are standard deviation parameters. Furthermore, we have $|c_{ikjl}| \leq d_{ik}d_{jl}$.

THEOREM 1. The system $\dot{\mathbf{X}} = (A + \mathbf{W})\mathbf{X}$ is reliable if $\mu(A) < 0$ and $\mu(D) < (-2\mu(A))^{1/2}$.

Proof. In view of our previous remarks, all that needs to be shown is that the coefficient matrix for the linear system equivalent to (4) has negative measure, i.e. for

$$R_{ij} = a_{ii} + a_{jj} + c_{iijj} + \Sigma_{k\neq i}|a_{ik} + c_{ikjj}| + \Sigma_{l\neq j}|a_{jl} + c_{iijl}| + \Sigma_{k\neq i,l\neq j}|c_{ikjl}|$$

we have $R_{ij} < 0$. But,

$$R_{ij} \le a_{ii} + \Sigma_{k\ne i}|a_{ik}| + a_{jj} + \Sigma_{l\ne j}|a_{jl}| + \Sigma_{k,l}|c_{ikjl}| \le 2\mu(A) + \Sigma|c_{ikjl}|$$

$$\le 2\mu(A) + (\Sigma d_{ik})(\Sigma d_{jl}) \le 2\mu(A) + (\mu(D))^2$$

and the result follows.

These estimates are somewhat crude, in that the measure of a matrix may be positive, but the real parts of all its eigenvalues are negative. For the measure to be negative, the diagonal entries must be negative and small enough to dominate the absolute values of the other entries in the corresponding rows. For this reason, one might say that the theorem is useful only in systems where each state is the main contributor to its own decay. On the other hand, the estimates are very simple to compute and have meaningful physical interpretation, in that the entries of D are the standard deviation of the noise for corresponding entries in A.

If one is willing to tolerate more computation then it is possible to avoid the problem caused by A having positive measure by transforming the system. If we let $\mathbf{Y} = B\mathbf{X}$ for a non-singular matrix. B, then $\mathbf{Y}$ satisfies $\dot{\mathbf{Y}} = (BAB^{-1} + B\mathbf{W}B^{-1})\mathbf{Y}$ and $E(\mathbf{Y}\mathbf{Y}^T) = BSB^T$. Therefore, if the system for Y is reliable, then so is the system for $\mathbf{X}$. We denote by D_B the matrix of deviations for the noise matrix $B\mathbf{W}B^{-1}$.

THEOREM 2. If there is a non-singular matrix B so that $\mu(BAB^{-1}) < 0$ and $\mu(D_B) < (-2\mu(BAB^{-1}))^{1/2}$ then the system $\dot{\mathbf{X}} = (A + \mathbf{W})\mathbf{X}$ is reliable.

The new states, $\mathbf{Y}$, and the matrix, D_B, may have no meaningful physical interpretation. In particular, the second inequality in Theorem 2 may be difficult to interpret in terms of tolerance levels for the original system noises. However, information about the original noise deviations, D, can be recovered because

$$\mu(D_B) \le M(B)\mu(D)$$

where $M(B)$ is the condition number of the matrix B. The condition number is defined by $M(B) = \sup\{\Sigma_{k,l}|b_{ik}||b^{il}|\}$, where b_{ij} and b^{ij} are the entries of B and B^{-1}, resp. We can then replace $\mu(D_B)$ in Theorem 2 by $M(B)\mu(D)$ and the conclusion still follows.

If $\mu(BAB^{-1}) < 0$ for some B then the real parts of all the eigenvalues of A must be negative. It follows from the proof of the next theorem that the converse is also true.

If we let $l = -\sup \text{Re}(\lambda)$ for λ an eigenvalue of A then $-\mu(BAB^{-1}) \le l$ for any B. In order to have a generous estimate for $\mu(D)$, it is reasonable to choose B to make $-\mu(BAB^{-1})$ as close to l as possible; this means making BAB^{-1} as close to diagonal as possible. For this reason, we choose B to be a matrix which transforms A to Jordan cannonical form, $\Lambda + H$, where Λ is diagonal and H has ones and zeros on the superdiagonal and zeroes elsewhere. Let $k + 1$ denote the size of the largest Jordan block in $\Lambda + H$, or in other words, k is the length of the longest string of consecutive ones in H. If A is diagonal, then $k = 0$, and conversely. Finally, if $k > 0$, then let $L = -\sup \text{Re}(\lambda)$ for λ an eigenvalue with at least one non-zero entry on the superdiagonal of one of its corresponding Jordan blocks. We can then give tolerance levels for the noise matrix D_B in terms of the numbers, k, l and L.

The nature of the estimates depends on relationships between k, l and L which we distinguish as follows:

Case I. $k = 0$ or $L \ge l + 1$.

Case II. $k > 0$ and $l + 1 > L \ge (2k + 1)/2k$.

Case III. $k > 0$ and $l + 1 > L$ and $(2k + 1)/2k > L \ge (2k + 1)l$.

Case IV. $k > 0$ and $l + 1 > L$ and $(2k + 1)/2k > L$ and $(2k + 1)l > L$.

THEOREM 3. If $l > 0$ then the inequalities below are sufficient conditions for $\dot{\mathbf{X}} = (A + \mathbf{W})\mathbf{X}$ to be reliable for the case indicated.

Case I. $\mu(D_B) < (2l)^{1/2}$.

Case II. $\mu(D_B) < (2(L - 1))^{1/2}$.

Case III. $\mu(D_B) < (2l)^{1/2}(L-l)^k$.
Case IV. $\mu(D_B) < (2L)^{1/2}(2kL)^k/(2k+1)^{k+1/2}$.

Proof. The assumption, $l > 0$, implies that the eigenvalues of A have negative real parts and that solutions to $\dot{x} = Ax$ are asymptotically stable.

If $k = 0$, then A is diagonalizable, i.e. $BAB^{-1} = \Lambda$, so $\mu(BAB^{-1}) = \mu(\Lambda) = -l$ and this part of Case I follows immediately from Theorem 2. For the remainder we have $k > 0$ so that $H \neq 0$.

For $r > 0$, we construct a diagonal matrix, R, whose diagonal entries are powers of r, determined by the structure of $\Lambda + H$. For an $m \times m$ Jordan block in $\Lambda + H$, the $m \times m$ diagonal block in the same location in R is constructed as follows. For $m = 1$ the entry is 1; for $m > 1$ the ith diagonal entry is r^{m-i}. Since $r > 0$, R is non-singular and the largest power of r that appears is r^k. If we transform $\dot{\mathbf{X}} = (A + \mathbf{W})\mathbf{X}$ by RB, the coefficient matrix is $\Lambda + rH$, and for $r < L$ we have $\mu(\Lambda + rH) < 0$. Moreover, $\mu(D_{RB}) \leq \mu(D_B) \max(r^k, r^{-k})$. So if, for $0 < r < L$,

$$\mu(D_B) < (-2\mu(\Lambda + rH))^{1/2} \min(r^k, r^{-k}) \tag{6}$$

then $\dot{\mathbf{X}} = (A + \mathbf{W})\mathbf{X}$ is reliable. But, the quantity on the right in (6) is a continuous, non-negative function for $0 \leq r \leq L$, and vanishes for $r = 0$ or $r = L$. Furthermore, it has at most three local maxima which, if they occur, are at $r = L - l$, 1 and $2kL/(2k+1)$. For each of the cases described in the theorem, the right hand side of the inequality is the global maximum for this function for $0 < r < L$ under the condition described in that case; so each is a special case of (6).

As was the case with Theorem 2, if $\mu(D_B)$ is replaced by $M(B)\mu(D)$ the new estimates also imply that the system is reliable.

In applications to modeling, the situation which occurs most often is that the eigenvalues of A are distinct. In this case A is diagonalizable and Theorem 3 assures the best possible tolerance level, $(2l)^{1/2}$.

REFERENCES

1. W. A. Coppel, *Stability and Asymptotic Behaviour of Differential Equations.* D. C. Heath, Boston (1965).
2. A. H. Jazwinski, *Stochastic Processes and Filtering Theory.* Academic Press, New York (1970).
3. M. W. Lorenzen, D. J. Smith and L. V. Kimmel, A Long-Term Phosphorus Model for Lakes. *Modeling Biochemical Processes in Aquatic Ecosystems* (Edited by R. P. Canale), pp. 75–92. Ann Arbor Science Publishers, Ann Arbor, Michigan (1976).
4. T. T. Soong, *Random Differential Equations in Science and Engineering.* Academic Press, New York (1973).